UCLA School of Law

Public Law & Legal Theory Research Paper No. 22-23

UNFAIR ARTIFICIAL INTELLIGENCE: HOW FTC INTERVENTION CAN OVERCOME THE LIMITATIONS OF DISCRIMINATION LAW

BY

ANDREW D. SELBST
ASSISTANT PROFESSOR OF LAW

AND

SOLON BOROCAS
MICROSOFT RESEARCH
CORNELL UNIVERSITY

UNFAIR ARTIFICIAL INTELLIGENCE: HOW FTC INTERVENTION CAN OVERCOME THE LIMITATIONS OF DISCRIMINATION LAW

Andrew D. Selbst[*]
Solon Barocas[**]

The Federal Trade Commission has indicated that it intends to regulate discriminatory AI products and services. This is a welcome development, but its true significance has not been appreciated to date. This Article argues that the FTC's flexible authority to regulate "unfair and deceptive acts and practices" offers several distinct advantages over traditional discrimination law when applied to AI. The Commission can reach a wider range of commercial domains, a larger set of possible actors, a more diverse set of harms, and a broader set of business practices than are currently covered or recognized by discrimination law. For example, while most discrimination laws can address neither vendors that sell discriminatory software to decision makers nor consumer products that work less well for certain demographic groups than others, the Commission could address both. The Commission's investigative and enforcement powers can also overcome many of the practical and legal challenges that have limited plaintiffs' ability to successfully seek remedies under discrimination law. The Article demonstrates that the FTC has the existing authority to address the harms of discriminatory AI and offers a method for the Commission to tackle the problem, based on its existing approach to data security.

---

[*] Assistant Professor, UCLA School of Law.

[**] Principal Researcher, Microsoft Research; Adjunct Assistant Professor, Information Science, Cornell University. The authors thank Maria Brincker, Lizzie Brown, Woody Hartzog, Pauline Kim, Chris Morten, and the participants of the 2022 Privacy Law Scholars Conference and Electronic Privacy Information Center AI Policy Symposium for comments and feedback. Thanks also to Ellie Baines for excellent research assistance.

INTRODUCTION

Discriminatory artificial intelligence is unfair.[1] Yet, until very recently, the Federal Trade Commission, the United States federal agency charged with regulating "unfair and deceptive acts and practices" in commerce,[2] has not sought to address harms from discriminatory AI. Since the confirmation of Chair Lina Khan, however, the FTC has demonstrated a willingness to take on a much more aggressive enforcement stance than at any time in the last 40 years, and that includes a clear interest in using its existing authorities to rein in discriminatory AI.[3]

In this Article, we argue that FTC intervention in this space is a positive and overdue development. The Commission can do a lot of good by

[1] Certainly, discrimination would seem to fit under any commonly understood definition of unfairness. *See Unfair*, MERRIAM-WEBSTER ONLINE (defining unfair as either "marked by injustice, partiality, or deception," or "not equitable in business dealings); STEPHEN HAYES & KALI SCHELLENBERG, DISCRIMINATION IS "UNFAIR": INTERPRETING UDA(A)P TO PROHIBIT DISCRIMINATION 14 (2021), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3832022 (pointing out that discrimination and unfairness are often treated as synonyms).

[2] 45 U.S.C. §15(a).

[3] *See* Elisa Jillson, *Aiming for Truth, Fairness, and Equity in your Company's Use of AI*, FTC BUSINESS BLOG (April 19, 2021), https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai (blog post by FTC attorney stating that the FTC will address algorithmic discrimination); Samuel Levine, Keynote, Cleveland-Marshall College of Law Cybersecurity and Privacy Protection Conference 3, https://www.ftc.gov/system/files/ftc_gov/pdf/Remarks-Samuel-Levine-Cleveland-Marshall-College-of-Law.pdf (Director of the Consumer Protection Bureau stating discrimination that results from surveillance is a top concern of the FTC); Rebecca Kelly Slaughter, Janice Kopec & Mohamad Batal, *Algorithms and Economic Justice: A Taxonomy of Harms and a Path Forward for the Federal Trade Commission*, 23 YALE J.L. & TECH 1, 40–41 (2021) (An article by FTC Commissioner Slaughter arguing that "[t]he FTC can and should be aggressive in its use of unfairness to target conduct that harms consumers [including] discrimination, as well as . . . other algorithmic harms.").

applying its authority to address unfair and deceptive acts and practices to discriminatory AI.[4] Surprisingly, though the discriminatory harms of commercial AI have been frequently discussed in the last decade of legal literature[5] and scholars have occasionally suggested a possible role for the FTC,[6] there has never been an exploration of the unique benefits of the Commission's involvement in algorithmic discrimination or a full treatment of the legal authority that would allow the Commission to address it. We provide that exploration and treatment here.

The FTC's consumer protection authority is most useful when the Commission can fill gaps in existing legal regimes. The Commission's flexible authority allows it to be nimble and adjust to new developments and changing

[4] 15. U.S.C. 45(a).

[5] *See generally* Ifeoma Ajunwa, *An Auditing Imperative for Automated Hiring Systems*, 34 HARV. J.L. & TECH. 621 (2021); Thomas B. Nachbar, *Algorithmic Fairness, Algorithmic Discrimination*, 48 FLA. ST. U. L. REV. 509 (2021); Andrés Páez, *Negligent Algorithmic Discrimination*, LAW & CONTEMP. PROBS., 2021, at 19; Sandra Wachter, Brent Mittelstadt & Chris Russell, *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*, 41 COMPUTER L. & SECURITY REV. 105567 (2021); Jason R. Bent, *Is Algorithmic Affirmative Action Legal?,* 108 Geo L.J. 830 (2020); Deborah Hellman, *Measuring Algorithmic Fairness*, 106 VA. L. REV. 811 (2020); Aziz Z. Huq, *A Right to a Human Decision,* 106 VA. L. REV. 611 (2020); Anya E.R. Prince & Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data,* 105 IOWA L. REV. (2020); Ifeoma Ajunwa, *Age Discrimination by Platforms*, 40 BERKELEY J. EMP. & LAB. L. 1 (2019); Ifeoma Ajunwa, *The Paradox of Automation as Anti-Bias Intervention*, 41 CARDOZO L. REV. 1671 (2019) [hereinafter, Ajunwa, *Paradox*]; Sonia K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54 (2019); Sandra G. Mayson, *Bias In, Bias Out,* 128 YALE L.J. 2218 (2019); Stephanie Bornstein, *Antidiscriminatory Algorithms*, 70 ALA. L. REV. 519 (2018); Stephanie Bornstein, *Reckless Discrimination*, 105 CALIF. L. REV. 1055 (2017); Anupam Chander, *The Racist Algorithm?,* 115 MICH. L. REV. 1023 (2017); Pauline T. Kim, *Data-Driven Discrimination at Work,* 58 WM. & MARY L. REV. 857 (2017); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning,* 51 U.C. DAVIS L. REV. 653 (2017); Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions,*89 WASH. L. REV. 1 (2014); Tal Z. Zarsky, *Understanding Discrimination in the Scored Society*, 89 WASH. L. REV. 1375 (2014).

[6] Dennis D. Hirsch, *From Individual Control to Social Protection: New Paradigms for Privacy Law in the Age of Predictive Analytics*, 79 MD. L. REV. 439, 491–92 (2020) [hereinafter Hirsch, *New Paradigms*]; Dennis D. Hirsch, *That's Unfair! Or Is It? Big Data, Discrimination and the FTC's Unfairness Authority*, 103 KY. L.J. 345, 354 (2015) [hereinafter, Hirsch, *That's Unfair!*]; Michael Spiro, *The FTC and AI Governance: A Regulatory Proposal*, 10 SEATTLE J. TECH. ENV. & INNOVATION L. 26, 52 (2020); Lauren E. Willis, *Deception by Design*, 34 HARV. J.L. & TECH. 115, 177 (2020); Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785, 822 (2015).

circumstances.[7] This is why the FTC has always been a central regulator of new technological development[8] and has become the de facto data protection regulator in the United States[9]—the Commission reacts and fills the gaps in existing law. But this presents a puzzle: Why does the FTC want to get involved in regulating discriminatory AI when we already have an extensive list of civil rights laws? And why is that a good idea?

As with other instances of technological development, the landscape of decisionmaking has changed. Where discrimination laws were designed to prevent harmful decisions by individuals or groups of people, today many of these decisions are technologically mediated and infinitely more complex, with parts of decisions delegated to AI. There are some aspects of this that traditional civil rights laws can likely address, but a new technological reality means new gaps to fill. Indeed, while public statements by the FTC suggest that its initial focus is on regulation of those activities that can already be addressed by discrimination law,[10] there are also hints that it aims to go further. In Commissioner Slaughter's words:

> Civil rights laws are the logical starting point for addressing discriminatory consequences of algorithmic decision-making, [but] in many cases, existing civil-rights jurisprudence may be difficult to apply to algorithmic bias . . . . So, we must consider what other legal protections currently exist outside of direct

[7] Atl. Ref. Co. v. Fed. Trade Comm'n, 381 U.S. 357, 367 (1965) ("The Congress intentionally left development of the term 'unfair' to the Commission rather than attempting to define 'the many and variable unfair practices which prevail in commerce."); Fed. Trade Comm'n v. Bunte Bros., 312 U.S. 349, 353 (1941) (" '[U]nfair competition' was designed by Congress as a flexible concept with evolving content.")

[8] CHRIS JAY HOOFNAGLE, FEDERAL TRADE COMMISSION PRIVACY LAW AND POLICY 25–26 (2015).

[9] Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 600 (2014); Steven Hetcher, *The De Facto Federal Privacy Commission*, 19 J. MARSHALL J. COMPUTER & INFO. L. 109 (2000).

[10] For example, the Commission's 2021 blog post addressing issues of unfairness, Jillson, *supra* note 3, notably focuses on cases of algorithmic bias in the context of employment, credit, housing, and healthcare—all domains subject to laws that seek to guard against discrimination. *See* 42 U.S.C. § 2000e; 15 U.S.C. § 1691; 42 U.S.C. § 1301; 42 U.S.C. §§ 2000d–d-7. While the blog post also touches on advertising, it focuses specifically on ads related to these regulated domains (e.g., ads for jobs), which are themselves covered by discrimination law. *See* Amit Datta, Anupam Datta, Jael Makagon, Deirdre K. Mulligan & Michael Carl Tschantz, *Discrimination in Online Advertising: A Multidisciplinary Inquiry*, 81 Proceedings of Machine Learning Research 1.

civil rights statutes.[11]

Thinking about discrimination more broadly than the civil rights statutes—as unfairness—confers several advantages that have so far been overlooked. It allows the Commission to regulate commercial domains, actors, injuries, and business practices that cannot otherwise be addressed. These benefits may not be obvious because we typically think of discrimination as a separate problem from consumer protection, but where a large commercial industry creates products that discriminate or enable discrimination, the two merge, necessitating new ways to think about the problem. And though it might on its face seem incongruous to address discrimination with consumer protection authority, the FTC's actual charge is to regulate commerce broadly,[12] so there is no reason that the FTC's authority could not apply.[13]

To be sure, the problem of discriminatory AI is multifaceted, and we do not argue here that FTC oversight or the common law approach is the best or only way to address it. We have in the past argued for application of and revisions to discrimination law,[14] the use of additional procedural protections,[15] the adoption of documentation standards,[16] and the introduction of impact assessment requirements.[17] Instead, we argue that the FTC can play a unique—and so far overlooked—role in the larger set of regulatory responses that are necessary to rein in discriminatory AI and that the Commission should focus its efforts there.

The Article proceeds in three Parts. In Part I, we identify five advantages of FTC intervention. First is its ability to address activities outside

---

[11] Slaughter, et al., *supra* note 3, at 38.

[12] *See* 15 U.S.C. 45(a) ("[U]nfair or deceptive acts or practices in or affecting commerce, are hereby declared unlawful.) "Consumer" is not defined in the statute itself, but consumers are simply individual people acting in their commercial capacity. *See* Meg Leta Jones, *The Characters of Consent* (forthcoming) (discussing information law's various descriptions of people as "users," "consumers," and "data subjects").

[13] *See e.g.*, In the Matter of Napleton Automotive Group Commission, File No. 2023195, 2022 WL 1039797.

[14] *See generally* Barocas & Selbst, *supra* note 5.

[15] Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson, & Harlan Yu, *Accountable Algorithms*, 165 U. PA. L. REV. 633 (2017).

[16] Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1129 (2018).

[17] Andrew D. Selbst, *An Institutional View of Algorithmic Impact Assessments*, 35 HARV. J. L. & TECH. 118 (2021).

domains explicitly regulated by discrimination law, such as the development and sale of technology that enables downstream discrimination or consumer products that perform systematically worse for certain demographic groups than others. Second is the ability to reach certain actors that discrimination law often cannot—often the developers of technology that will later be used to discriminate. Third is the ability to recognize and address different types of discriminatory harm than discrimination law's primarily procedural definitions allow. We identify three different types of harm—allocative (the traditional concern of most discrimination law), quality-of-service (a classic consumer protection harm, in this case experienced more by certain demographic groups than others), and representational harms (depictions of demographic groups that harm their social standing overall), though the FTC's ability to address representational harm is debatable. Fourth is the FTC's ability to address business practices that are not about discrimination directly but could make discrimination more likely to occur or more difficult to address, such as failure to evaluate products and services for discrimination and disclose those results or take appropriate remedial actions. Fifth is the FTC's advantages as a litigant, which inhere in its status as an enforcement agency rather than a discrimination plaintiff or plaintiff class.

Part II addresses the Commission's legal authority. The FTC Act defines unfair acts and practices with a three-part test: (1) a likelihood of substantial injury to consumers; (2) the injury is not reasonably avoidable by consumers themselves; and (3) it is not outweighed by countervailing benefits to consumers.[18] Generally, to be unfair, something need not be otherwise illegal, but the Commission may take inspiration from established public policies, and for the most part, it will not be difficult to find that discriminatory AI constitutes an unfair business practice.[19] Part II also addresses deception and the potential for challenge under the "major questions doctrine."

Part III examines one way that the FTC could go about this work. Over the last two decades, the Commission has addressed data security by developing a program of enforcement that amounts to a pseudo-common law approach, by issuing guidance and settling enforcement actions with public consent decrees that demonstrate the best and worst practices respectively. There are several interesting parallels between data security harms and discriminatory AI that lead us to think that the data security model could be

[18] 15 U.S.C. § 45(n).
[19] *Id.*

a good one to build on. The other regulatory model that the FTC could pursue—which it seems likely to—is the creation of "trade regulation rules" to define affirmatively what practices are unfair.[20] We focus on the common law approach because of the interesting parallels and because it is an established model that would probably be the easiest for the FTC to replicate, but we briefly discuss the rulemaking option as well.

I. THE BENEFITS OF FTC INTERVENTION

There are several reasons that FTC intervention into the regulation of discriminatory AI would be a positive development. Currently, the regulatory scheme that is most directly applicable to discriminatory AI is the set of civil rights laws, such as Title VII, the Equal Credit Opportunity Act (ECOA), and the Fair Housing Act (FHA). Algorithmic an AI-based decision-making poses challenges to enforcement of those laws in ways that have been documented at length in other scholarship.[21] Our interest here is slightly different. The FTC's Section 5 authority will indeed allow it to replicate some of the successes of those civil rights laws, but more importantly, the FTC's reach is broader. In this Part we identify five benefits to FTC intervention that can help it reach beyond the structural and procedural limitations of discrimination law as applied to AI.

A. Discrimination Domains

One major benefit of FTC intervention to address discriminatory AI is the ability to reach domains other than those already regulated by discrimination law. Most discrimination law applies to specific domains such as employment,[22] credit,[23] and housing,[24] and it is focused on concrete decisions within those domains, such as whether to hire, lend, or rent a home. Many commercial applications of AI take place in domains outside those regulated by discrimination law or upstream from the regulated decision, yet nonetheless create discriminatory harms. As a result, troubling instances of discriminatory AI often escape the reach of discrimination law, including

[20] *See* 15 U.S.C. § 57a.
[21] *See* sources cited in note 5, *supra*.
[22] 42 U.S.C. § 2000e.
[23] 15 U.S.C. § 1691.
[24] 42 U.S.C. § 1301.

those that have been the subject of some of the most well-known studies of algorithmic bias.

Consider the famous Gender Shades study by Joy Buolamwini and Timnit Gebru, which demonstrated that many commercially available gender classification tools exhibited significant performance disparities by skin tone and gender, performing especially poorly for darker-skinned women.[25] Gender classification tools are unlikely to be adopted in domains regulated by discrimination laws because these laws forbid disparate treatment on the basis of gender.[26] This is cold comfort because there remain many other potential applications of gender classification and related computer vision tools outside these domains where such biases would be cause for serious concern. Buolamwini herself has pointed out that "[t]he same data-centric techniques that can be used to try to determine somebody's gender are also used . . . to unlock your phone."[27] There is no discrimination law that would hold phone manufacturers liable if the facial recognition that is commonly used to unlock users' devices demonstrates systematic differences in performance across different demographic groups. The same holds for the automated speech recognition that powers the virtual assistants on many phones (e.g., Apple's Siri, Google Assistant, etc.), which has likewise been shown to exhibit disparities in accuracy by race.[28] These activities are, however, well within the scope of the FTC's authority. The Commission can reach AI-based products and services targeted at everyday consumers, not just those targeted at the decision makers in the set of domains regulated by discrimination law.

---

[25] Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*. PROCEEDINGS OF THE 2018 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 77.

[26] There are, of course, many efforts to use computer vision in decision-making in regulated domains, especially employment. *See* Luke Stark & Jevan Hutson, *Physiognomic Artificial Intelligence*, 32 FORDHAM INTELL. PROP. MEDIA & ENT. L.J., 922 (2022); Ifeoma Ajunwa, *Automated Video Interviewing as the New Phrenology*, 36 BERKELEY TECH. L.J. __ (forthcoming 2022),

[27] Larry Hardesty, *Study Finds Gender and Skin-Type Bias in Commercial Artificial-Intelligence Systems,* MIT NEWS (Feb. 11, 2018), https://news.mit.edu/2018/study-finds-gender-skin-type-bias-artificial-intelligence-systems-0212. Another common application is policing, *see generally* CLAIRE GARVIE, ALVARO BEDOYA & JONATHAN FRANKLE, THE PERPETUAL LINEUP (2016), where there is no law to regulate discriminatory AI. *See* Andrew D. Selbst, *Disparate Impact in Big Data* Policing, 52 GA. L. REV. 109, 144 (2017).

[28] Allison Koenecke, et al., *Racial Disparities in Automated Speech Recognition*, 117 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, 7684 (2020).

Consider another canonical example from the research on algorithmic bias. Several early studies showed that natural language processing (NLP) tools can encode a range of troubling biases.[29] These studies revealed that NLP tools using "word embeddings"—a commonly used mathematical representation of how words relate to each other—would learn stereotypical associations between gender and occupation, as highlighted in the very name of one of the papers: "Man is to Computer Programmer as Woman is to Homemaker."[30] Yet again, tools that exhibit such properties remain outside the scope of discrimination law unless they are adopted for decision making in a regulated domain. Only once word embeddings have been incorporated into a particular task like ranking job applicants according to the content of their resumes would such tools implicate discrimination law. In contrast, the FTC has the authority to go after general-purpose AI tools before they have been incorporated into a decision-making process in a regulated domain. The Commission can reach these tools, outside of specified domains, if there is good reason to believe that they will cause harm in the future.[31] This should be an easy argument to make when these general-purpose tools are explicitly marketed to decision-makers in the domains regulated by discrimination law and where the marketing stresses the utility of the tools for making such decisions.[32]

B. Possible Defendants

Just as discrimination law is limited in the range of domains that it can reach, so, too, is it limited in the range of actors that it can hold liable. These are related points—if the focus of discrimination law is on decision points, then the main actors regulated will be the decision makers. And just as the FTC is well positioned to overcome the limits of regulated domains, so, too, is it well positioned to overcome the limits of regulated actors.

---

[29] Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, and Adam T. Kalai, *Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings*, 29 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 1 (2016); Aylin Caliskan, Joanna J. Bryson & Arvind Narayanan, *Semantics Derived Automatically from Language Corpora Contain Human-like Biases* 356 SCI. 183 (2017).

[30] Bolukbasi, et al., *supra* note 29.

[31] 15 U.S.C. § 45(n).

[32] *See, e.g.,* Manish Raghavan, Solon Barocas, Jon Kleinberg & Karen Levy, *Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices*, PROCEEDINGS OF THE 2020 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FACCT) 469.

To begin, consider the range of actors subject to Title VII, ECOA, and FHA. Title VII applies to employers, employment agencies, labor organizations, or joint labor-management committees overseeing training programs.[33] It does not extend to the various entities that might provide support to these actors, such as vendors of employment assessments.[34] Similarly, ECOA's list of defendants is limited to creditors, defined as "any person who regularly extends, renews, or continues credit; any person who regularly arranges for the extension, renewal, or continuation of credit; or any assignee of an original creditor who participates in the decision to extend, renew, or continue credit."[35] ECOA does not appear to apply to actors that provide information informing these decisions, such as vendors of credit scores,[36] despite Congress's and federal agencies' long-standing interest in uncovering potential racial disparities in credit scores.[37]

The FHA is more flexible than Title VII and ECOA. With two exceptions, the law does not explicitly define a set of covered entities.[38]

---

[33] 42 U.S.C. § 2000e-2(a)-(d).

[34] *See* Stewart v. Hannon, 469 F. Supp. 1142, 1148 n.3 (N.D. Ill. 1979) ("Plaintiffs have conceded that ETS cannot properly be charged with a Title VII violation because it is neither plaintiffs' employer, an employment agency, nor a labor organization, and accordingly, it is not covered by the substantive provisions of Title VII."); Though while federal law does not contain any prohibitions on aiding and abetting employment discrimination, some states do. Datta, et al, *supra* note 10, at 9 ("Several states including California, New York and Pennsylvania, prohibit any person from aiding, abetting, inciting, compelling, or coercing discriminatory employment practices.").

[35] 15 U.S.C. § 1691a(f).

[36] *See* Hilton v. Fair Isaacs, Inc., 2005 WL 8177639, at *3 (N.D. Cal. 2005) ("Hilton's complaint against Fair Isaac, the County, and Buffington fails to state an ECOA claim […] Hilton does not allege that Fair Isaac, the County, and Buffington are creditors for purposes of 15 U.S.C. § 1691."); *see also* Arikat v. JP Morgan Chase & Co., 430 F.Supp.2d 1013, 1020 n.9 (2006)("[P]laintiffs conceded their claims under the ECOA as to Fair Isaac because Fair Isaac 'does not act as a direct creditor with respect to plaintiffs.'").

[37] Board of Governors of the Federal Reserve System, Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit (Aug. 2007); Federal Trade Commission. Credit-Based Insurance Scores: Impacts on Consumers of Automobile Insurance (Jul. 2007).

[38] ROBERT G. SCHWEMM, HOUSING DISCRIMINATION LAW AND LITIGATION § 12B:1 Who may be liable under the Fair Housing Act ("The statute makes little effort to define the scope of proper defendants. With the exception of § 3605's ban of discrimination in certain real estate-related transactions and § 3608's affirmative command to federal agencies, the substantive provisions of the statute (§§ 3604, 3605, 3606, 3617) simply declare certain housing practices to be unlawful without specifying who may be held responsible for these practices." (parenthetical cross-references deleted)).

Instead, the FHA describes a range of actions that fall under its purview, regardless of the specific actor who might be undertaking them. In a notable departure from Title VII and ECOA, FHA has already been read to extend to vendors of tenant screening tools. In *Connecticut Fair Housing Center v. Corelogic Rental Property Solutions*, plaintiffs sued a vendor, Corelogic, claiming that the company "had a duty not to sell a product to [a landlord] which would unwittingly cause [the landlord] to violate federal housing law and regulations".[39] In reaching this decision, the court stressed the close nexus between the vendors' conduct and the denial of housing, noting that landlords defer critical aspects of their decision-making process to the vendor, even if the ultimate decision regarding a prospective tenant rests with the landlords.[40]

The more expansive scope of the FHA was also apparent in two high-profile lawsuits against Facebook. *National Fair Housing Alliance, et al. v. Facebook* addressed a practice first publicized by ProPublica, in which Facebook provided tools to advertisers that allowed them to engage in disparate treatment by intentionally exposing members of protected classes to employment-, credit-, and housing-related ads at different rates.[41] A second suit, initiated by the Department of Housing and Urban Development (HUD) and eventually referred to the Department of Justice, alleged that Facebook was itself engaging in unlawful discrimination under the FHA because its methods for optimizing the delivery of ads could cause the platform to display housing-related ads at much different rates to different groups of people, even where the advertisers had not set out to differentially target users on such a basis.[42] Both suits have now settled, with Facebook agreeing to reform its advertising tools.[43]

While the FHA seems to be up to the task of tackling housing

[39] 369 F. Supp. 3d 362, 372 (D. Conn. 2019).

[40] *Id.* ("Defendant cannot downplay its role in the screening process. It was Defendant's form, Defendant's screening process and Defendant's adverse action letter that contributed to the denial of [the plaintiff's] application.")

[41] First Amended Complaint, Nat'l Fair Hous. All. v. Facebook, Inc., 2018 WL 8343918 (S.D.N.Y. June 25, 2018).

[42] Facebook, Inc., No. 018-0323-8 (Dep't of Hous. & Urban Dev. Mar. 28, 2019), https://www.hud.gov/sites/dfiles/Main/documents/HUD_v_Facebook.pdf.

[43] Facebook Settlement, NAT'L FAIR HOUSING ALLIANCE, https://nationalfairhousing.org/facebook-settlement/; Proposed Settlement Agreement, United States v. Meta Platforms, Inc., No. 1:22-cv-05187, ECF No. 5-1, (S.D.N.Y June 21, 2022).

discrimination caused by AI,[44] this does not seem to be the case with Title VII and ECOA.[45] Actors in the context of employment or credit that would seem to be equivalent to the entities covered by FHA are not subject to Title VII and ECOA. Vendors of employment assessments and credit scores seem to be beyond their scope, despite the decisive role that each might play in employment and credit decisions. This split in the apparent reach of Title VII and ECOA, on the one hand, and FHA, on the other, reveals a gap in the potential reach of discrimination law that the FTC would be well positioned to fill. To reach the vendors that sell to employers and creditors, the FTC could rely on its unfairness authority to hold these actors responsible for their roles in perpetuating discrimination, in much the same way that the FHA has been read to do so. More broadly, the FTC could seek to hold accountable any actor that the agency finds to have unfairly contributed to discriminatory outcomes, whether these are recognized as covered entities under any existing discrimination law.

One benefit to focusing on upstream actors in the AI pipeline is that they are often the cheapest cost avoider, better positioned to identify and address the source of discriminatory outcomes. For example, rather than targeting individual employers who purposefully rely on the technical expertise of vendors of employment assessments or hiring software, the FTC can target the vendors themselves, thereby placing regulatory pressure on the actor who is much more likely to have the relevant knowledge and practical skills to address the source of discriminatory outcomes. Addressing vendors also brings the benefits of scale. In targeting vendors, the FTC's enforcement actions could have a much greater effect in reducing discrimination in various markets because any remedial actions taken by the vendor would cascade down to all its clients. In many cases, going after the upstream suppliers of AI

[44] Even so, a recent story from ProPublica does not paint a very favorable picture of the current state of affairs, noting that "tenants often get the runaround when they complain about screening decisions. The screening companies say landlords decide what criteria to use. Landlords say screening companies make the decision." Erin Smith and Heather Vogell, *How Your Shadow Credit Score Could Decide Whether You Get an Apartment*, PROPUBLICA (Mar. 29, 2022), https://www.propublica.org/article/how-your-shadow-credit-score-could-decide-whether-you-get-an-apartment.

[45] It's worth noting that in Facebook's settlement with the National Fair Housing Alliance, it settled multiple cases simultaneously, involving claims about discriminatory ads for not just housing, but employment and credit as well. But it is likely that once Facebook had to make changes to its platform to settle the housing suit, it was essentially costless to agree to make the same changes in all five suits, regardless of whether it faced a real prospect of liability outside the FHA.

tools will be both more effective and more efficient in limiting discrimination than going after the downstream users of these tools.

Of course, the Commission need not pin all responsibility on just one actor. It can hold both vendors *and* clients responsible. The ability to do so may be especially useful because the distinction between developers and users of AI systems has begun to blur in many real-world applications, where vendors provide a rather general-purpose AI capability that is then adapted by clients to their particular requirements or circumstances.[46] Rather than having to decide in advance where it will allocate legal responsibility—a difficult question currently being debated by legislators in the European Union drafting the proposed AI Act[47]—the FTC can target its interventions at the actors that its investigations reveal to be responsible in any given case.

C. Discrimination Harms

In addition to expanded domains and defendants, the idea of discrimination harm or injury is also more capacious under the FTC's Section 5 authority than under discrimination law. We identify three types of discrimination harms that could concern the Commission: allocative, quality-of-service, and representational harms. Allocative harms are those that concern the distribution of a desirable resource or opportunity, such as a job, credit, or a home. These are the types of harms that are the concerns of traditional discrimination law. Quality-of-service harms are the injuries caused by consumer products and services that simply work less well for certain demographic groups than others. Representational harms capture cases where certain demographic groups are represented in a stereotypical or demeaning manner or where they are not acknowledged at all, harming their social standing in society.[48] Neither quality-of-service nor representational

---

[46] Jennifer Cobbe & Jatinder Singh, *Artificial Intelligence as a Service: Legal Responsibilities, Liabilities, and Policy Challenges*, 42 COMPUTER L. & SECURITY REV. 1 (2021).

[47] Michael Veale & Frederik Zuiderveen Borgesius, *Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach*, 22 COMPUTER L. REV. INT'L 97 (2021).

[48] Solon Barocas, Kate Crawford, Aaron Shapiro & Hanna Wallach, *The Problem with Bias: Allocative Versus Representational Harms in Machine Learning*, 9TH ANNUAL CONF. OF THE SPECIAL INTEREST GROUP FOR COMPUTING, INFO. & SOC'Y (2017); Kate Crawford, *The Trouble with Bias*, NeurIPS Keynote 2017,

harms is directly actionable under existing discrimination law.[49]

We begin with allocative harms because they are both the most familiar, and—perhaps surprisingly—the most complicated case. It is one thing to simply state that "discrimination is unfair"—a statement few would disagree with—and another to understand how the legal expression of discrimination contained within the civil rights statutes translates to unfairness under Section 5. As a starting point, the Commission could certainly rely on determinations of existing courts and other agencies with respect to discrimination law, holding that violations of existing discrimination law are unfair. It would be hard to argue that the Commission should not be free to adopt such determinations as a baseline. But the Commission is also not beholden to such determinations, as unfairness is inherently more flexible than discrimination law. But then what does an unfairness determination look like?

An important difference in the structure of these laws makes the translation challenging. Discrimination law lacks a concept of discrimination injury that is distinct from liability for it. As currently practiced, discrimination law is committed to procedural understandings of discrimination rooted in disparate treatment and disparate impact.[50] Disparate treatment is determined by whether a decision-maker had improper motives for an adverse action against a member of a protected class,[51] and disparate impact about whether a decision-maker had an acceptable reason to use a facially neutral decision mechanism that nonetheless had a disproportionate impact on people in a protected class.[52] Both of these ideas are rooted in a question of decisionmaker fault, while

https://www.youtube.com/watch?v=fMym_BKWQzk. Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. *Language (Technology) is Power: A Critical Survey of "Bias" in NLP*. PROCEEDINGS OF THE 58TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL), 5454-5476.

[49] Representational harms are relevant to existing discrimination law in the sense that decisions that appear based in stereotypes can form the basis for a determination of disparate treatment after a separate adverse action is taken, *see e.g.,* Chadwick v. WellPoint, Inc., 561 F.3d 38, 46–47 (1st Cir. 2009); Back v. Hastings On Hudson Union Free Sch. Dist., 365 F.3d 107, 120–21 (2d Cir. 2004), but business practices that create, reinforce, or propagate stereotypes are not themselves actionable under existing law.

[50] *See, e.g.*, Kim, *supra* note 5, at 902–03 ("Judges, litigants, and scholars commonly recite that Title VII prohibits two types of discrimination: disparate treatment and disparate impact.").

[51] 42 U.S.C. § 2000e-2(m).

[52] 42 U.S.C. § 2000e-2(k).

mostly indifferent to the outcomes faced by a person who suffered the harm—what we might otherwise call the injury.[53] Think about the way we speak about discrimination casually: we ask whether a victim was "discriminated against" with the "by a decisionmaker" part silently appended. Compare this to a legal regime like negligence where we ask the question of whether someone broke their leg separately from whether the driver of the car that hit them should bear liability. Discrimination law merges the fact of an injury with the question of responsibility for it, and as a result, there is no stable answer within discrimination law for what the "discrimination injury" is.[54]

Unfairness works differently, requiring that the Commission find a "significant injury"[55] as a predicate to a determination that something is unfair. Now, this does not pose a problem for the Commission if it merely adopts the holdings of discrimination law; the issue is not that discrimination law lacks a *claim* of injury, but rather that it has subsumed it into the question of liability. But there is a good conceptual reason to separate the issues. Specifically, the modern concept of discrimination, based in sociological understandings, is *far* more expansive than the legal definition. Ideas like structural, historical, and institutional discrimination are definitionally concepts of discrimination that lack a discriminator to blame.[56] Even where

---

[53] Alan David Freeman, *Legitimizing Racial Discrimination Through Antidiscrimination Law: A Critical Review of Supreme Court Doctrine* 62 MINN. L. REV. 1049, 1053 (1977).

[54] Noah D. Zatz, *Disparate Impact and the Unity of Equality Law*, 97 B.U. L. REV. 1357, 1379 (2017) ("[T]he fundamental question for the field is the nature of the relevant injury.") The first step of the disparate impact test is arguably *a* definition of injury, but that step is itself unstable and ill-defined. The EEOC long ago created the four-fifths rule in its Uniform Guidelines on Selection Procedures, 29 C.F.R. § 1607.4(D), but that rule is just a somewhat-arbitrary guideline, not followed in many contexts, and not theoretically justified by anything other than a need to draw a line. *See* Katie Eissenstat, *Lies, Damned Lies, and Statistics: The Case to Require "Practical Significance" to Establish a Prima Facie Case of Disparate Impact Discrimination*, 68 OKLA. L. REV. 641, 648 (2016) ("The advantages of the four-fifths rule are clear: it is an easy-to-calculate, simple test that puts responsible parties on notice of the relative balance an employer must achieve in its workforce to avoid liability. . . [But r]ather than directly addressing causation, many courts believe the four-fifths rule sets a seemingly arbitrary and unachievable evaluation of impact difficult for plaintiffs to meet." (footnotes omitted)); Scott W. McKinley, *The Need for Legislative or Judicial Clarity on the Four-Fifths Rule and How Employers in the Sixth Circuit Can Survive the Ambiguity*, 37 CAP. U. L. REV. 171, 177 (2008) ("As the courts continue to construe the four-fifths rule as having little validity, this federal "guideline" can no longer safely be used as a guideline for employers in administering promotional examinations.").

[55] 45 U.S.C. §15(n).

[56] *See generally* EDUARDO BONILLA-SILVA, RACISM WITHOUT RACISTS (6th. ed. 2022).

it would not amount to liability under discrimination law, it would also not be incoherent to hold that certain actions and business practices that subject consumers to greater harms from structural, historical, and institutional discrimination are unfair.

Thus, we come to the question of how to define a discrimination injury separate from liability. This is a much bigger question than we can resolve here, and as an essentially contested concept,[57] no answer we give would be truly satisfying. But conceptually, if liability is gated later in the process—if the mere fact of an injury caused by a business practice does not imply that it is inherently unfair—then it is best to consider the initial injury capaciously. To the extent we need a concrete answer, Professor Noah Zatz's treatment of the question is useful one from a practical perspective. He describes the relevant injury as "status causation," defining it as an individual suffering harm because of such individual's status with respect to protected class.[58] Zatz argues that this injury is consistent across different versions of equality law, with the different approaches—individual disparate treatment, systemic disparate treatment, disparate impact, and non-accommodation—differing only in the question of how the injury of status causation occurred. Thus, if a person is, say, not granted a loan, and that result is attributable in some part to their protected class status, that constitutes the injury of "status causation" but that fact would not constitute disparate treatment or impact until a court determines that the lender was in the relevant sense *responsible* for the harm.[59] While other definitions could be appropriate, defining the injury capaciously opens up a broader space to explore the question of when an allocative choice should be considered unfair.[60]

Whereas discrimination law might require a rewrite to accomplish this because there is so much precedent already established, the FTC is bound by neither the precedent nor the limiting language of the civil rights statutes. Conceptually, this opens a good deal of space to find unfairness, but in practice, we anticipate smaller departures.[61] Take, for example, the validation

---

[57] George Rutherglen, *Disparate Impact, Discrimination, and the Essentially Contested Concept of Equality*, 74 FORDHAM L. REV. 2313 (2006).

[58] Zatz, *supra* note 54, at 1359.

[59] *Id.*

[60] *See generally* Andrew D. Selbst, *Injury, Liability, and Responsibility in Discrimination Law* (in progress).

[61] We believe this partly because despite not being limited by the narrowness existing discrimination law, the politics of a decision to unilaterally define discrimination liability in ways that go far beyond will be a challenge.

of employment assessments for disparate impact cases. Validation is focused primarily on ensuring the overall empirical quality of the assessments.[62] The Society for Industrial and Organizational Psychology (SIOP), the professional society for many of the developers of such selection procedures, has long criticized the EEOC's Uniform Guidelines on Selection Procedures for failing to independently require differential validity—the question of whether a selection procedure is more accurate in assessing members of some groups than others.[63] The crux of the conflict is that the Guidelines are about compliance with disparate impact law while the SIOP is focused on detailing accepted professional practices for the fairest possible procedures.[64] This means that a selection procedure with wildly different accuracy rates between different groups would be considered unfair under professional standards, but if the outcome does not result in a large discrepancy between groups, it will not trigger disparate impact scrutiny because it will not fail the first prong of the disparate impact test.[65] Already, we are seeing companies exploit this distinction, by mathematically constraining their prediction models to satisfy the four-fifths rule, without consideration of how such constraints affect differential validity.[66] Such models will not trigger disparate impact law, but would surely be considered unfair under prevailing standards. If the FTC recognized differential validity independently as unfairness, it would then be a departure from existing discrimination law, but one solidly in line with professional standards of selection procedures.[67]

So far we have addressed only allocative harms. Another benefit of FTC involvement is that it can address quality-of-service harms. These are still concerning because they are discriminatory, but they are a different category of harm entirely. As mentioned earlier, if Apple's Siri or Google

[62] Uniform Guidelines on Employment Selection Procedures, 29 C.F.R. § 1607.5

[63] AMERICAN PSYCHOLOGICAL ASSOCIATION, PRINCIPLES FOR THE VALIDATION AND USE OF PERSONNEL SELECTION PROCEDURES __ (5th. ed. 2018)

[64] *See generally* Kelly Trindel, Sara Kassir & Jasen Bent, *Fairness in Algorithmic Employment Selection: How to Comply with Title VII*, 35 ABA J. LAB. & EMP. L. 241, 248–70 (2021) (comparing the Guidelines and SOIP Principles documents).

[65] *Id.* at 248 ("For the past several decades, a strict read of the SIOP Principles has seemingly suggested that employers can skip past the first prong of adverse impact analysis, relying rather completely on the second prong 'validity defense.'")

[66] *See e.g.*, Manish Raghavan, et al., *supra* note 32, at 472.

[67] Interestingly, the Uniform Guidelines do discuss differential validity—using the language of unfairness no less! 29 C.F.R. § 1607.14(B)(8). But unfairness is considered of secondary importance and only exists as part of validation, meaning only after the first prong triggers an inquiry.

Assistant are simply less capable of understanding the spoken language of certain demographic groups than others, the Commission might assert that the sale of such products is unfair. Consumers undoubtedly suffer a harm when products and services don't work for them as promised; they don't get what they pay for and they may bear the cost of inconvenience, either because they must take additional steps to get the products or services to work or because they must find some other way to achieve their goals. This is not an allocative harm as traditionally understood by discrimination law; it is a difference in the value that consumers are able to derive from the AI products and services that they have paid for. This is the most classic of consumer harms, an even more natural fit for the Commission as a consumer protection agency than the injuries covered by discrimination law.

Interestingly, the same activity can cause allocative harms to some consumers and quality of service harms to others. Specifically, if a decisionmaker—an employer, a lender, or a landlord—in good faith buys an AI product that is advertised as "fair" but is not in reality, then that decisionmaker has not gotten what they paid for—a quality-of-service harm. They have purchased a product that does not work, which itself constitutes a consumer harm, in addition to the resulting litigation risk or expense if they subject someone else to discrimination.[68] And those subject to such discrimination will, of course, experience this as an allocative harm.

Representational harms are the final type of injury that the FTC might seek to address, though, as explained below, such harms might not fall legally within the scope of unfairness.[69] In her article, Commissioner Slaughter raised the prospect of the FTC regulating the type of harms that Professor Safiya Noble has famously raised with regard to Google's search results.[70] Professor Noble reported that searches for "black girls" resulted in

[68] In addition, AI products and services that fail more often for some consumers than others can also cause dignitary harms if such failures communicate to consumers that the specific groups to which they belong are simply not worthy of having these product and services work well for them. Thus, even beyond the material costs of failure, AI products and services that exhibit performance disparities across groups can inflict psychic harms, undermining consumers' self-worth. But except in extreme cases, such individual dignitary harms will not be considered injuries that the FTC may address under its unfairness authority. Mark MacCarthy, *New Directions in Privacy: Disclosure, Unfairness and Externalities*, 6 I/S: J.L. & POL'Y FOR INFO. SOC'Y 425, 484 (2011) ("Emotional distress, mental anguish, loss of dignity and other harms are not ruled out by this criterion, but they must be effects that all or most or reasonable persons would construe as genuine harms.")

[69] *See* Part II.A.1 *infra* (discussing the "substantial harm" requirement).

[70] Slaughter et al., *supra* note 3, at 19.

a slew of pornographic and sexual content, while searches for "three black teenagers" showed images of violence.[71] In Commissioner Slaughter's article, she suggests that these harms could be construed as a failure to adequately test algorithms, but does not offer a theory as to why they constitute consumer injury.[72] We refer to these as representational harms because while they can certainly affect the dignity of the specific person that is being portrayed in harmful light, they are better understood as affecting the social standing of the entire group to which the person belongs.[73] Other notable examples of representational harms include AI systems that have misgendered people,[74] output ethnic slurs,[75] labeled people as animals,[76] or failed to recognize people in counter-stereotypical roles.[77]

D. Other Harmful Business Practices

Whereas discrimination law must start from a claim of discrimination, the FTC's focus is on unfair business practices in general. This means that it can address certain business practices that make discrimination more likely to occur or more difficult to address—business practices that would not by themselves amount to a claim of discrimination under any theory but could be nonetheless considered unfair because they will very likely lead to discriminatory injuries in commerce.

Such business practices could include the development and marketing of algorithmic systems that are untested for discrimination or where results of the testing are inadequately disclosed. Any time a company releases a product

---

[71] *See* Safiya Noble, *Google Has a Striking History of Bias Against Black Girls*, TIME (Mar. 26, 2018), https://time.com/5209144/google-search-engine-algorithm-bias-racism/

[72] Slaughter et al., *supra* note 3, at 19.

[73] Barocas, et al., *supra* note 48; Kate Crawford, *The Trouble with Bias*, NeurIPS Keynote 2017, https://www.youtube.com/watch?v=fMym_BKWQzk.

[74] James Vincent, *Automatic Gender Recognition Tech Is Dangerous, Say Campaigners: It's Time to Ban It*, THE VERGE (Apr. 14, 2021), https://www.theverge.com/2021/4/14/22381370/automatic-gender-recognition-sexual-orientation-facial-ai-analysis-ban-campaign.

[75] Amy Kraft, *Microsoft Shuts Down AI Chatbot After It Turned into a Nazi*, CBS NEWS (Mar. 25, 2016), https://www.cbsnews.com/news/microsoft-shuts-down-ai-chatbot-after-it-turned-into-racist-nazi/.

[76] Alistair Barr, *Google Mistakenly Tags Black People as 'Gorillas,' Showing Limits of Algorithms*, WALL. ST. J. (July 1. 2015), https://www.wsj.com/articles/BL-DGB-42522.

[77] Lisa Anne Hendricks, Kaylee Burns, Kate Saenko, Trevor Darrell, and Anna Rohrbach. *Women also Snowboard: Overcoming Bias in Captioning Models*, PROCEEDINGS OF THE EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV) 771-787 (2008).

that fails to test for likely harms or fails to adequately disclose facts that consumers need to safely use the product, the Commission does not need to wait for the harms to unfold before bringing an unfairness claim; the failure to adequately test or disclose could itself be the unfair business practice, rather than the downstream discrimination.[78] Another such practice could be a failure to anticipate or prevent foreseeable misuses. In one notable example, the ACLU of Northern California showed that Amazon's facial recognition product, Rekognition, had falsely matched 28 members of the U.S. Congress against a database of mugshots, with mismatches falling disproportionately on people of color.[79] The company responded by asserting that the study was flawed because it failed to adopt the appropriate confidence threshold for police use of 99%.[80] The default threshold, however, was set to 80%, and subsequent reporting revealed that police were unaware that they needed to change it, as the instructions were buried in documentation that predictably went unread.[81] The reporting further revealed that Amazon, as a matter of policy, does not suspend police use of the product even if the police fail to use the 99% confidence threshold.[82] Both the design of a product that will be used in a foreseeably harmful way and the failure to prevent the ongoing harm can be considered unfair.[83]

Practically, these claims will often coincide with discrimination claims. The Commission is unlikely to file a claim against a developer without discrimination having occurred, if only for the practical reason that such complaints are how the Commission would even discover the practice. One scenario that does present itself as likely to lead to one of these claims is where a decisionmaker such as an employer who relies on algorithmic software is sued for discrimination and defends itself (likely successfully) by pointing to a

[78] *Cf., e.g., In the Matter of Apple Inc., A Corp.*., No. 112-3108, 2014 WL 253519, at *1 (MSNET Jan. 15, 2014) (bringing a case based on a failure to disclose an aspect of technological design, rather than a claim that the technological design was itself injurious).

[79] Jacob Snow, *Amazon's Face Recognition Falsely Matched 28 Members of Congress with Mugshots*, FREE FUTURE (July 26, 2018), https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28.

[80] Matt Wood, *Thoughts on Machine Learning Accuracy*, AWS NEWS BLOG (July 27, 2018), https://aws.amazon.com/blogs/aws/thoughts-on-machine-learning-accuracy/.

[81] Bryan Menegus, *Defense of Amazon's Face Recognition Tool Undermined by Its Only Known Police Client*, GIZMODO (Jan. 31, 2019), https://gizmodo.com/defense-of-amazons-face-recognition-tool-undermined-by-1832238149.

[82] *Id.*

[83] *See generally* Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785, 819–20 (2015) (discussing the FTC's focus on design and defaults as unfair).

lack of disclosure by the developer or tools to mitigate the discrimination. Here, the FTC would be alerted to the possibility of discrimination but proving the actual harms to discriminated-against consumers would be much harder than proving the harm to the employers who were sued and faced litigation risk and cost. This would be a scenario where the failure to test, disclose, and mitigate would make sense to claim as its own business practice.[84]

There are also other types of business practices that could be considered unfair, but which are farther afield from actual discrimination claims. For example, Jenny Yang, former Chair of the Equal Employment Opportunity Commission has noted that vendors of hiring software have been known to sign contracts that indemnify their clients should their clients be sued for employment discrimination.[85] Such contractual terms might have been necessary for the industry to get off the ground, as potential clients might not have been willing to adopt these tools if they faced all the legal risk for adopting a new and legally untested technology. Or it might have been necessary for clients to take seriously vendors' claims that their offerings comply with the law: unless the vendors were willing to bear the costs of failure, then why should clients trust their claims? But Yang also suggests that indemnification by these vendors—usually small companies—is unreliable because multiple large cases could simply bankrupt them.[86] If software vendors are offering indemnification to sell their products, knowing that they will be unable to fulfill such guarantees because the follow-on claims after the first successful one will result in bankruptcy,[87] then it might be reasonable to hold that the practice is unfair. As the sector matures further, other business arrangements may become common which leave businesses better off but end up shifting risk to consumers.[88] Those kinds of concerns are classically in the

[84] Failure to disclose can be considered either deceptive or unfair. *See infra* Part II.B (discussing deceptive practices).

[85] Statement of Jenny R. Yang before the Civil Rights and Human Services Subcommittee, Committee on Education and Labor, United States House of Representatives (Feb. 5, 2020), https://edlabor.house.gov/imo/media/doc/YangTestimony02052020.pdf.

[86] *Id*. at 18 ("[E]mployers [cannot] rely on a vendor's promise of indemnification because multiple large cases could render a vendor unable to satisfy this promise.").

[87] *See supra* notes 85–86 and accompanying text.

[88] *Cf.* Blake E. Reid, *Internet Architecture and Disability*, 95 IND. L.J. 591, 621 (arguing that platforms similarly use contract law to shift responsibility for accessibility to the users of the platforms).

domain of consumer protection, and the FTC could get involved.

E. Advantages as a Litigant

The FTC possesses certain advantages in litigation as compared to individual plaintiffs or plaintiff classes bringing discrimination claims. First, where individual plaintiffs might not even know the discrimination is occurring, the Commission can collect complaints from multiple sources. As an enforcer with the power to open an investigation, the Commission can also gain access to the necessary information from a defendant before bringing a formal claim. Second, discrimination claims by individual litigants are often viewed skeptically by courts. There is an overriding concern that unless courts heavily police discrimination claims, innocent decisionmakers will end up buried in costly discovery on meritless cases. This skepticism has been present for a long time, with a result that discrimination plaintiffs rarely win.[89] Where a court might be skeptical of the initial claim by a single plaintiff, a claim by a government agency that a product or company discriminates against many people will likely not seem fanciful.

Third, the legal standard of injury under Section 5 is that an act or practice must either cause or be "*likely* to cause substantial injury to consumers"[90]—an inherently less demanding standard of proof than demonstrating that discrimination already did occur in the past. Moreover, as an enforcement agency, the Commission will not face a standing hurdle, even for harm that has not yet occurred. In recent years, the Supreme Court has expanded standing doctrine to make it ever harder for plaintiffs to bring claims of intangible harm. In *TransUnion* v. *Ramirez*, decided in 2021, the Court held that to satisfy standing, an injury must be both "concrete" and "particularized," where the concreteness requirement is not well-defined, but

[89] *See e.g.*, Michael Selmi, *Why Are Employment Discrimination Cases So Hard to Win?*, 61 LA. L. REV. 555, 562 (2001) ([C]ourts tend to view the claims of race plaintiffs skeptically…"); Melissa Hart, *Skepticism and Expertise: The Supreme Court and the EEOC*, 74 FORDHAM L. REV. 1937, 1938 (2006) ("The Court's reluctance to defer to EEOC interpretations may also reflect a broader skepticism about the scope of the problem of discrimination and the appropriateness of empowering a federal agency to define the problem and its possible solutions."); Kerri Lynn Stone, *Shortcuts in Employment Discrimination Law,* 56 ST. LOUIS U. L.J. 111, 115 (2011) (noting "the skepticism and hostility with which judges have regarded employment discrimination plaintiffs, as opposed to the way in which they have regarded traditional tort plaintiffs").

[90] 15 U.S.C. § 45(n) (emphasis added).

it is contrasted with a "bare procedural violation"[91] and asks whether there exists a "close historical or common-law analogue" to the claimed injury.[92] This may make it harder for individual plaintiffs to bring discrimination claims,[93] but it will not affect the FTC.

Fourth, the Commission need not rely on class actions. Many of the claims we discuss here are too small to support the expense of individual litigation. The class action mechanism exists to allow plaintiffs to band together and take advantage of economies of scale. In 2011, the Court decided *Wal-Mart* v. *Dukes*, a case that limits the ability of discrimination plaintiffs to claim the commonality of injury necessary to certify a class.[94] Scholars immediately noted that this case made class actions immensely harder for discrimination plaintiffs.[95] So, too, here. Some algorithmic discrimination cases might include claims that are amenable to class certification, such as a consumer suit against a company where the software was systematically worse for a class of purchasers in exactly the same way. Here the company's action was identical with respect to each consumer. But other types of cases will falter on class certification. For example, in a case where many different landlords buy tenant screening systems from the same vendor and then customize them, it is unlikely that a multi-jurisdictional class of rental applicants would have enough commonality to be certified under *Wal-Mart*, even if only one AI company developed the offending software. As scholars noted immediately with respect to the EEOC and employment discrimination, enforcement by government agencies is one way to get around the *Wal-Mart* problem.[96]

Many of these advantages are inherent to the fact that the FTC is an enforcement agency rather than an individual plaintiff, but they are still worth

---

[91] *Id.* at 2213 (quoting Spokeo, Inc. v. Robins, 578 U.S. 330 (2016))

[92] *Id.* at 2204.

[93] *See, e.g.*. Erwin Chemerinsky, *What's Standing After* Transunion LLC v. Ramirez*?*, 96 N.Y.U. L. REV. ONLINE 269, 283–84 (2021).

[94] Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338 (2011).

[95] *See* Joseph A. Seiner, *Weathering Wal-Mart*, 89 NOTRE DAME L. REV. 1343, 1350 & n.61 (2014) ("Scholars immediately—and correctly—denounced the case as one that undermines the rights of workplace discrimination victims.") (collecting sources).

[96] *Id.* at 1352 ("Perhaps the most obvious response to *Wal-Mart* is insisting that the case applies only to private plaintiffs bringing suit pursuant to Federal Rule of Civil Procedure 23. Thus, governmental agencies, such as the U.S. Equal Employment Opportunity Commission (EEOC), are not subject to the decision."); Tristin K. Green, *The Future of Systemic Disparate Treatment Law*, 32 Berkeley J. Emp. & Lab. L. 395, 397 (2011). ("Alternatively, the [EEOC] . . . . can sue on behalf of a class of women without obtaining class certification.").

noting because discrimination relies heavily on individual plaintiffs. Not only have courts been making such claims harder for plaintiffs for years now, but the threats of discriminatory AI occur much more at scale than on an individual basis. Reliance on individual claims to regulate large scale harms becomes a serious structural weakness of discrimination law when it comes to AI.

II. THE FTC'S AUTHORITY TO REGULATE DISCRIMINATORY AI

Section 5 of the FTC Act provides the Commission with authority to regulate "unfair and deceptive acts and practices" in commerce.[97] These concepts are expansive,[98] and Congress's goal in the original FTC Act was to leave the concepts flexible, giving the Commission the authority to determine what practices stood out as unfair or deceptive in a way that evolves over time as required.[99] As a result, this authority can reach essentially any type of activity in commerce that can injure individuals acting in a commercial capacity.[100]

While the deceptive practices authority is not particularly controversial or contested, the Commission's unfairness authority has been the subject of much controversy. To lawyers acquainted with the history of the FTC, the tale is familiar: In the 1970s, the Commission sought to aggressively expand the use of its unfairness authority and failed spectacularly when it turned out that government control of so much of the economy was unpopular. By 1980, the Commission had retreated, chastened. The Commission adopted a policy statement seeking to bound its own authority

---

[97] 45 U.S.C. § 15(a).

[98] Woodrow Hartzog & Daniel J. Solove, *The Scope and Potential of FTC Data Protection*, 83 GEO. WASH. L. REV. 2230, 2246 (2015) ("Rather than attempt to define the specific consumer protection issues that the FTC should focus on, Congress created two broad categories—practices that are deceptive and practices that are unfair—with virtually no hard boundary lines.")

[99] Atl. Ref. Co. v. FTC, 381 U.S. 357, 367 (1965) ("The Congress intentionally left development of the term 'unfair' to the Commission rather than attempting to define 'the many and variable unfair practices which prevail in commerce."); Fed. Trade Comm'n v. Bunte Bros., 312 U.S. 349, 353 (1941) (" '[U]nfair competition' was designed by Congress as a flexible concept with evolving content.")

[100] *See* 45 U.S.C. § 15(n);

(the "Unfairness Statement"),[101] and then, in 1994, Congress incorporated those limitations into the FTC Act,[102] leading to the modern understanding of unfairness as focused on consumer injury and consumer sovereignty.[103] Since then, the FTC has focused much more on enforcing deception than unfairness. The major exception would be the FTC's moves over the last decade to bring enforcement actions for unreasonably lax data security practices—moves which have engendered fierce opposition from big business which once again argues that the FTC is overstepping its bounds.[104]

This received wisdom has been enormously influential, shaping lawyers' and even commissioners' understanding of the FTC's role and the limits of its power for decades.[105] Fear of blowback from aggressive use of the Section 5 authority has starkly limited the Commission's use of its unfairness authority. While such a fear may be sound as a political matter—big business is very influential and high-profile members of both parties still prefer market self-regulation to government oversight of markets—it's about politics rather than a legal requirement. As Luke Herrine has persuasively argued, "[t]he main limitation on the use of the unfairness authority . . . has been the ideology of regulators charged with its enforcement," rather than any purely legal constraint within the FTC Act.[106] While the 1994 Amendments did put some formal constraints on what the FTC can do under its unfairness authority, it remains quite expansive, and whether because of conviction in the correctness of the consumer sovereignty view or timidity in the face of hostile politics, the Commission has never truly sought to test the authority after 1994.

In this Part, we explain why even though the FTC has never addressed discriminatory AI in the past, doing so falls squarely within its Section 5 power to regulate unfair practices as well as deceptive practices where appropriate. Nonetheless, while the current legal authority is relatively clear, the Supreme Court has become quite hostile to administrative agencies,

---

[101] FED. TRADE COMM'N, COMMISSION STATEMENT OF POLICY ON THE SCOPE OF THE CONSUMER UNFAIRNESS JURISDICTION (1980), *reprinted in* In re Int'l Harvester Co., 104 F.T.C. 949, 1072, 1073 & n.16 (1984) [hereinafter, UNFAIRNESS STATEMENT] at 1072–88.

[102] *See* Federal Trade Commission Act Amendments of 1994, Pub. L. No. 103-312, § 9, 108 Stat. 1691, 1695 (codified at 15 U.S.C. § 45(n)).

[103] Herrine, *supra* note 88, at 441.

[104] *See* Fed. Trade Comm'n v. Wyndham Worldwide Corp., 799 F.3d 236 (3d Cir. 2015), LabMD, Inc. v. Fed. Trade Comm'n, 894 F.3d 1221, 1229 (11th Cir. 2018).

[105] *See generally* Herrine , *supra* note 88, at 439–44.

[106] *Id.* at 433.

especially when they try to address novel problems. Thus, this Part will also discuss the potential impact of the major questions doctrine on Commission efforts to address discriminatory AI.

A. Unfair Practices

The FTC has wide latitude to determine that a particular business practice is unfair. There are no subject matter restrictions, as long as the practice is "in or affecting commerce."[107] The only legal restrictions are set out in Section 5(n) of the FTC Act. To be considered unfair, an act or practice must (1) cause or be likely to cause "substantial injury to consumers," which is (2) "not reasonably avoidable by consumers themselves," and (3) "not outweighed by countervailing benefits to consumers . . . ."[108] Nonetheless, the FTC has historically been cautious in using its unfairness authority, instead relying principally on its deception authority.[109] Below, we illustrate why, if the Commission pursues discriminatory AI under an unfairness theory, it should not have much trouble satisfying the legal requirements.

1. Substantial Injury

The first requirement is that an unfair act or practice must cause or be likely to cause "substantial injury to consumers." There is little concrete law regarding what counts as a substantial injury.[110] We know that the type of injury matters. Tangible harms, including economic, monetary, or health-related, all count.[111] Intangible harms are more controversial. While the Commission has brought an action arguing that secret monitoring inside a home was unfair,[112] purely emotional harms are usually only considered in

---

[107] 15 U.S.C. § 45(a).

[108] *Id.* § 45(n).

[109] Terrell McSweeny, *Psychographics, Predictive Analytics, Artificial Intelligence, & Bots: Is the FTC Keeping Pace?*, 2 GEO. L. TECH. REV. 514, 522 (2018).

[110] *See* Concurring Statement of Acting Chair Maureen Ohlhausen, In the Matter of VIZIO, Inc., https://www.ftc.gov/public-statements/2017/02/concurring-statement-acting-chairman-maureen-k-ohlhausen-matter-vizio-inc ("This case demonstrates the need for the FTC to examine more rigorously what constitutes 'substantial injury' in the context of information about consumers.").

[111] HOOFNAGLE, *supra* note 8, at 132; Hirsch, *New Paradigms*, *supra* note 6, at, 483.

[112] Complaint, In the Matter of DesignerWare, LLC, FTC File No. 1123151 (April 15, 2013).

extreme cases.[113] The degree of harm also matters. The harm cannot be trivial or speculative.[114] But the harm to an individual need not be large on its own if the Commission can find substantial harm by aggregating across many consumers.[115]

Allocative harms have clear and concrete economic consequences. If an employer, creditor, or landlord purchases AI-based screening software, and that screen blocks people in protected classes from opportunities for employment, loans, or housing, there is no doubt that any injury is substantial because people are denied a major life opportunity.[116] The FTC's focus, at least so far, has been on allocative harms, likely because they are the clearest to name as injurious.[117]

Quality-of-service harms also result in concrete economic harm. Recall the hypothetical in which a phone manufacturer implements a facial recognition feature to help users more easily unlock their devices—but where that feature fails to work much more often for certain demographic groups than others. If part of the value of these devices is that it offers such a feature, then the people paying for them who find that they are unable to make use of the feature will have suffered a classic consumer injury: they do not enjoy some benefit that they have paid for. Of course, the injury must be substantial. But this should pose no problem. Aggregating even small injuries adds up. Consider what this might mean concretely in the case of phones. Some estimates put the last several years of annual phone sales around the $55–75-billion-dollar mark in the United States,[118] so if such a feature adds even a tenth of a percent of the value of the phone, and even five percent of people are unable to use the feature, then the total cost to consumers is in the ballpark

---

[113] MacCarthy, *supra* note 68, at 484 ("Emotional distress, mental anguish, loss of dignity and other harms are not ruled out by this criterion, but they must be effects that all or most or reasonable persons would construe as genuine harms."); UNFAIRNESS STATEMENT, *supra* note 101, at 1073 & n.16 (noting that emotional harm may be sufficient only in an "extreme case").

[114] UNFAIRNESS STATEMENT, *supra* note 101, at 1073.

[115] Hirsch, *New Paradigms*, *supra* note 6, at 483, HOOFNAGLE, *supra* note 8, at 132.

[116] Hirsch, *That's Unfair*, *supra* note 6, at 354.

[117] *See* Jillson, *supra* note 3 (discussing examples like employment, credit, and health care).

[118] Statista, *Smartphone Sales Forecasts in the United States from 2005 to 2022* https://www.statista.com/statistics/191985/sales-of-smartphones-in-the-us-since-2005/; Chris Kolmar, *U.S. Smartphone Industry Statistics [2022]: Facts, Growth, Trends, and Forecasts* (Jan. 30 2022), https://www.zippia.com/advice/us-smartphone-industry-statistics/#:~:text=The%20United%20States%20is%20the,around%20the%20world%20every%20year.

of $3 million. That's not out of line with the FTC's consumer protection cases.[119] Perhaps more importantly, the substantiality requirement refers more to the type of injury than the degree of harm.[120] Physical and economic harms are recognized as substantial injuries, but "[e]motional impact and other more subjective types of harm, on the other hand, will not ordinarily make a practice unfair."[121] Because quality of service harms are economic in nature, they will easily be established as substantive.

Last is representational harms, such as what Commissioner Slaughter alluded to in terms of harms from search engine results. This is certainly trickier. Perpetuating harmful stereotypes and associations writ large is undoubtedly harmful. Individuals do suffer psychic and dignitary harms that arise when they encounter such representations. These types of harms, real though they are, are inchoate and likely fall within the subjective and thus insubstantial category. While we believe it is overly dismissive to treat these experiences as "merely . . . offend[ing] the tastes or social beliefs of some [consumers],"[122] many would argue that that is the most appropriate treatment.[123]

Representational harms such as those described can also be seen as harms to consumers as a whole. The textual requirement of Section 5 states that an unfair practice must include a "substantial injury to consumers."[124] While it is arguably most natural to read the plural form of consumer as the aggregate of individual consumers, and thus the substantial injury requirement as an aggregation, the language can also be read to include injuries to consumers as a body. In this way we distinguish representational harms from dignitary harms.[125] Consumers in the aggregate can suffer many individual, subjective dignitary harms, but representational harms are of a

[119] *Cf.* Fed. Trade Comm'n v. Wyndham Worldwide Corp., 799 F.3d 236, 242 (3d Cir. 2015) ($10.6 million); *In the Matter of Dave & Busters, Inc.*, 149 F.T.C. 1449 (2010) ("several hundred thousand dollars in fraudulent charges").

[120] Maureen K. Ohlhausen, *Weigh the Label, Not the Tractor: What Goes on the Scale in an FTC Unfairness Cost-Benefit Analysis?*, 83 GEO. WASH. L. REV. 1999, 2017 (2015) (noting that in *International Harvester*, the case that established the current unfairness requirements most clearly, the FTC focused on the "character of the injury, rather than its magnitude.").

[121] UNFAIRNESS STATEMENT, *supra* note 101, at 1073.

[122] *Id.*

[123] *Cf.* MARTHA CHAMALLAS & JENNIFER B. WRIGGINS, THE MEASURE OF INJURY: RACE, GENDER, AND TORT LAW 89–93 (discussing the history of dismissing emotional and mental injuries in tort law).

[124] 15 U.S.C § 45(n).

[125] *See* note 68, *supra*.

different character.[126] Representational harms demonstrate that a group of people are fundamentally second-class citizens, a concept that has been recognized as harmful to the affected group and to society as a whole.[127] By considering consumers as a group, rather than an aggregate, the Commission may be able to argue that this type of harm is not subjective like emotional harm, and should be counted as an injury.[128]

Finally, as to all three types of injury, there is a good argument that the FTC would be within its rights to call discrimination per se injurious. Indeed, according to *In re Napleton Automotive Group Commission*, a recent enforcement action, Chair Lina Khan and Commissioner Slaughter seem to interpret it exactly this way.[129] To understand why, it is useful to examine how the Commission may consider public policy in its determinations. According to Section 5(n): "In determining whether an act or practice is unfair, the Commission may consider established public policies as evidence to be considered with all other evidence. Such public policy considerations may not serve as a primary basis for such determination."[130] On the one hand, public policy may be considered; on the other hand, it may not be

---

[126] Richard H. Pildes, *Why Rights Are Not Trumps: Social Meanings, Expressive Harms, and Constitutionalism*, 27 J. Legal Stud. 725, 754–55 (1998) ("Expressive harms are . . . social rather than individual. Their primary effect is not the tangible burdens they impose on particular individuals but the way in which they undermine collective understandings.").

[127] Richard A. Primus, *Equal Protection and Disparate Impact: Round Three*, 117 HARV. L. REV. 493, 566–67 (2003) ("Leading decisions from Strauder v. West Virginia to Brown v. Board of Education turned at least in part on the anti-egalitarian social meanings of the practices at issue. . . . Disparate impact doctrine has not traditionally been thought of as something that might give rise to expressive harms, but . . . [o]nce the question is asked, it seems plausible . . . ."); Deborah Hellman, *The Expressive Dimension of Equal Protection*, 85 MINN. L. REV. 1, 3 (2000) (arguing that equal protection "inheres in what the law expresses"); Charles L. Black, Jr., *The Lawfulness of the Segregation Decisions*, 69 YALE L.J. 421, 428 (1960) (arguing that segregation is unlawful because of social meaning); MARI MATSUDA, WORDS THAT WOUND ("*Brown* . . . articulates the nature of the injury inflicted by the racist message of segregation.)

[128] The courts, however, may be skeptical about this. While the Supreme Court has been willing to recognize racial discrimination as an expressive harm, it seems much more likely to do so in affirmative action cases brought by white plaintiffs than disparate impact cases brought by people of color. *See* William M. Carter, Jr., *Affirmative Action as Government Speech*, 59 UCLA L. REV. 2, 19–29 (2011).

[129] In the Matter of Napleton Automotive Group Commission, File No. 2023195, 2022 WL 1039797, at *3. (stating categorically that "discrimination based on protected status is a substantial injury to consumers").

[130] 15 U.S.C § 45(n).

primary. What does this mean?

The language, as with the rest of Section 5(n), came from the Unfairness Statement. Because Section 5(n) codified the exact test in the Unfairness Statement and the legislative history is otherwise scant, the Unfairness Statement is usually treated as the legislative history of Section 5(n). Prior to that, the FTC's position on whether a practice was unfair had been governed by the Cigarette Rule, which held that a practice is unfair if it "offends public policy," is "immoral, unethical, oppressive, or unscrupulous," or "causes substantial injury to consumers."[131] In 1972, the Supreme Court upheld this broad understanding, holding that the FTC should, "like a court of equity, consider[] public values beyond simply those enshrined in the letter or encompassed in the spirit of the . . . laws,"[132] but it was seen by many as unprincipled, allowing commissioners to substitute their "personal values" for any sort of measured determination of unfairness.[133] Ultimately, the Unfairness Statement—and thus Section 5(n)—was a reaction to the breadth of this authority.

The purpose of the amendment was to concretize the FTC's unfairness authority by tethering it to the three elements that were later codified, structuring it beyond regulating a "general sense of national values." As Luke Herrine observes, however, while the Unfairness Statement sought to narrow the FTC's authority, it did so in a way that "emphasize[d] continuity":

> Although the Unfairness Policy Statement is now commonly treated as a break from past statements on the meaning of unfair acts, its purpose was to emphasize continuity. On its face, the Statement presents itself as a synthesis of "the most important principles of general applicability" that can be drawn from "decided cases and rules." Rather than using the history of consumer unfairness to articulate a *new* standard, the Statement uses that history to clarify the meaning of the closest thing to an old standard: the Cigarette Rule. . . . That is not to say that the Statement simply restates the Cigarette Rule. It instead elevates the first ("consumer injury") prong, cabins the meaning of the second ("public policy") prong, and tosses aside the third ("immoral, unethical, oppressive, or

[131] Unfair or Deceptive Advertising and Labeling of Cigarettes in Relation to the Health Hazards of Smoking, 29 Fed. Reg. 8324 (July 2, 1964) (codified at 16 C.F.R. § 408.1)

[132] Fed. Trade Comm'n v. Sperry & Hutchinson Co., 405 U.S. 233, 244 (1972).

[133] J. Howard Beales, *The FTC's Use of Unfairness Authority: Its Rise, Fall, and Resurrection*, FED. TRADE COMM'N, (May 30, 2003), https://www.ftc.gov/public-statements/2003/05/ftcs-use-unfairness-authority-its-rise-fall-and-resurrection (speech at Marketing and Public Policy Conference, Washington, D.C.).

> unscrupulous") prong. But it does so on the grounds that the FTC's experience applying the unfairness standard has clarified the relevance of each prong . . . [and] the third prong [was] "largely duplicative."[134]

While the amendment sought to structure and limit the FTC's unfairness analysis, it did so not by rejecting old interpretations or limiting the types of practices it could find as unfair, nor—most importantly—by rejecting public policy as a consideration. Indeed, the Unfairness Statement discussed the use of public policy explicitly, describing two ways it should be used. First, it said that public policy may inform whether a practice is injurious or unfair, in either direction.[135] Second, public policy may "independently support a Commission action," overriding the injury question, "when the policy is so clear that it will entirely determine the question of consumer injury, so there is little need for separate analysis by the Commission."[136] Thus, Section 5(n)'s prohibition on using public policy as a "primary basis" for an unfairness determination sounds like a stronger prohibition than it is.[137]

---

[134] Herrine, *supra* note 88, at 509 (footnotes omitted).

[135] UNFAIRNESS STATEMENT, *supra* note 101, at 1075 ("[T]he Commission wishes to emphasize the importance of examining outside statutory policies and established judicial principles for assistance in helping the agency ascertain whether a particular form of conduct does in fact tend to harm consumers. Thus the agency has referred to First Amendment decisions upholding consumers' rights to receive information, for example, to confirm that restrictions on advertising tend unfairly to hinder the informed exercise of consumer choice. Conversely, statutes or other sources of public policy may affirmatively allow for a practice that the Commission tentatively views as unfair.")

[136] *Id.*

[137] Unfortunately, this language has sometimes been read as an additional constraint on the unfairness power beyond the three-part test written into the statute. Take *LabMD*, for example. While evaluating the FTC's authority to regulate data security under its unfairness authority, the Court wrote, based on the Unfairness Statement, that "an act or practice's 'unfairness' must be grounded in statute, judicial decisions—*i.e.*, the common law—or the Constitution. An act or practice that causes substantial injury but lacks such grounding is not unfair within Section 5(a)'s meaning." LabMD, Inc. v. Fed. Trade Comm'n, 894 F.3d 1221, 1229 (11th Cir. 2018). This statement turned out to be dicta because the court ended up assuming without deciding that the FTC did have such authority, *id.* at 1231, and it's a good thing too, because it's just not true. Nothing in the statute or the history creates an independent requirement that unfairness must be premised on a violation of other law. The Supreme Court expressly rejected that idea in 1972, holding that the FTC's unfairness authority was meant to be broad and flexible, that "Congress [did not] intend[] to confine the forbidden methods to fixed and unyielding categories." Fed. Trade Comm'n v. Sperry &

Because at the time of the Unfairness Statement there was no injury requirement at all, that prohibition is best read as a limitation solely on where public policy may be held to *substitute* for the brand-new injury requirement of Section 5(n).[138] Thus, either the policy is to be used to support the injury decision, or the policy is so clear that the demonstration of injury is simply unnecessary at all. Thus, for any practice that would create discrimination liability under law, it's not just that the FTC may find it unfair, but that the FTC need not even demonstrate any other concrete injury because the injury can be presumed. At least for that subset, the FTC can argue injury per se.

The remaining question for quality-of-service and representational harms is how anti-discrimination public policy either informs or supplants the injury determination. Here, the picture is unclear. On the one hand, current discrimination law doesn't capture these harms. There is no law that currently imposes liability for quality-of-service or representational harms alone. This could be seen not only to fail to support a determination of per se unfairness, but to actively cut against it. On the other hand, we have many, many, laws against discrimination, demonstrating that if something is considered discrimination, we should be confident that it should be treated as unfair.

The key question, it seems, is whether the recognition of discrimination injuries for the purpose of finding a per se injury is limited to those injuries recognized by existing anti-discrimination law. As to quality-of-service harms, we do not see much difference between these in principle than any other type of discriminatory harm. As with prohibitions on discrimination in public accommodations,[139] discrimination is not merely wrongful in cases that affect major life opportunities. Lesser quality of service in a purchased good due to race is no less an injury than lesser service in a lunch counter for the same reason.

Representational harms are, again, a tricky case. They are not universally recognized as harms in the same way as allocative discrimination is. But nor are these types of harms unfamiliar. Scholars since *Brown v. Board*

Hutchinson Co., 405 U.S. 233 (1972). The 1994 amendment did not disturb that finding, instead observing that everything that the FTC did find unfair was already covered primarily by consumer injury and secondarily by public policy. Rather than disturb the Supreme Court's reading of the breadth with respect to subject matter, the amendment changed how unfairness should be presented and defended.

[138] UNFAIRNESS STATEMENT, *supra* note 101, at 1075 ("In these cases the legislature or court, in announcing the policy, has already determined that such injury does exist and thus it need not be expressly proved in each instance.")

[139] 42 U.S.C. § 2000a.

of Education have discussed the case in terms of the expressive harm of segregation.[140] The principle of expressive or social meaning harm lives on in various other contexts as well.[141] So while public policy here may not enable the Commission to label representational harm discrimination per se, it does seem that there is support out there in existing policy for the recognition of the injury. There is also a good reason that representational harms are not included in existing discrimination law outside of the equal protection context: Discrimination law as written targets individual injury and expressive harm is not individual. Even if it were written into the law, no plaintiff could bring a case. So its absence should not be thought to conclusively reject the idea as a harm. But of course, if the FTC brought a case on this theory, it could implicate First Amendment concerns as well.[142]

Ultimately, the Commission should have no trouble at all counting allocative and quality of service harms as substantial injuries, and it would have an argument for counting representational harms as an injury as well.

2. Not Reasonably Avoidable

The second requirement under Section 5(n) is that the injury must be "not reasonably avoidable by consumers themselves."[143] This requirement enshrines the idea of consumer sovereignty and consumer choice.[144] If something would injure consumers but they could nonetheless choose whether to encounter it, then, by the reasoning of the statute, calling such a situation unfair would be unreasonably paternalistic, presuming consumers are not capable of balancing pros and cons on their own.

Here, again, we can split the analysis into allocative, quality-of-service, and representational harms. Allocative harms are again the easy cases. Where an employer, lender, or landlord uses a discriminatory AI

[140] *See* sources cited *supra* note 127.

[141] *See* Hellman, *supra* note 125; C. Edwin Baker, *Outcome Equality or Equality of Respect: The Substantive Content of Equal Protection Fna1*, 131 U. PA. L. REV. 933, 934 (1983) (arguing for "equality of respect" model of equal protection."); Richard H. Pildes & Richard G. Niemi, *Expressive Harms, "Bizarre Districts," and Voting Rights: Evaluating Election-District Appearances After Shaw v. Reno*, 92 MICH. L. REV. 483, 506–07 (1993) ("An expressive harm is one that results from the ideas or attitudes expressed through a governmental action, rather than from the more tangible or material consequences the action brings about.");

[142] *See, e.g.*, Meg Leta Jones, *Silencing Bad Bots: Global, Legal and Political Questions for Mean Machine Communication*, 23 COMM. L. & POL'Y 159, 184 (2018).

[143] 15 U.S.C § 45(n).

[144] Herrine, *supra* note 88, at 441–42.

product that harms consumers, the harmed consumers cannot avoid such harm, because it is not their choice of whether the product is used or not, and most of the time they will not even know. Even in the cases where a decision maker is up front about its use of AI tools, the only way for an applicant to avoid them is to first investigate their potentially discriminatory effects, then decline to apply for a job, apply for a loan, or apply for a residence.[145] While this suggests that there is some limited sense in which these algorithmic harms are avoidable, a "reasonably avoidable" requirement cannot be thought to mandate such extreme measures.

Only slightly harder are cases of quality-of-service harms. If the harm in these cases is that certain consumers will fail to receive the benefit of the purchased products and services, then such a harm will not be unfair so long as these consumers can reasonably avoid it by purchasing alternatives available in the market. But there are several reasons that consumers will generally be unable to do so. The first problem is that whether a consumer AI product and service is discriminatory is not generally known, even to their developers and vendors. Most consumer AI products and services are not

[145] A pair of recent laws has tried to address some of this problem in the case of hiring. In 2020, Illinois passed a law that requires that job applicants receive notice when their video interviews are subject to "artificial intelligence analysis". Artificial Intelligence Video Interview Act, 820 Ill. Comp. Stat. 42/1 *et seq.* https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=4015&ChapterID=68. In 2021, New York City passed a law that requires that job applicants receive notice when they are subject to an "automated employment decision tool"; it also requires that such tools be subject to audit and that a summary of the results of the audit be made publicly available. N.Y.C. Admin Code §20-870, https://legistar.council.nyc.gov/ViewReport.ashx?M=R&N=Text&GID=61&ID=4550092&GUID=FA34B1CB-AD88-41F1-9179-CAFE5A238810&Title=Legislation+Text. Both laws require that applicants receive notice *prior* to being subject to evaluation and that applicants have the option of requesting an alternative means of evaluation—that is, of opting-out of the evaluation performed by AI or an automated tool. Even under these circumstances, advocates have worried that applicants might lack the necessary knowledge to make well-informed decisions about whether to agree to such evaluations, largely because the laws do not compel sufficient disclosure to determine whether the tools might exhibit bias. For more discussion of these laws, *see generally* Brittany Kammerer, *Hired by a Robot: The Legal Implications of Artificial Intelligence Video Interviews and Advocating for Greater Protection of Job Applicants*, 107 IOWA L. REV. 817 (2021); Matt Scherer & Ridhi Shetty, *NY City Council Rams Through Once-Promising but Deeply Flawed Bill on AI Hiring Tools* (November 12, 2021), https://cdt.org/insights/ny-city-council-rams-through-once-promising-but-deeply-flawed-bill-on-ai-hiring-tools/.

currently tested for bias.[146] Rather than performing a thoroughgoing audit of their products and services—grappling with challenging questions about various ways their product might discriminate—many companies appear to declare them discrimination-free based on simplistic and often faulty understandings of discrimination.[147] When confronted with claims of discrimination, some companies have responded by denying the accusation and further obfuscating how their products and services work.[148] But even when a company wants to audit its products, doing so can be quite challenging, and no standard frameworks exist with which to evaluate the quality of the audit.[149]

There is a push within the academy and industry for regularized auditing and better documentation, so there is a chance that auditing may

---

[146] Inioluwa Deborah Raji et al, *Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing*, PROCEEDINGS OF THE 2020 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FACCT) 33 (arguing for the necessity of conducting internal audits before product release).

[147] Raghavan, et al., *supra* note 32; Elizabeth Anne Watkins, Michael McKenna & Jiahao Chen, *The Four-Fifths Rule is Not Disparate Impact: A Woeful Tale of Epistemic Trespassing in Algorithmic Fairness* (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4037022.

[148] Joy Buolamwini, *Response: Racial and Gender bias in Amazon Rekognition — Commercial AI System for Analyzing Faces*, MEDIUM (Jan. 25, 2019), https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced; Inioluwa Deborah Raji & Joy Buolamwini, *Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products*, PROCEEDINGS OF THE 2019 AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY 429.

[149] McKane Andrus, Elena Spitzer, Jeffrey Brown & Alice Xiang, *What We Can't Measure, We Can't Understand: Challenges to Demographic Data Procurement in the Pursuit of Fairness*, PROCEEDINGS OF THE 2021 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FACCT) 249; Inioluwa Deborah Raji, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee & Emily Denton, *Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing*, PROCEEDINGS OF THE AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY 145 (2020); Solon Barocas, Anhong Guo, Ece Kamar, Jacquelyn Krones, Meredith Ringel Morris, Jennifer Wortman Vaughan, W. Duncan Wadsworth & Hanna Wallach, *Designing Disaggregated Evaluations of AI Systems: Choices, Considerations, and Tradeoffs*, PROCEEDINGS OF THE 2021 AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY 368; Inioluwa Deborah Raji & Jingying Yang *ABOUT ML: Annotation and Benchmarking on Understanding and Transparency of Machine Learning Lifecycles*, https://arxiv.org/abs/1912.06166 (manuscript at 2); Alfred Ng, *Can Auditing Eliminate Bias from Algorithms?*, THE MARKUP (Feb. 23, 2021), https://themarkup.org/ask-the-markup/2021/02/23/can-auditing-eliminate-bias-from- algorithms.

soon become a more common and more standardized practice. This brings us to the second problem. Even if audits become commonplace, consumers will still not be aware of the discriminatory nature of the products they use. Companies currently have plenty of incentives to keep them private, such as fear of liability or the possibility of exposing trade secrets. Legislation seeking to mandate impact assessments and documentation is often designed with limited transparency as an explicit trade-off to enhance the likelihood of procedural compliance, so these incentives are being built into the legal responses that we are likely to see, making them unlikely to significantly change.[150] Even when audits are made public, they may be difficult to interpret,[151] and most consumers will lack the sophistication to understand their implications for the products they buy.[152] In the hypothetical future world where every company somehow makes a an audit publicly available and easily understandable, discrimination audits may become yet another document—like privacy policies[153]—that consumers cannot possibly read for each and every product they use, making the harms nonetheless unavoidable.

Even setting aside these serious problems with consumer knowledge and understanding, the ability for consumers to avoid biased products and services requires that there be sufficient competition that consumers have viable alternative choices. This might pose a particular challenge in the case of AI. By their very nature, AI and other tools that facilitate automation tend toward scale and standardization. In replacing a large set of hiring managers, loan officers, or landlords with software, hiring, lending, and housing decisions become far more uniform across applications. As Cathy O'Neil has argued, problematic software has the potential to affect a much larger scale

---

[150] *See* Algorithmic Accountability Act of 2022, S.3752, 117th Cong. § 6 (2022) (requiring no publication of impact assessments but requiring the FTC to publish annual reports and a limited repository of public information); Working Party on the Protection of Personal Data 95/46/EC, Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679, art. 29, WP 248 (Apr. 4, 2017), at 18 (Official guidance stating that DPIAs under GDPR Art. 35 require no publication, but recommending publication of a summary), *see also* Margot E. Kaminski, *Understanding Transparency in Algorithmic Accountability*, *in* CAMBRIDGE HANDBOOK OF THE LAW OF ALGORITHMS (Woodrow Barfield, ed. 2020) 137.

[151] Barocas et al., *supra* note 149 at __.

[152] *See* Hirsch, *That's Unfair!*, *supra* note 6, at 354.

[153] *See generally* Aleecia M. McDonald and Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 I/S: J.L. & POL'Y FOR INFO. SOC'Y 543 (2009).

of people than the faulty decision-making of any given human.[154] Scholars have also recently warned of an algorithmic monoculture that leaves unsuccessful applicants, perhaps subject to biased assessment, with no place to turn when the natural diversity present in human decision making is replaced with the homogeneity of a fixed decision rule.[155] Much the same applies for AI-based products and services because advances in these technologies depends on having access to the kinds of large datasets that are very difficult for new market entrants to develop, limiting the degree of competition that is likely to occur. For these tools to be reasonably avoidable by consumers, there needs to be meaningful differences in the decision making of the various actors in the market and in the products and services on offer. AI, as a technology, is likely to work against this goal.

Finally, representational harms are inherently unavoidable. Representational harms work to undermine the social standing of members of certain groups by affecting the beliefs and attitudes that others hold about these groups. There is no way for members of these groups to avoid such harms because they are in no position to control the representations to which these other people are exposed.

3. Cost-Benefit Analysis

The final requirement of Section 5(n) is the cost-benefit analysis. For a practice to be deemed unfair, the harms must not be "outweighed by countervailing benefits to consumers . . . ."[156] There is surprisingly little law on how the FTC should conduct such an analysis. The statute says nothing at all about what the cost-benefit analysis should entail or how it should be performed. The Unfairness Statement is also light on specifics, though it stresses that the analysis needs to consider the cost of possible remedies, not just the costs of the challenged practice. It states, for example, that failing to disclose information that might aid consumers in avoiding some harm could be justified if doing so allows a company to charge consumers lower prices

[154] CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016).

[155] Jon Kleinberg and Manish Raghavan, *Algorithmic Monoculture and Social Welfare*, 118 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES 1 (2021); Kathleen Creel & Deborah Hellman, *The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision Making Systems*, CANADIAN JOURNAL OF PHILOSOPHY 1 (March 4, 2022).

[156] 15 U.S.C § 45(n); *see also* Hirsch, *New Paradigms*, *supra* note 6, at 484 ("The FTC and commentators have interpreted this to require a cost-benefit analysis.").

(on the idea that furnishing disclosure can be a costly undertaking and might raise operating costs and thus prices).[157] It further notes that the Commission must consider the cost of "increased paperwork, increased regulatory burdens on the flow of information, reduced incentives to innovation and capital formation, and similar matters."[158] While there is an explicit focus on the ancillary costs of a remedy, there is no mention of the ancillary costs to consumers of failing to remedy consumer harm, including increased time and resources spent to avoid the harm, where possible.

Although cost-benefit requirements generally evince a deregulatory posture,[159] the history of FTC enforcement suggests that they do not present a significant hurdle to intervention. In most of the cases that the FTC brings under its unfairness authority, it alleges that the cost-benefit test is satisfied without any real argument, and such claims seem rarely to be challenged,[160] which at least partly explains the lack of legal development. The precedent set by these cases suggests that little analysis is necessary to satisfy the cost-benefit test—and that the law affords the Commission flexibility in setting up and conducting its analysis.

On the few occasions when the FTC has engaged in a more substantial cost-benefit analysis, much of the analysis has focused on questions of scoping. Cost-benefit analysis is notoriously sensitive to how relevant costs and benefits are scoped and measured; different choices can give rise to

---

[157] UNFAIRNESS STATEMENT, *supra* note 101, at 1073 ("A seller's failure to present complex technical data on his product may lessen a consumer's ability to choose, for example, but may also reduce the initial price he must pay for the article. The Commission is aware of these tradeoffs and will not find that a practice unfairly injures consumers unless it is injurious in its net effects.")

[158] *Id.*

[159] *See generally* David M. Driesen, *Is Cost-Benefit Analysis Neutral?*, 77 U. COLO. L. REV. 335, 402 (2006) (finding that cost-benefit analysis is deregulatory as a matter of historical practice); Nicholas Bagley & Richard L. Revesz, *Centralized Oversight of the Regulatory State*, 106 COLUM. L. REV. 1260, 1268–83 (2006) (OIRA budget review is a deregulatory "one-way ratchet").

[160] *See, e.g.*, In re CafePress, FTC Docket No. 1923209 (recent data security complaint alleging unfair practices without any cost-benefit analysis); Fed. Trade Comm'n v. Wyndham Worldwide Corp., 799 F.3d 236, 256 (3d Cir. 2015) ("Wyndham does not argue that its cybersecurity practices survive a reasonable interpretation of the cost-benefit analysis required by § 45(n) . . . . [W]e leave for another day whether Wyndham's alleged cybersecurity practices do in fact fail, an issue the parties did not brief."); Fed. Trade Comm'n v. Neovi, Inc., 604 F.3d 1150, 1159 (9th Cir. 2010) (holding the FTC met its burden by producing an expert to show the challenged practice held no substantial benefit to consumers, while the defendant did not offer any evidence to challenge the expert testimony.)

opposing conclusions. Cost-benefit analysis is also indifferent to the uneven distribution of costs and benefits; highly concentrated and substantial harms could be excused by diffuse and minor benefits, so long as the total costs are not outweighed by the total benefits. As a result, which and whose costs and benefits are counted will almost entirely determine the outcome.

In a series of enforcement actions, the Commission has suggested that the relevant costs and benefits should be scoped to (1) the offending practice at issue, (2) the population at risk of harm, or, in a discrimination case specifically, (3) the costs that are strictly necessary to achieve the benefits. In what follows, we consider the motivation behind each approach to scoping, and we explore how they could be applied to discriminatory AI.

In re International Harvester is the first case to discuss cost-benefit analysis in any depth and the case that also introduced the Unfairness Statement into the FTC's record.[161] International Harvester was a manufacturer of agricultural equipment, and the case concerned the company's lack of effective disclosure alerting users of its tractors to the risk of fuel geysering, which can cause serious injury. To support its claim that the company's lack of sufficient disclosure was unfair, the Commission argued that the proper way to conduct a cost-benefit analysis was to compare the cost to those harmed by the lack of sufficient disclosure to the benefits—the cost savings—specifically due to not providing more effective disclosure.[162] Notably, the Commission did not weigh the costs of the ineffective disclosure against the benefit provided by the company's tractors overall or by its business altogether, which would have tipped the analysis in favor of benefits. Instead, it narrowly scoped its analysis—both costs and benefits—to the offending practice at issue. In later commentary, former Commissioner Ohlhausen argued that the Unfairness Statement itself dictates this interpretation.[163]

A similar maneuver can be observed in *In re Apple*,[164] a more recent case in which the FTC brought an unfairness claim related to Apple's practice

[161] In the Matter of Int'l Harvester Co., 104 F.T.C. 949 (1984).

[162] Ohlhausen, *supra* note 120, at 2019 ("[T]he principal tradeoff to consider was compliance costs—how much money had IHC saved by not notifying consumers about the risk of fuel geysering?").

[163] *Id.* at 2018–19 ("[T]he language of the Unfairness Statement is clear: '[T]he injury must not be outweighed by any offsetting consumer or competitive benefits that the sales practice also produces.' And '[t]he Commission . . . will not find that *a practice* unfairly injures consumers unless *it* is injurious in *its* net effects.'" (quoting UNFAIRNESS STATEMENT, *supra* note 101, at 1073 (emphasis supplied))).

[164] In the Matter of Apple Inc., A Corp.., 2014 WL 253519, at *1 (2014).

of storing a password for fifteen minutes after an in-app purchase. Apple failed to provide notice that the password remained active, and children were able to rack up bills for purchases that their parents never intended to authorize. Much like the charges in *International Harvester*, the FTC's claim was not that the practice of storing the password was unfair, but rather that the lack of notice was an unfair practice that created a risk of consumer harm. And once again, the Commission scoped its cost-benefit analysis to the ineffective disclosure. Reflecting on the case, Commissioner Ohlhausen explained:

> [I]t would have been incorrect for the Commission to compare the harm caused by the failure to notify consumers with the benefits of the design choice to use a fifteen-minute purchase window, or to compare the harm to the overall sales of the iPhone or iPad or total Apple sales more broadly. . . . [S]uch an approach would stack the deck against consumers, in favor of large companies. As long as a company's extensive line of products benefited consumers overall, the company would be free to inflict a significant amount of consumer harm with impunity.[165]

As Commissioner Ohlhausen's comments make clear, if the FTC were compelled to incorporate such a broad scope of benefits, it could easily lead to absurd situations in which the Commission would be kept from reaching even the most egregious cases of consumer harm. The FTC must therefore possess the freedom to scope the cost-benefit analysis in such a way that the benefit provided by a product or service overall does not excuse the harms caused by a particular feature.

Applying this principle to discriminatory AI, the FTC would not have to conduct its cost-benefit analysis by comparing the overall benefits of an AI product or service to the costs imposed on those consumers subject to discrimination. Instead, the FTC would compare the benefits of abstaining from efforts to reduce the discriminatory impact of the AI product and service against the costs imposed by the current level of discriminatory impact. Scoped in this way, the cost-benefit analysis might not present a serious a hurdle in seeking to address instances of discriminatory AI. There may very well be cases where the cost of reducing discrimination in an offending product or service translates into a benefit of equal or greater value for consumers as a whole. Where the harm is sufficiently costly, even seemingly expensive remedial actions could generate a value for consumers far greater

[165] Ohlhausen, *supra* note 120, at 2024.

than the expense involved. These would be the easy cases for the Commission. But of course, the reverse is possible as well. Some harms—even particularly costly harms—will affect such a small number of people that the benefits of addressing the harm will fail to outweigh the expense of doing so. As mentioned earlier, the lack of distributional considerations in cost-benefit analysis means that a practice that inflicts harm on a small number of people for the benefit of a much larger group would still be defensible because it enhances overall welfare. If reducing discrimination in AI products and services only benefits a small number of people, but the cost of doing so gets passed along to all consumers, then it is possible to imagine situations where consumers do not benefit overall, largely because many consumers, for whom a product or service might have been working perfectly fine already, enjoy no additional benefits for the additional cost.

But further examination of *Apple* shows this is not the correct interpretation for harms applied to small groups. In dissent, Commissioner Joshua Wright made such an argument. He compared the total costs to the consumers who would have not made purchases had they received effective notice with the benefits to the people who avoided the inconvenience of an unnecessary notice and paid a slightly lower price when Apple did not have to pursue expensive consumer research to determine how much more disclosure would have been necessary to avoid unauthorized purchases.[166] Because the victims constituted a "miniscule" fraction of overall users, he concluded that there is no evidence that the costs outweigh benefits.[167] But Commissioner Wright's reasoning would lead to precisely the outcome that Commissioner Ohlhausen rejects as absurd: that it is reasonable for a huge harm to befall a small group in exchange for a widespread but tiny benefit. The cost-benefit analysis requirement should not be understood to give companies license to, as Commissioner Ohlhausen says, "inflict a significant amount of consumer harm with impunity," and any analysis that counts benefits to a population many times the size of the injured population would set up such a result.

The second scoping approach that the Commission has adopted follows from that observation. In their joint statement for the *Apple* majority, Chair Edith Ramirez and Commissioner Julie Brill wrote:

---

[166] *Id.* at *13.

[167] *Id* at *12–14. He also argues that they cannot even complete the cost-benefit analysis because FTC staff did not conduct a study to determine what fraction of the victims would even have changed their behavior with more notice.

> [O]ur complaint focuses on conduct affecting Apple account holders whose children may unwittingly incur in-app charges in games likely to be played by kids. The proportion of complaints about children's in-app purchases as compared to total app downloads . . . sheds no light on the extent of harm alleged in this case. More fundamentally, the FTC Act does not give a company with a vast user base and product offerings license to injure large numbers of consumers or inflict millions of dollars of harm merely because the injury affects a small percentage of its customers or relates to a fraction of its product offerings.[168]

Combined with Commissioner Ohlhausen's comments, this suggests that, where the harmed population is a small subset of the total population that could be considered, the Commission could restrict the costs and benefits considered to only those experienced by the harmed group. Applying this principle to discriminatory AI, the Commission would not compare the benefits that all consumers enjoy when a company abstains from efforts to reduce the discriminatory impact of its AI product and service against the costs imposed on the subset of consumers impacted by the current level of discrimination. Instead, it would only consider the benefits of abstention (e.g., lower prices) that accrue to the subset of consumers who also experience the costs of discrimination.

Finally, in the recent case of *In re Napleton Automotive Group*,[169] the Commission addressed cost-benefit analysis in discrimination specifically, proposing yet another approach to scoping. In this case, the FTC brought an enforcement action against a franchise of car dealerships alleged to charge consumers for various add-ons without consumer consent, and—most important for our purposes—in a way that affected Black consumers more than white consumers.[170] In the discussion of cost-benefit analysis, Chair Khan and Commissioner Slaughter assert that the only defensible costs are those that are strictly necessary to achieve the benefit: "[A]ny purported benefit that can be achieved without engaging in the [discriminatory] conduct causing substantial injury is not countervailing, and does not overcome the costs associated with discrimination."[171] This version of scoping is not based on the practice at issue or the population at risk of harm. This version instead seemingly aligns the cost-benefit analysis with an idealized version of

[168] *Id.* at *26.
[169] 2022 WL 1044863.
[170] *Id.* at *1.
[171] *Id.* at *3.

disparate impact law, which holds there to be liability where a practice could have achieved the decisionmaker's goals equally well, but with less adverse impact on the plaintiffs.[172] Applying this principle to discriminatory AI should be relatively straightforward.[173]

Any of these scoping methods could allow the commission to call discriminatory AI unfair. Ultimately, the degree of flexibility that the Commission has—assuming it chooses to pursue these actions—will be determined by the amount of deference decisions are granted if and when they are challenged in court—and there is no obvious reason that the FTC would not be accorded deference when construing Section 5.[174] Thus, cost-benefit analysis should not present a barrier to an FTC determination that discriminatory AI is unfair.

B. Deceptive Practices

Our focus on unfairness should not be taken to suggest that regulating deceptive practices is unimportant to discriminatory AI. The Commission's authority to regulate deceptive practices is far-reaching, touching any type of material deception, made in any form, whether explicit or implicit.[175] And

---

[172] 42 U.S.C. 2000e-2(k).

[173] The operation of cases using this principle will not necessarily look the same as disparate impact law and may be more effective. Whereas in discrimination law, the plaintiff must prove the existence of such a less discriminatory alternative without access to information or many resources, the FTC may conduct an investigation beforehand, obtaining information from the defendant and determining the existence of less discriminatory alternatives before bringing an enforcement action. *See* Part I.E, *supra*. This functionally accomplishes a similar result sought by scholars who advocate a burden-shifting approach in discrimination law because of how poorly positioned plaintiffs are to prove less discriminatory alternatives. *See* Ajunwa, *Paradox*, *supra* note 5, at 1728; James Grimmelmann & Daniel Westreich, *Incomprehensible Discrimination*, 7 CALIF. L. REV. ONLINE 164, 170 (2017); Kim, *supra* note 5, at 921.

[174] *See* Rohit Chopra & Lina M. Khan, *The Case for "Unfair Methods of Competition" Rulemaking*, 87 U. CHI. L. REV. 357, 375 (2020) ("Rulemaking under "unfair methods of competition" is governed by the Administrative Procedure Act and is eligible for *Chevron* deference."); Justin (Gus) Hurwitz, *Chevron and the Limits of Administrative Antitrust*, 76 U. PITT. L. REV. 209, 250–58 (2014) (addressing common arguments as to why Chevron should not apply to Section 5); Daniel G. Lloyd, *The Magnuson-Moss Warranty Act v. the Federal Arbitration Act: The Quintessential Chevron Case*, 16 LOY. CONSUMER L. REV. 1, 26–27 (2003) (arguing that the FTC has an especially good case for *Chevron* deference, especially when issuing rules under Magnuson-Moss).

[175] *See* Hartzog & Solove, *supra* note 98, at 2247.

because the authority is less contested, cases are easier to prosecute, and the FTC has historically been more willing to pursue novel cases on deception grounds than unfairness.[176]

In the case of discriminatory AI, deceptive practices are most likely to take the form of products that are deemed to have some discrimination mitigation measure built in, and are thus marketed as "fair," "unbiased," "equitable," or some similar claim. This is a limited set of cases, as the majority of AI products and services on the market are not even tested for bias, let alone designed with discrimination mitigation measures that would justify marketing them as "fair". Over time, however, this may change. Awareness of the discriminatory potential of AI has exploded in the last decade, and fairness in algorithm solutions is already becoming a selling point in the vein of ethical consumerism.[177]

Deceptive fairness claims would likely take one of two forms. If a product is advertised as "fair" with reference to some verifiable standard, then the company producing that product can be held to that standard. Where the AI products and services do not function as the developers claim, the FTC can consider that a deceptive practice.[178] For example, some companies appear to be creating "fair" algorithmic solutions by constraining their models to satisfy the four-fifths rule.[179] If a company declares publicly that this is their metric, and then if they fail to live up to that standard, it is a clear deceptive practice. In addition, as part of the overall algorithmic accountability discourse, some scholars have proposed certification measures for fairness.[180] If certification—and especially self-certification—becomes an accepted method to establish fairness, the FTC will have the authority to

---

[176] *See e.g.*, Jillson, *supra* note 3 (warning companies not to exaggerate what their algorithms can or cannot do).

[177] *See* Daniel Greene, Anna Lauren Hoffmann & Luke Stark. *Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning*, PROCEEDINGS OF THE 52ND HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES (2019).

[178] Inioluwa Deborah Raji, I. Elizabeth Kumar , Aaron Horowitz & Andrew Selbst, *The Fallacy of AI Functionality*, PROCEEDINGS OF THE 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FACCT) 959.

[179] *See e.g.*, Manish Raghavan, et al., *supra* note 32, at 472.

[180] Tatiana Tommasi, Silvia Bucci, Barbara Caputo & Pietro Asinari, *Towards Fairness Certification in Artificial Intelligence* (unpublished manuscript) https://arxiv.org/abs/2106.02498; Gianclaudio Malgieri & Frank A. Pasquale, *From Transparency to Justification: Toward Ex Ante Accountability for AI* (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4099657.

check the veracity of those certifications.[181] Note that this does not suggest that a company can simply declare their models fair because they meet an arbitrary standard—an intolerable standard would trigger an unfairness claim as surely as no standard at all. But the deception claim would be limited to ensuring that the company meets their representations.

The second form of deception claim could come about if a company declared that its product is fair or unbiased and did not provide any reference point. Here, to call this deceptive, the Commission would have to have its own reference point for a fairness baseline. If the company made such claims but made no effort to test or audit their own systems, the Commission will not have a hard time calling that deception. But if a company has some mitigation measures in place, and calls its product fair, then for this be deceptive, the FTC would have to rule that it does not meet a standard of fairness that the company knew—or perhaps should have known—applies in order to justify such a claim.

This last case, then, merges the deception and unfairness issues. The conduct can be described alternately as deception (the company represented its product as fair and that was untrue because it failed to meet some minimum fairness threshold of which the company had notice) or unfairness (the company created a fair algorithm according to its own metric, so it did not deceive the public, but its internal definition of fairness does not meet the minimum threshold required to be fair). Either way, there would need to be a minimum threshold of fairness that the Commission envisions. Thus this last example ties the deception claim to the standards of unfairness, which echoes many of the FTC's enforcement actions in which they have brought both types of claims.

C. Major Questions Doctrine

Aside from statutory authorization, we must consider an additional hurdle for FTC intervention: the major questions doctrine. According to *West Virginia v. EPA*, in "certain extraordinary cases" of great "economic and political significance," courts should apply a clear statement rule to determine whether an agency action was in the scope of the authority they received from

[181] *Cf.* Robert R. Schriver, Note, *You Cheated, You Lied: The Safe Harbor Agreement and Its Enforcement by the Federal Trade Commission*, 70 FORDHAM L. REV. 2777, 2791–92 (2002) (discussing the FTC's ability to enforce data privacy self-certifications under its deceptive practice authority).

Congress.[182] The precise contours of the doctrine are not clear, as it is still emerging.[183] The Court did not provide a test in *West Virginia*, so it is hard to know how to apply the doctrine.[184]

The doctrine arose in the context of *Chevron* deference, where the concern of courts is whether a Congressional grant of authority to an agency is ambiguous, and if so, whether the agency interpretation is reasonable.[185] But in more recent years, the Court has stopped treating the doctrine as if it is about resolving ambiguities, and instead treated it as a question prior to *Chevron* analysis.[186] In *King v. Burwell*, the Court held that *Chevron* deference did not apply to an Internal Revenue Service interpretation of a tax credit provision of the Affordable Care Act because it involved "billions of dollars in spending each year and affect[ed] the price of health insurance for millions of people,"[187]—even though it was the IRS interpreting the tax code. No longer was the doctrine just another interpretive tool, rather it was a separate and prior injunction against dramatic and costly agency action.[188] Over the last two years, the Court has further expanded the doctrine, striking down an EPA policy in *West Virginia* and using it to block agency actions twice during

[182] W. Virginia v. Env't Prot. Agency, 142 S. Ct. 2587, 2594–95; *id* at 2616 (Gorsuch, J., concurring) (describing the holding as a clear-statement rule).

[183] *See e.g.*, Daniel Deacon & Leah Litman, *The New Major Questions Doctrine*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4165724 (manuscript at 3); Cass R. Sunstein, *There Are Two "Major Questions" Doctrines*, 73 ADMIN. L. REV. 475, 480-82 (2021) (arguing that the Court has deployed two different formulations of the doctrine).

[184] *See generally* Deacon & Litman, *supra* note 183.

[185] *See* FDA v. Brown & Williamson Tobacco Corp., 529 U.S. 120, 132–33 (2000). The doctrine has usually been applied to the question of whether the statute unambiguously grants authority in the first place (*Chevron* step one), but it may also apply in the context of whether the agency interpretation was reasonable (*Chevron* step two). *See* Utility Air Regulatory Group v. Envt'l Protection Agency, 573 U.S. 302 (2014).

[186] Alison Gocke, *Chevron's Next Chapter: A Fig Leaf for the Nondelegation Doctrine*, 55 U.C. DAVIS L. REV. 955, 978 (2021) (calling it an "all-out assault" on *Chevron*).

[187] *King v. Burwell*, 576 U.S. 473, 485 (2015). *See* Christopher J. Walker, *What King v. Burwell Means for Administrative Law* https://www.yalejreg.com/nc/what-king-v-burwell-means-for-administrative-law-by-chris-walker/ (noting the novelty of *King*'s holding for administrative law).

[188] Deacon & Litman, *supra* note 183 (manuscript at 23) ("Rather than resolving an ambiguity or even placing a thumb on the scale as the Court attempts to discern the meaning of a statute, the new major questions doctrine functions as a kind of carve out to an agency's authority.").

the COVID-19 pandemic.[189]

After *West Virginia*, there are two major questions about major questions doctrine: What counts as "major" and how clear must a statement by Congress be to satisfy the Court? Daniel Deacon and Leah Litman identify "three indicia of 'majorness,' in addition to the costs imposed by the agency policy," on which the court relies: political significance or controversy, novelty of a policy, and essentially a slippery slope argument about what approving such a broad policy could lead to in the future.[190] Then, once the Court finds a policy to implicate a major question, the Court in theory applies a clear-statement rule, but as Deacon and Litman put it, "[e]ven broadly worded, otherwise unambiguous statutes do not appear good enough when it comes to policies the Court deems 'major.'"[191]

If we take seriously the language of *West Virginia*, the FTC might not actually have a problem. If any agency could justifiably take refuge in the breadth of the authority expressly delegated to it by Congress, it is the FTC. The FTC's authority is based on the incredibly general phrase: "unfair or deceptive acts or practices in or affecting commerce."[192] Congress clearly intended the FTC to be able to regulate large swaths of the economy and to adapt to previously unforeseen harms, and the Supreme Court has on several occasions noted how expansive the authority is.[193] Ultimately, the expansiveness and adaptability is the whole point of the unfairness authority. In theory, this would suggest that the FTC is on solid footing with respect to

---

[189] Alabama Ass'n of Realtors v. Dep't of Health & Hum. Servs., 141 S. Ct. 2485 (2021) (blocking the Center for Disease Control's extension of an eviction moratorium); Nat'l Fed'n of Indep. Bus. v. Dep't of Lab., Occupational Safety & Health Admin., 142 S. Ct. 661 (2022).

[190] Deacon & Litman, *supra* note 183 (manuscript at 3).

[191] *Id.*

[192] 15 USC § 45(a).

[193] Fed. Trade Comm'n v. Sperry & Hutchinson Co., 405 U.S. 233, 239–40 (1972) ("When Congress created the Federal Trade Commission in 1914 and charted its power and responsibility under § 5, it explicitly considered, and rejected, the notion that it reduce the ambiguity of the phrase 'unfair methods of competition' by tying the concept of unfairness to a common-law or statutory standard or by enumerating the particular practices to which it was intended to apply."); Atl. Ref. Co. v. F.T.C., 381 U.S. 357, 367, 85 S. Ct. 1498, 1505 (1965); ("The Congress intentionally left development of the term 'unfair' to the Commission rather than attempting to define 'the many and variable unfair practices which prevail in commerce.' . . . In thus divining that there is no limit to business ingenuity and legal gymnastics the Congress displayed much foresight.") Fed. Trade Comm'n v. Bunte Bros., 312 U.S. 349, 353 (1941) ("Unlike the relatively precise situation presented by rate discrimination, 'unfair competition' was designed by Congress as a flexible concept with evolving content.")

major questions.

But to treat the expansion of this doctrine as a normal exercise of legal reasoning would be naïve and disingenuous. The Court isn't really *doing law* here.[194] Both of the COVID-19-era cases were non-merits cases on the "shadow docket," so neither of them were fully reasoned, and the Court played fast and loose with precedent in *West Virginia*, treating the shadow docket cases as precedential, while providing little in the way of reasoning to support the changes to the doctrine.[195] As Justice Kagan notes in her *West Virginia* dissent, these cases have become law-by-judicial eyebrow raise,[196] and as Mark Lemley notes, the case is one part of the Court's broader effort to "strip[] power from every political entity except the Supreme Court itself."[197]

If the Supreme Court decides that the FTC should not be regulating discriminatory AI, then there's not much the FTC can do to stop that determination. And there is good reason to believe such a move by the Commission would rub the Court's conservatives the wrong way. For one, they're famously hostile to discrimination law and its goals. Even more concerning is the centrality of the Court's "antinovelty" line of reasoning, in which the Court displays skepticism of novel government action on the grounds of novelty itself.[198] At several points in *West Virginia*, the Court and the concurrence observe that the EPA has never sought to regulate under Section 111(d) of the Clean Air Act before—a regulatory "little-used

[194] W. Virginia v. Env't Prot. Agency, 142 S. Ct. 2587, 2634 (2022) (Kagan, J., dissenting) ("The majority claims it is just following precedent, but that is not so. The Court has never even used the term 'major questions doctrine' before.")

[195] *See* Blake Emerson, The Real Target of the Supreme Court's EPA Decision, SLATE https://slate.com/news-and-politics/2022/06/west-virginia-environmental-protection-agency-climate-change-clean-air.html ("Chief Justice John Roberts makes no serious effort to defend his assertion that EPA exercised a 'power beyond what Congress could reasonably be understood to have granted.'")

[196] *W. Virginia*, 142 S. Ct. at 2636 ("The eyebrow-raise is . . . a consistent presence in these cases")

[197] Mark A. Lemley, *The Imperial Supreme Court*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4175554 (forthcoming).

[198] *See generally* Leah M. Litman, *Debunking Antinovelty*, 66 DUKE L.J. 1407 (2017) (critiquing antinovelty reasoning in the context of statutes' constitutionality).

backwater"[199]—as if the novelty is self-discrediting.[200] This is the eyebrow raise in action: gesture at the novelty, expecting the reader to nod along in agreement at the self-evident absurdity of the government action. As a result, it is not hard to imagine the Court reacting to the FTC's decision to venture into discriminatory AI in the same way. Legally speaking it's nonsense, but practically, it is not hard to imagine just such an outcome if the FTC pursues these cases and is challenged.

Once we acknowledge this possibility, however, it is not clear that it changes much for the FTC. Perhaps blowback will affect other Commission priorities, and that fear is rational. But if the Court wants to cut agency authority, it can find a vehicle from another agency—after all, *West Virginia* was a case about the EPA, yet it could easily affect the FTC. Pursuing these priorities won't change that. Meanwhile, a failure to act due to speculative concerns would come at the cost of the continuation of all the algorithmic harms the FTC would have sought to address. That does not seem a good tradeoff for a Commission interested in addressing new kinds of harms in any context.

Ultimately, the implication of major questions doctrine is that the FTC's position is more precarious, but given that the choice is merely to stop reacting to new consumer threats—essentially giving up on the entire mission of the FTC—or to risk blowback, we believe the best course of action is just to proceed and meet the challenge head on.

III. ADAPTING THE DATA SECURITY MODEL

Assuming the FTC proceeds, the remaining question is how the FTC's approach should work in practice. How should the Commission determine what constitutes unfair discrimination? There are two main options: a case-by-case approach that mirrors common law development or the creation of rules to define practices that are unfair.

---

[199] *West Virginia*, slip op. at 26; *accord id.* at 6, (referring to "Section 111(d) as an "obscure, never-used section of the law.'"), *id* at 15, (Gorsuch, J., concurring) (same), *id* at 22 (majority op.) (referring to "an unbroken list of prior Section 111 rules" as evidence that the Section 111 rule at issue, which different in kind, was invalid), *id.* at 27 (referring to the "the regulatory writ EPA newly uncovered")

[200] *See id.* (Kagan, J., dissenting) ("The majority's decision rests on one claim alone: that generation shifting is just too new and too big a deal for Congress to have authorized it in Section 111's general terms.").

Canvassing recent literature on discriminatory AI, many scholars have proposed solutions to help regulate AI, but none offer a framework that the FTC can apply to determine the merits of when something is unfair. For example, looking to reform discrimination law, several scholars have proposed burden-shifting approaches.[201] Ifeoma Ajunwa, for example, proposes to create a "discrimination per se" standard to reform Title VII, which "would shift the burden of proof from the plaintiff to the defendant (employer) to show that its practice is non-discriminatory."[202] But when it comes to determining what substantively counts as discrimination per se, Ajunwa would defer to a common law approach.[203] Gianclaudio Malgieri and Frank Pasquale argue for a pre-deployment self-certification approach.[204] From the FTC's perspective, this would certainly change the landscape, as it would functionally entwine deception and unfairness; a system that fails to meet the standard would come with a false declaration that it meets the standard.[205] While this would make the FTC's life easier when bringing a case, it still does not provide an answer for what actions are or are not unfair. Paul Ohm's proposal for "forthright code" would ratchet up the deception standard to require affirmative disclosure of harm.[206] This would similarly bring deception and unfairness closer together but does not provide an answer on the merits. Even a strict liability approach would not be concrete where the injury is unspecified.

This leaves two possible approaches to defining unfairly discriminatory AI: either the common law approach that many scholars implicitly or explicitly rely on or issuing affirmative rules defining unfairness. In this Part, we are focused on the common law approach, but the Commission does have the authority to pass "trade regulation rules" that affirmatively define practices that are unfair or deceptive.[207] We discuss this very briefly at the end of the Part.

The reason we focus here on direct enforcement is that over the last

---

[201] Ajunwa, *Paradox*, *supra* note 5, at 1728; Grimmelmann & Westreich, *supra* note 173, at 170; Kim, *supra* note 5, at 921.

[202] Ajunwa, *Paradox*, *supra* note 5, at 1728.

[203] *Id.* So would the various scholars discussing flavors of negligent discrimination. *See* Páez, *supra* note 5.

[204] *See* Malgieri & Pasquale, *supra* note 180.

[205] Cf. Schriver, supra note 181, at 2791–92 (2002) (discussing data privacy self-certifications under its deceptive practice authority).

[206] Paul Ohm, *Forthright Code*, 56 HOUS. L. REV. 471 (2018).

[207] 15 U.S.C. § 57a.

two decades, the FTC has developed just such an approach to a different problem: regulating data security. The Commission has brought direct actions against companies that have failed to provide reasonable data security under the theory that such a failure is an unfair act or practice, drawing on negligence principles to define inadequate security.[208] The FTC brings enforcement actions against the worst actors, which everyone agrees should fall below any reasonable threshold of data security, and almost invariably settles with them ending the enforcement action with a public consent decree.[209] The Commission then uses the consent decrees, in conjunction with industry best practices and published guidance, to develop a body of knowledge about what constitutes inadequate security.[210] Armed with this knowledge, businesses can understand what constitutes the outer bounds of reasonableness, while still retaining a large degree of flexibility such that they can adapt their approach to their particular size and security needs. While there are some important differences between this approach and a "true" common law—most notably that the FTC is both prosecutor and judge, and it can in theory change the standards unilaterally[211]—the use of an incremental style of reasoning to define the merits is the key point here and is the reason that this method has been referred to as a sort of common law.[212]

We are interested in this approach because it is now a well-worn path for the Commission, and it turns out that there are a surprising number of parallels between data security and algorithmic discrimination, all of which we discuss below.

A. A Risk Mitigation Approach

The first parallel between algorithmic discrimination is conceptual. Both are fundamentally about risk recognition and mitigation, where liability should attach where defendants fail to do enough to mitigate harm or risk, rather than those who simply fail to prevent any harm at all. In both cases, there is likely no possibility of total prevention of what amounts to an inescapable background risk.

---

[208] *See generally* William McGeveran, *The Duty of Data Security*, 103 MINN. L. REV. 1135, 1194 (2019); Solove & Hartzog, *supra* note 9, at 648.

[209] Solove & Hartzog, *supra* note 9, at __.

[210] *Id.* at 676; McGeveran, *supra* note 208, at 1193–95.

[211] *See* Justin (Gus) Hurwitz, *Data Security and the FTC's Uncommon Law*, 101 IOWA L. REV. 955, 984–87 (2016).

[212] Solove & Hartzog, *supra* note 9, at __.

The parallel here may not be obvious, as the source of background risk is quite different in each case. The FTC's approach to data security recognizes that the hackers that steal people's data are nearly impossible to find, and if located are usually outside the United States jurisdiction. In addressing data security by assigning responsibility to those companies who suffer data breaches, the Commission treats the fact of hackers' existence as a background level of risk that must be mitigated by businesses as stewards of consumer data. No one knows when and where a company will suffer a breach, and it is not entirely the company's fault when it happens, but this does not mean the company bears no responsibility to make the intrusion more difficult and less costly to consumers. Because this is the FTC's theory, it makes sense to adopt a responsibility-focused approach on the question of security *practices*, not liability based on the final result. The question becomes not "Was the company breached?" but "Did the company do enough to mitigate the background risk and harms of potential breaches?"

The parallel in discriminatory AI comes from a recognition of extant disenfranchisement of subordinated groups. It is overwhelmingly likely, if not inevitable, that allocative algorithmic decision systems will evince some degree of bias in decisions whenever applied to people in different demographic groups. But there is no way to fully debias these systems. Appeals to accuracy do not work. An inherent aspect of predicting the future from past data is the absence of a ground truth by which to arbitrate the accuracy of different models. Because accuracy is therefore in a real sense undefined, every predictive model requires some element of persuasive justification for its use and part of that persuasive justification is how discriminatory it turns out to be, and in what ways. But just as there is no one answer on the accuracy question, there is no one answer on the bias question. There are many reasonable yet often incompatible ways to measure bias,[213] and therefore by some metric, every allocative decision will evince *some* bias. Hence, in allocative decisions, there is no such thing as zero bias. It is a persistent background risk that the FTC cannot do anything about, so if the commission seeks to hold companies accountable for it, the question must become one of whether they did enough to mitigate the risk of or harm from

[213] Sorelle A. Friedler, Carlos Scheidegger & Suresh Venkatasubramanian, *The (Im)possibility of Fairness: Different Value Systems Require Different Mechanisms for Fair Decision Making* 64 COMM. ACM 136 (2021); Jon Kleinberg, Sendhil Mullainathan & Manish Raghavan, *Inherent Trade-Offs in the Fair Determination of Risk Scores* PROCEEDINGS OF INNOVATIONS IN THEORETICAL COMPUTER SCIENCE (ITCS), (2017).

the likelihood of discrimination.

The case of quality-of-service harms is a little different in that it should be possible, at least in theory, to require that products work as well for underrepresented groups as the dominant groups. As the authors of the Gender Shades study showed, these products are fixable to a large degree. When they confronted the companies that they studied with their results, many of them went back and fixed their products.[214] But at what point is the product "fixed"? If the FTC wants to hold companies liable for even *de minimis* differential quality of service, then unfairness becomes essentially strict liability. But assuming companies can get partway there at less cost, there may come a point where the effort to make a product perform with perfectly equal accuracy across groups requires a great deal more research and development cost. In that case, the FTC will be forced to apply cost-benefit analysis and ask whether the company did enough to get most of the way there. Thus this question, too, sounds in negligence: Did the company do enough to satisfy their duty of nondiscrimination in their products?

Finally, due to the nature of how representational harms occur, they are often more difficult to foresee. But if a company fails to fix their system—if Microsoft allowed Tay to continue operating after transforming into a digital hatebot overnight[215]—then, again assuming this is a legitimate injury the FTC can address,[216] it would certainly seem fair game to call the failure to address it unfair.

Thus, there is something fundamentally similar about data security and discrimination. The targets of FTC enforcement in both cases are operating in a world where there is a persistent risk of the relevant harm for which it would be unreasonable to hold companies fully responsible. Thus, the question is transformed into one in which we separate the fact of injury from the question of responsibility. There may be different ways to do that, but a common one is negligence, and that is the basis for the FTC's data security approach. It should work similarly with discrimination.

It is also worth noting that this idea of discrimination as negligence—where defendants have essentially a duty to prevent or mitigate a background

---

[214] Matt O'Brien, *Face Recognition Researcher Fights Amazon Over Biased AI*, ASSOCIATED PRESS (April 3, 2019) ("Months after her first study . . . all three companies showed major improvements.")

[215] James, Vincent, *Twitter Taught Microsoft's AI Chatbot to Be a Racist Asshole in Less than a Day*, THE VERGE (Mar. 24, 2016) https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist.

[216] *See* Part II.A.1, *supra*.

risk of discriminatory harm—is not new and not particular to algorithmic harms. Disparate impact law asks whether a decisionmaker used a facially neutral decision tool with a discriminatory effect, but nonetheless had a good enough reason for it. Formally, this is a burden-shifting test about "business necessity" and less discriminatory alternatives. But as David Oppenheimer famously observed three decades ago, many courts functionally treat disparate impact as a negligence inquiry.[217] This makes sense. Burden shifting, in practical terms, can either be a strong version where almost nothing is justified or a weak version where almost anything rationally related to the desired outcome passes muster.[218] As written, such a test flip-flops unstably between strict liability and no liability, neither of which is satisfactory because there is no moment to ask what would make a decisionmaker responsible for the harm.[219] Negligence gets more directly at the fault question that a discrimination suit seeks to answer. Discrimination law then becomes a determination about whether a defendant is responsible for compounding the injustice[220] or failing to mitigate it.[221] Though certainly not a universally accepted view of discrimination law, the idea of discrimination as negligence is not restricted to algorithms.

B. The Flexibility of a Common Law Approach

The Commission's common law approach to data security also offers a few concrete benefits that are mirrored in the discrimination context. The first is its ability to be sensitive to context. The appropriate level of security is highly contextual. It depends on the type of data being stored, the quantity of data stored, the business that the company is in, how much reason the company has to believe they are likely to be hacked, and the resources a company has available to devote to security. The FTC's approach sets a

---

[217] David Benjamin Oppenheimer, *Negligent Discrimination*, 141 U. PA. L. REV. 899 (1993); *see also* Richard Thompson Ford, *Bias in the Air: Rethinking Employment Discrimination Law*, 66 STAN. L. REV. 1381, 1389 (2014) (comparing discrimination to negligence); Noah D. Zatz, *Managing the Macaw: Third-Party Harassers, Accommodation, and the Disaggregation of Discriminatory Intent*, 109 COLUM. L. REV. 1357, 1364 (2009) (arguing that third-party harassment cases can functionally be seen as negligence-based); Páez, *supra* note 5.

[218] Barocas & Selbst, *supra* note 5, at 704–05.

[219] *See* Part I.C, *infra*.

[220] Deborah Hellman, *Sex, Causation, and Algorithms: How Equal Protection Prohibits Compounding Prior Injustice*, 98 WASH. U. L. REV. 481 (2020).

[221] *See* Benjamin Eidelson, *Patterned Inequality, Compounding Injustice, and Algorithmic Prediction*, AMERICAN JOURNAL OF LAW AND EQUALITY 252 (2021).

minimum absolute baseline as a calibration point, but then offers individual companies leeway to tailor their operations to their particular circumstances, for which they may be the best judge.[222] Reasonableness admits of a world where there are multiple permissible solutions to the problem that are defensible, as long as the overall result satisfies some minimum baseline. In a world where the right result is so fact-laden, it is quite difficult to take a more prescriptive, rule-based approach.[223]

The same context-sensitivity is necessary for algorithmic fairness.[224] The specific determinations of whether something is or is not discriminatory will depend on the type of harm, the specific sector it is being deployed in, what decisions and tradeoffs were made in design, what attempts were made to address the discriminatory harm, what remedies might have been practically available, and a host of other considerations. An approach focused on reasonableness will allow for leeway in attempts to solve the discriminatory harms that are sure to arise, while a minimum baseline would at least give some teeth to the requirement.

The second benefit is related, and that is that the approach is based in negative determinations. In short, it is much easier in the face of uncertainty and disagreement to say when a specific set of facts does *not* satisfy a threshold of risk mitigation than to offer a prescriptive set of rules for when activities will satisfy it. That is, it's easier to reject practices that are *unfair* than have a rule dictating which practices are *fair*. It is perhaps not coincidental that the language of the FTC's mandate prohibits unfair practices rather than permits fair ones.[225]

---

[222] Solove & Hartzog, *supra* note 9, at 661; Justin (Gus) Hurwitz, *Response to McGeveran's the Duty of Data Security: Not the Objective Duty He Wants, Maybe the Subjective Duty We Need*, 103 MINN. L. REV. HEADNOTES 139, 143 (2019).

[223] Gus Hurwitz objects to the description of the FTC's actions *as* negligence because of this subjectivity, arguing that it "roughly corresponds to a subjective reasonableness standard backed by a *per se* negligence standard for extremely objectionable conduct." Hurwitz, *supra* note 223, at 143.

[224] *See* Doaa Abu Elyounes, *Contextual Fairness: A Legal and Policy Analysis of Algorithmic Fairness*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3478296; Andrew D Selbst, danah boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi, *Fairness and Abstraction in Sociotechnical Systems*, PROCEEDINGS OF THE 2019 CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 59.

[225] This is also in keeping with Ben Hutchinson and Meg Mitchell's suggestion that *technical* work be oriented around *un*fairness rather than fairness, as attempts to develop

Whether for practical or legal reasons, the FTC's approach results in bringing enforcement actions against the "worst practices" in data security. There are certain accepted ways to mitigate data security risk, and if a company does not even try, then they are so far from the line it is easy enough to call them out and bring an action—much easier than trying to say where the line would have been in the abstract. Same with discrimination. Take allocative harms, where recent research has shown, for example, that it is often possible to develop many different machine learning models that each exhibit the same degree of accuracy overall, yet differ in the degree to which they result in disparities in outcomes across groups.[226] Those developing or procuring algorithmic employment assessments, credit scoring models, and tenant screening software, among many other such tools, will frequently find that they do not need to forgo a commitment to maximizing the accuracy of their decisions to reduce disparities in hiring, lending, and leasing rates.[227] Now that this is a known possibility, firms should be expected to make reasonable efforts to figure out if such models exist, and a failure to even try should be seen as an unfair practice.

The situation is much the same with quality-of-service harms. It is now widely understood that the standard way of evaluating the overall performance of machine learning models can easily conceal that such models may perform much less well when applied to certain demographic groups than others. A model reported to have 95% overall accuracy could be accurate 100% of the time for a majority group that constitutes 95% of the population and 0% of the time for a minority group that constitutes 5% of

methods for producing fair machine learning models have been much less successful and productive than attempts to develop techniques for determining when machine learning models are unfair. *See* Ben Hutchinson and Margaret Mitchell, *50 Years of Test (Un)fairness: Lessons for Machine Learning*, PROCEEDINGS OF THE 2019 CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 49, 50, 56.

[226] Emily Black, Manish Raghavan, and Solon Barocas, *Model Multiplicity: Opportunities, Concerns, and Solutions*, PROCEEDINGS OF THE 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY (FACCT) 850.

[227] This observation speaks directly to the idea in disparate impact doctrine that decision makers should face liability if they fail to adopt an alternative business practice that serves their goals equally well but generates a less disparate impact. It also echoes Chair Khan and Commissioner Slaughter's argument that firms cannot assert that a business practice provides an overall benefit to consumers if it generates a disparate impact that could have been avoided at no cost. *See* notes 169-173 *supra* and accompanying text.

the population.[228] Standard performance metrics would fail to reveal this extreme disparity. As a result, it has become increasingly common to expect firms to perform disaggregated evaluations: breaking apart the reported performance of machine learning models by group.[229] While it might turn out, in some cases, that it is unreasonably difficult for firms to conduct a thorough disaggregated evaluation or to address revealed performance disparities (due to the difficulty of obtaining the necessary demographic information by which to disaggregate the evaluation, the cost involved in performing the evaluation or remedying the disparity, or other technical limitations),[230] it is unacceptable for firms to fail to even perform such assessments. Of course, as the cost of conducting an assessment and remedying any revealed performance disparities drops, these costs may become a less reasonable justification for the failure to uncover and address any gaps in performance, rendering the continued use of a model unfair.

The final point in the previous paragraph highlights how a reasonableness standard with a minimum baseline can evolve over time. With data security, standards have changed over time. Technology evolves; the threats, the available responses, and the cost of responses change over time. We have a better sense of which responses work and which don't, and how to evaluate risk tradeoffs. As Gus Hurwitz has observed, "[t]he duty of data security . . . is . . . akin to keeping apace of advancements in the field of cybersecurity, of constantly monitoring, updating, testing, and replacing the locked box that data is secured into."[231]

Once again, the same is true here. Right now we do not know all the failure modes of AI systems, including which types of decisions lead to particularly bad outcomes for discrimination, but we are learning over time how to test for and mitigate discriminatory harms.[232] Just like in the data security context, technology to perform some of this measurement and mitigation will likely develop, standardize, and become less expensive over time. As possible responses become less expensive, they should and will become part of the expected suite of mitigations.[233]

Inherent in this approach is also a foreseeability limitation: firms may

---

[228] Moritz Hardt, How Big Data Is Unfair, MEDIUM (Sept. 26, 2014), https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de.

[229] *See* Buolamwini & Gebru, *supra* note 25; Barocas, et al., *supra* note 149.

[230] *See* sources cited *supra* note 149.

[231] Hurwitz, *supra* note 222, at 148.

[232] Selbst, *supra* note 17, at 121.

[233] Black et al., *supra* note 226; Raghavan, et al., *supra* note 32.

only be held responsible for harms that they are reasonably able to anticipate. On the one hand, this means that the FTC may struggle to declare unfair any business practice that results in harms that were difficult or impossible to foresee. This limitation will be especially relevant to representational harms, which can manifest in a wide variety of ways, many of which will be difficult to fully anticipate. Take the example of Google Photos tagging an image of a Black person with the label 'gorilla'.[234] While this controversy has fostered broad recognition of the harm that might be caused by image tagging systems that mislabel people as animals—especially mislabeling a Black person as a gorilla—it remains very difficult to imagine the full range of labels whose misapplication to a particular type of image might be similarly demeaning. Given the remarkably expressive capacity of language and the seemingly infinite possible variation in the composition of photos, enumerating all the label-image pairs that might be widely perceived as harmful is an enormously challenging, likely impossible task.[235] On the other hand, a regulatory regime based on foreseeability also means that as the range of foreseeable harms continues to grow, firms will be expected to address more of them. Thus, not only would Google be expected to address the foreseeable harm of mislabeling people as animals—and specifically mislabeling Black people as gorillas—it would also be expected to address the foreseeable harm of returning pornographic images in response to a search query for "black girls".[236] Once harms along these lines are no longer unforeseeable, the FTC could ask, in negligence terms, whether enough was done to test for and remedy the problem such that Google bears no liability.

As a general matter, it is much easier to use a common law approach to extrapolate whether a particular practice is so foreseeable that efforts should have been made to address it than it is to proactively predict all possible harms. An approach that allows the FTC to start with the worst of the worst and slowly ratchet up the baseline will also give companies ample time to stay ahead of their responsibilities and will allow the law to adjust as the set of known harms expands and our understanding of the issues evolve.

---

[234] Barr, *supra* note 76.

[235] Solon Barocas, Su Lin Blodgett, Jared Katzman, Kristen Laird, Morgan Scheuerman, & Hanna Wallach, *Representational Harms in Image Tagging*, WORKSHOP ON BEYOND FAIRNESS: TOWARDS A JUST, EQUITABLE, AND ACCOUNTABLE COMPUTER VISION, CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2021; Angelina Wang, Solon Barocas, Kristen Laird & Hanna Wallach, PROCEEDINGS OF 2022 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 324.

[236] *See* Noble, *supra* note 71.

Of course, such an approach may also create perverse incentives: if firms are only held responsible for foreseeable harms, they may purposefully avoid investing the effort to discover more about the harms that their systems could bring about.[237] This, in turn, may place much of the burden for uncovering and raising awareness of such harms on the people who directly experience them or other outsiders, such as advocates, journalists, and researchers, who are generally less well-resourced and less well-positioned to undertake such investigations.

C. Parallel Challenges to the FTC's Authority

Despite its utility, the FTC's approach to data security has proved contentious, with businesses furious at the lack of clearly articulated data security rules—or less charitably, that the FTC dared to regulate data security at all. Ultimately, the FTC has survived two major challenges to its authority to regulate data security in this manner, with the Third Circuit ruling in the Commission's favor in *FTC v. Wyndham*,[238] and the Eleventh Circuit avoiding the issue in *LabMD v. FTC.*[239] Though neither of these cases was a full-throated endorsement, the FTC's approach continues unchanged. This suggests that a similar approach to discrimination should also be legal. But just as important for our purposes is that despite bitter disagreement over the legality of the data security model, one issue was never even raised: Whether the fact that only a small population bore the costs of bad data security rendered the cost-benefit analysis problematic for the Commission.

In *Wyndham*, the hotel chain raised three challenges to the data security model of regulation, all of which were rightly rejected by the court. First, it argued that unfairness means more than what is included in the language of Section 5, that it must include unscrupulous or unethical

[237] This is a major difference between an unfairness regime limited by reasonable foreseeability, and a requirement for affirmative investment in research, like an impact assessment regime. For example, the Algorithmic Accountability Act of 2022 would also have sought to regulate AI through the FTC, but largely through an impact assessment regime, violation of which would independently be considered an unfair practice. Algorithmic Accountability Act of 2022 S.3752, 117th Cong. §§ 3(b), 9 (2022).

[238] 799 F.3d 236, 247 (3d Cir. 2015) (holding that bad cybersecurity can be included in unfairness); *id* at 255 (holding that the FTC's reasonableness approach and lack of fixed standards did not violate principles of fair notice).

[239] 894 F.3d 1221, 1231 (11th Cir. 2018) (assuming without deciding that negligent data security constitutes an unfair act or practice).

behavior.[240] Wyndham argued that it was not unscrupulous—in fact, because it got hacked it was really a victim![241] The Court rejected this argument, noting that unfairness was envisioned by Congress as an open-ended and flexible concept that was meant to be adapted to the problems of the day, and that cybersecurity was not outside the plain meaning of "unfair."[242] This line of reasoning applies equally well to discriminatory AI: Unfairness is flexible and adaptable and should pose no problem here.

The second challenge was a statutory argument, claiming that Congress had granted the FTC limited authority over cybersecurity in other specific contexts, including children's privacy, finance, and credit, with the implication that Congress would not have done so if the Commission already had such authority.[243] The Court rejected this argument as well, recognizing that the statutes granted authority to the FTC to regulate with different kinds of procedures, so the fact that there may have been overlapping conceptual authority should not be read to preclude a more general approach under Section 5.[244] To the extent we should take anything away from an argument that is rather specific to the data security context, it is that the existence of other anti-discrimination statutes with their own procedures does not hinder the FTC's use of Section 5 in this context.

Finally, Wyndham argued that the Commission's reasonableness approach violated principles of fair notice. The Court performs some legal jiujitsu here, trapping Wyndham in its own argument. Wyndham essentially argued that the FTC did not make a concrete enough determination and was thus owed no deference.[245] But the Court pointed out that if Wyndham is correct that the FTC is due no deference, the implication is not that the conduct was not unfair (the result Wyndham wanted), but that the *court* must decide in the first instance. And when it comes to courts, there is no "ascertainable certainty" standard—courts decide reasonableness questions all the time without vagueness problems.[246] Thus, "Wyndham was not entitled to know with ascertainable certainty the FTC's interpretation of what cybersecurity practices are required by § 45(a). Instead, the relevant question in this appeal is whether Wyndham had fair notice that its conduct *could* fall

[240] *Wyndham*, 799 F.3d at 244.
[241] *Id.* at 246.
[242] *Id.* at 243, 247.
[243] *Id.* at 247.
[244] *Id.* at 248.
[245] *Id.* at 252.
[246] *Id.* at 255.

within the meaning of the statute."[247] And it had such notice.

The Court did not entirely let the FTC off the hook. Instead, it continued with the question of whether Wyndham had fair notice that its conduct *could* have been considered unfair, treating the vagueness claim as an as-applied challenge.[248] It is at this point that the Court notes just how bad the facts are for Wyndham—the hotel had been hacked three times, and had implemented next to no security measures even after the first two incidents.[249] The Court suggested that it is possible that one day the FTC could bring a case that is close enough to the line that a litigant might not have fair notice, but told Wyndham that its case just isn't a close one.[250]

The *Wyndham* analysis applies directly to a future discriminatory AI case. If the FTC declares that it intends to address discriminatory AI through Section 5, publishes guidance, and goes after the worst offenders, it should not have any vagueness problem. Perhaps if it gets too aggressive, a court will push back, but until then there is no real problem with the approach.

The Eleventh Circuit in *LabMD v. FTC* was much more skeptical. The opinion ultimately expresses its skepticism about the FTC's authority in dicta, assuming without deciding that the FTC had the authority it claimed.[251] So the Commission survived its challenge to its authority intact. But then the Court vacated the Commission's cease and desist order, stating that the order "contains no [specific] prohibitions. . . . Rather, it commands LabMD to overhaul and replace its data-security program to meet an indeterminable standard of reasonableness."[252] Ultimately, it's hard to know what to make of *LabMD*, as it is a bizarre opinion in many ways,[253] but the apparent takeaway

[247] *Id.* at 255.

[248] *Id.* at 256.

[249] *Id.* at 256–59.

[250] *Id.* at 259.

[251] *See* Part II.A.1, *supra*.

[252] *LabMD, Inc. v. Fed. Trade Comm'n*, 894 F.3d 1221, 1236 (11th Cir. 2018).

[253] The Court reasons based on hypothetical a future prosecution by the FTC for violation of the cease-and-desist order, noting that it could be pursued either internally or in district court. *Id.* at 1232. Then the Court says without supporting authority that the same standards should apply whether the future action is internal to the FTC or is in district court. *Id.* Next, the court says that a district court would have to enforce it via its contempt power, but it's too vague an order so it would exceed the contempt power. *Id.* Finally, the court says once again that the standards are the same in the FTC, so ipso facto it must be unenforceable by the FTC. *Id.* at 1237. This reasoning is largely unsupported and frankly bizarre. Adding

is that enforcement orders must contain some specificity as to the deficiencies the Commission would like corrected.[254] In the end, *LabMD* says little about FTC's authority to address data security, or *mutatis mutandis* to address discriminatory AI.[255]

Finally, one absence stands out in these cases, and that is the lack of challenge as to the cost-benefit prong of Section 5. As a reminder, the Commission may only consider something unfair if the costs are not outweighed by "countervailing benefits to consumers."[256] But as described above, whose costs and whose benefits count is entirely unspecified. In a discrimination case, the costs will be borne by a minority of consumers, where the concomitant benefits could be borne by everyone at once—testing for discriminatory harm is costly, and that cost will be theoretically passed back to consumers in the form of higher prices. An interpretation of cost-benefit analysis that looks only at global aggregate costs could suggest that even intentional discrimination is acceptable where the costs imposed on the minority group are less in aggregate than the benefits.

As explained above, we do not believe that is the correct way to interpret the cost-benefit test. The data security challenges are yet another data point for that view. In a case like *Wyndham*, "hundreds of thousands of consumers" were affected, leading to "over $10.6 million dollars in damage." But Wyndham is a huge corporation with many locations. Yet the Court never even asks whether the cost of improving Wyndham's data security practice outweighs the harm it did. The Court does note that the "costs to consumers that would arise from investment in stronger cybersecurity" is a relevant consideration, but only in the context of what Wyndham should be aware of for purposes of fair notice.[257] There was never an attempt to evaluate

to the weirdness, the Court in the same opinion (1) accuses the FTC of not giving LabMD any direction, telling them only to overhaul their entire data security program, and (2) expresses the fear that this will result in the FTC "micromanaging" LabMD's security practice. *Id.*

[254] *Id.* at 1236 ("In sum, the prohibitions contained in cease and desist orders and injunctions must be specific. Otherwise, they may be unenforceable.").

[255] This is true, despite much of the commentary after the decision by political opponents of FTC authority. *See, e.g.*, TechFreedom, *LabMD Court Blocks FTC's Approach to Data Security* (June 6, 2018), https://techfreedom.org/labmdftcdatasecurity/ (" 'The court could hardly have been more clear: the FTC has been acting unlawfully for well over a decade,' said Berin Szóka, President of TechFreedom.")

[256] 15 U.S.C. § 45(n).

[257] Fed. Trade Comm'n v. Wyndham Worldwide Corp., 799 F.3d 236, 255 (3d Cir. 2015)

the numbers, and Wyndham never even raised it in its briefs.[258] Given the aggressiveness of these challenges, we believe that if this was seen as a viable issue, it would have been raised. But as it was not, we can be more confident that the views expressed by the majority in *Apple* is widely shared and that is not the right way to interpret the cost-benefit test.[259]

D. The Alternative: Magnuson-Moss Rulemaking

In this Part, we have focused on developing a common law approach through direct enforcement of Section 5. But as noted above, the FTC does possess the authority to pass "trade regulation rules" that affirmatively define practices that are unfair or deceptive.[260] This rulemaking authority is much more onerous than APA rulemaking, with many additional requirements.[261] The Commission has therefore not used this authority to pass any new rules since 1980.[262] In 2021, however, the FTC voted to revise its internal Rules of Practice to make it easier to pass new trade regulation rules,[263] and the Commission does seem poised to pursue rulemaking either instead of or in addition to direct enforcement.[264]

We focused on the data security parallel because it is interesting and less straightforward than Magnuson-Moss rulemaking. But it is not obviously the right path. Magnuson-Moss rulemaking has some advantages: the authority to do it is less controversial, the Commission can more easily garner public engagement in a rulemaking, and rules allow businesses to better know where they stand. But Magnuson-Moss rulemaking also takes an

---

[258] Opening Brief, Fed. Trade Comm'n v. Wyndham Worldwide Corp., 2014 WL 5106183 (C.A.3); Reply Brief, Fed. Trade Comm'n v. Wyndham Worldwide Corp., 2014 WL 7036128 (C.A.3).

[259] *See* notes 161–168, *supra* and accompanying text.

[260] 15 U.S.C. § 57a.

[261] Solove & Hartzog, *supra* note 9, at 620 (calling it "so procedurally burdensome that it is largely ineffective").

[262] Jeffrey S. Lubbers, *It's Time to Remove the "Mossified" Procedures for FTC Rulemaking*, 83 GEO. WASH. L. REV. 1979, 1997–98 (2015).

[263] Press Release, FTC Votes to Update Rulemaking Procedures, Sets Stage for Stronger Deterrence of Corporate Misconduct (July 1, 2021), https://www.ftc.gov/news-events/news/press-releases/2021/07/ftc-votes-update-rulemaking-procedures-sets-stage-stronger-deterrence-corporate-misconduct.

[264] *See* Lina M. Khan, FTC Chair, *Remarks as Prepared for Delivery*, IAPP Global Privacy Summit 2022, at 5–6 (Apr. 11, 2022), https://www.ftc.gov/news-events/news/speeches/remarks-chair-lina-m-khan-prepared-delivery-iapp-global- privacy-summit-2022.

extraordinarily long time,[265] and agency rules are less adaptable in an area where technology is rapidly evolving. So we cannot claim truly that either approach is ultimately better—likely they should be pursued in parallel. But regulating discriminatory AI is a matter of some urgency, and if Magnuson Moss rules take too long, there is a legitimate concern that the Commission will act too slowly to be the regulator it needs to be.

CONCLUSION

The FTC's apparent desire to regulate discriminatory AI is a welcome development because the scope of FTC enforcement is far less constrained than discrimination law. It has the capacity to reach a broader set of domains than employment, credit, housing, among the few others subject to regulation. It can likewise hold accountable a broader set of actors—not just the ultimate decision maker targeted by most discrimination laws, but also the many other actors that support the decision-making process, including vendors of AI products and services. It can consider a broader set of possible discrimination injuries, escaping the narrow confines of the traditional allocative concerns of discrimination law, including cases where AI products and services exhibit systematic performance disparities across different demographic groups. It can target not only the act of discrimination, but the surrounding business practices that make discrimination more likely to occur. And as a regulatory agency, it can also avoid many of the procedural and structural challenges that limit individual plaintiffs' ability to vindicate their rights under existing discrimination laws. Such intervention by the FTC will be possible under its existing authority, and there are existing models for how they can go about it.

Taken together, this all suggests that the FTC can play a unique role by both filling gaps in discrimination law and helping to enforce existing discrimination laws more effectively. In taking on this role, though, the FTC should make special efforts to coordinate with other relevant agencies that enforce discrimination laws, working, for example, with the EEOC, the Consumer Financial Protection Bureau (CFPB), and HUD in employment,

[265] One study examined the time from proposed rule to completion of different procedures, finding that under APA informal rulemaking, the average rule took nine to ten months, while under Magnuson-Moss, it took over five years. Lubbers, *supra* note 262, at 1997–98.

credit, and housing cases. Rather than competing with these agencies or potentially stepping on their toes, the FTC should work with these agencies to determine the unique value that it has to offer in coordinated enforcement actions—and focus its efforts there.[266] Moreover, other agencies with a history of discrimination enforcement could be valuable partners in the Commission's effort to define unfairness standards. Many other agencies have been grappling with the questions raised by AI.[267] The FTC should work with them and learn from their experience and domain expertise. At the same time, the FTC should not act merely in a supportive role. Its authority grants it the ability to tackle issues well beyond the scope of these other agencies, and it likely possesses relevant technical expertise beyond that of other agencies.[268] Not only can it fill gaps, but it can chart new regulatory terrain where there is obvious unfairness, but no immediately relevant discrimination law.

While FTC intervention can be helpful, the Commission will have limitations. While it has the authority to initiate investigations of what it believes to be an unfair practice, it lacks the authority to compel businesses to routinely produce and share information that might reveal when their AI products or services are discriminatory.[269] As a result, it is likely to only use its investigative powers when there is already evidence that there is something to discover. While much of the FTC's enforcement actions have followed this pattern (i.e., only initiating investigations after some practice had been brought to light by earlier journalistic or academic investigations),

---

[266] The FTC regularly coordinates with other agencies where enforcement authority overlaps. *See* Prepared Statement of the Federal Trade Commission on Opportunities and Challenges in Advancing Health Information Technology at 3 n.6, https://www.ftc.gov/legal-library/browse/prepared-statement-federal-trade-commission-opportunities-challenges-advancing-health-information (describing cases in which the FTC has worked with Office of Civil Rights of Health and Human Services to address data security in the medical context).

[267] *See* note 42, *supra* and accompanying text (discussing HUD's lawsuit against Facebook; Artificial Intelligence and Algorithmic Fairness Initiative (EEOC), https://www.eeoc.gov/ai; Press Release, *CFPB Acts to Protect the Public from Black-Box Credit Models Using Complex Algorithms* (May 26, 2022), https://www.consumerfinance.gov/about-us/newsroom/cfpb-acts-to-protect-the-public-from-black-box-credit-models-using-complex-algorithms/.

[268] For example, the FTC has a rotating Chief Technologist role and are actively recruiting technologists. *See* Technologist Hiring Program, https://www.ftc.gov/about-ftc/careers/work-ftc/technologists.

[269] The Algorithmic Accountability Act of 2022 was proposed to give it exactly this authority by requiring the FTC to pass rules requiring impact assessments, Algorithmic Accountability Act of 2022 S.3752, 117th Cong. § 3(b) (2022).

discrimination is notoriously difficult to uncover via individual observations because it is about systematic differences in the treatment of entire groups. Stray observations and ad hoc studies will often be insufficient to determine whether there is something worthy of more serious and systematic investigation. As a result, FTC intervention is likely to be most effective when there are additional laws and policies in place that make the discovery of unfairness easier, such as legally required impact assessments, audits, or other forms of evaluation. In addition, the Commission has long been resource-constrained,[270] and as the market in AI grows, enforcement will only become more resource-intensive.

But as with other efforts to address discriminatory AI, the FTC does not need to solve the problem entirely. The FTC already has what it needs to get started addressing the problem of discriminatory AI, and this is a good thing.

[270] *See generally* Testimony of Chair Lina M. Khan Before the House Appropriations Subcommittee on Financial Services and General Government (May 18, 2022) (requesting greater funding); Tony Romm, *Will Congress fund Internet privacy?* POLITICO (June 3, 2011) (discussing the FTC's need for more resources).