

UNINTENTIONAL ALGORITHMIC DISCRIMINATION: HOW ARTIFICIAL INTELLIGENCE UNDERMINES DISPARATE IMPACT JURISPRUDENCE

VINCENT CALDERON[†]

ABSTRACT

Artificial intelligence holds the capacity to revolutionize the economy by capturing efficiencies. These benefits, ostensibly, should pass down to consumers, thereby benefitting the general public. But the immense complexity of AI systems is bound to introduce legal hurdles for plaintiffs and frustrate our disparate impact jurisprudence. Specifically, demonstrating causation and proffering a less discriminatory alternative are herculean tasks for a plaintiff seeking to prove a disparate impact upon which legal relief may be granted. The courts have already begun to wrestle with these issues, primarily in the housing and employment sectors. With the rapid surge of AI systems, courts should expect further inquiry into how these programs interfere with our established antidiscrimination framework. This Note outlines how each step of a plaintiff's successful disparate impact analysis is hindered by the opaque ways in which AI operates. This Note then proposes several policy reforms to mitigate these consequences.

INTRODUCTION

The United States “has not always measured up to its constitutional heritage of equality for all.”¹ In the 21st century, artificial intelligence (“AI”) is now contributing to this unfortunate history of discrimination, albeit unintentionally. When such discrimination occurs, a plaintiff could rely on a disparate impact analysis, but this is rarely

[†] Duke University School of Law, J.D. expected May, 2024; B.S. in Industrial and Labor Relations, Cornell University, May, 2020. Thank you to Professor Lee Tiedrich for her guidance and support in writing this Note.

¹ *Goldberg Says U.S. Will Sign U.N. Pact For Racial Equality*, N.Y. TIMES, July 7, 1966.

successful² and has only been further frustrated by the advent of complex and unexplainable³ AI technologies.

Employees and housing applicants are bearing the weight of these new technologies, and current disparate impact jurisprudence is insufficient to litigate against corporations' employment and housing practices. A disparate impact claim already places significant causation standards onto the plaintiff, and AI, by its opaque nature, further complicates a plaintiff's case. Although AI is critical to the US' strategy for economic development, these new technologies cannot become new ways to discriminate.⁴

I. WHAT IS ARTIFICIAL INTELLIGENCE?

Artificial intelligence is one of the most important technologies in the world today. It has the potential to diagnose diseases,⁵ prevent securities fraud,⁶ and unlock a sustainable future.⁷

But, what is it, and what are the associated societal and legal risks inherent in such potential? Artificial intelligence has eluded any single definition, but Congress recently defined it as “a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments.”⁸

² See Michael Selmi, *Was the Disparate Impact Theory a Mistake?*, 53 UCLA L. REV. 701, 738–39 (2006) (showing that a plaintiff's success rate in disparate impact claims in Court of Appeals decisions between 1984-2001 is 19.2%).

³ *The AI Black Box Problem*, THINKAUTOMATION, <https://www.thinkautomation.com/bots-and-ai/the-ai-black-box-problem/>.

⁴ Press Release, Off. Pub. Affs., Justice Department and EEOC Warn Against Disability Discrimination: Employers' Use of Artificial Intelligence Tools Can Violate the Americans with Disabilities Act (May 12, 2022), <https://www.justice.gov/opa/pr/justice-department-and-eeoc-warn-against-disability-discrimination>.

⁵ See generally Yogesh Kumar et al., *Artificial Intelligence in Disease Diagnosis: A Systematic Literature Review, Synthesizing Framework and Future Research Agenda*, 14 J. AMBIENT INTEL. & HUMANIZED COMPUTING 8459 (2022).

⁶ DAVID FREEMAN ENGSTROM ET AL., GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 22 (2020), <https://www.acus.gov/sites/default/files/documents/Government%20by%20Algorithm.pdf>.

⁷ Gretchen O'Hara, *My Next Chapter: Shaping a Sustainable Future With AI*, LINKEDIN (Mar. 2, 2020), <https://www.linkedin.com/pulse/my-next-chapter-shaping-sustainable-future-ai-gretchen-o-hara>.

⁸ National Artificial Intelligence Initiative Act of 2020, 15 U.S.C. § 9401(3), (2021).

The underlying goal of artificial intelligence is to create computer models that exhibit “intelligent behaviors” like humans without having to be explicitly programmed to do so.⁹ Businesses employ AI in a wide variety of practices, but are often unable to explain why the algorithm makes particular decisions. This is the “black box problem,” which refers to a decision-making process that is opaque and not understood by humans in a causal way.¹⁰ This lack of understandability has broad implications for how this technology is regulated, both domestically and internationally.

A. Current State of Artificial Intelligence Regulation Domestically and Abroad

The Biden Administration recently promulgated an AI Bill of Rights, which outlines five protections that Americans should be afforded in this new era of innovation: 1) safe systems, 2) data privacy, 3) notice and explanation, 4) human alternatives, and, most significantly, 5) protection from algorithmic discrimination.¹¹ While this represents a momentous acknowledgement of the potential negative implications of AI, it remains a voluntary framework.

By contrast, the European Union recently published its own proposal for Artificial Intelligence Regulation.¹² And in 2022, the European AI Liability Directive would relax plaintiffs’ evidentiary burdens, with the acknowledgement that AI’s complexity makes it prohibitively difficult or even impossible to hold any party responsible.¹³ Some academics have noted that the EU’s innovative approach is a “wake-

⁹ Sara Brown, *Machine Learning, Explained*, MIT SLOAN SCH. MANAGEMENT (Apr. 21, 2021), <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>.

¹⁰ Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a "Right to an Explanation" is Probably Not the Remedy You are Looking For*, 16 DUKE L. & TECH. REV. 18, 59 (2017).

¹¹ THE WHITE HOUSE, BLUEPRINT FOR AN AI BILL OF RIGHTS, <https://www.whitehouse.gov/ostp/ai-bill-of-rights/> (last visited Dec. 5, 2022).

¹² Jane Finlayson-Brown et al., *Key Provisions of the Draft AI Regulation*, ALLEN & OVERY (May 24, 2021), <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/key-provisions-of-the-draft-ai-regulation>.

¹³ *Commission Proposal for a Directive of the European Parliament and of the Council on Adapting Non-Contractual Civil Liability Rules to Artificial Intelligence*, at 16, COM (2022) 496 final (Sept. 28, 2022) [hereinafter *Commission Proposal*]

up call” for the US, which has been slower to regulate this new technology.¹⁴

The US has, however, actively sought collaboration with allies to establish principles and build trustworthy AI systems centered around human rights. In 2019, the US endorsed the principles put forth by the Organisation for Economic Co-Operation and Development (“OECD”).¹⁵ Then in 2020, the US joined the Global Partnership on Artificial Intelligence (“GPAI”) to foster innovation grounded in human rights, inclusion and diversity.¹⁶

B. Benefits of Artificial Intelligence: Employment and Housing

In the employment sector, companies have implemented AI both to micro-target job postings based on consumers’ browser histories and to screen candidates more efficiently through resume-reading algorithms.¹⁷ Additionally, chatbots and video interviewing have simplified the traditionally time-consuming interviewing process for both employers and employees.

In the housing sector, major companies such as Zillow and LoanSnap have deployed AI to match consumers with mortgages more efficiently than traditional real estate agents could.¹⁸ Before the advent of AI, obtaining the pertinent data and information was a burdensome process that involved parsing through masses of documentation at local offices.¹⁹ AI, however, optimizes this process by analyzing millions of documents in seconds.²⁰

¹⁴ Meredith Broadbent & Sean Arrieta-Kenna, *AI Regulation: Europe’s Latest Proposal is a Wake-Up Call for the United States*, CTR. FOR STRATEGIC & INT’L STUD. (May 18, 2021), <https://www.csis.org/analysis/ai-regulation-europes-latest-proposal-wake-call-united-states>.

¹⁵ Fiona Alexander, *U.S. Joins with OECD in Adopting Global AI Principles*, NAT’L TELECOMMS. & INFO. ADMIN. (May 22, 2019), <https://www.ntia.doc.gov/blog/2019/us-joins-oecd-adopting-global-ai-principles>.

¹⁶ THE GLOBAL PARTNERSHIP OF ARTIFICIAL INTELLIGENCE, <https://gpai.ai/about/> (last visited Jan. 5, 2022).

¹⁷ Russell Parsons, *Facebook to Serve Ads Based on Web Browsing History*, MARKETINGWEEK (June 12, 2014), <https://www.marketingweek.com/facebook-to-serve-ads-based-on-web-browsing-history/>.

¹⁸ Diana Olick, *Artificial Intelligence is Taking Over Real Estate – Here’s What That Means for Homebuyers*, CNBC (Sept. 17, 2021), <https://www.cnbc.com/2021/09/17/what-artificial-intelligence-means-for-homebuyers-real-estate-market.html>.

¹⁹ *Id.*

²⁰ *Id.*

Additionally, leasing companies employ algorithms that efficiently screen tenants or even outsource this task to vendors. CoreLogic Rental Property Solutions, LLC (“CoreLogic”) is one such company, which specializes in providing leasing companies and landlords copies of applicants’ criminal records.²¹ Employing vendor technology is ostensibly cheaper than conducting these activities in-house, especially for individual landlords. And because tenant-screening costs are often passed onto the applicant,²² the applicant should benefit from a lower cost. However, the discriminatory potential of AI drastically alters this calculation.

C. Drawbacks of Artificial Intelligence: Discrimination and Bias

Although there is demonstrated potential for AI to unlock unprecedented economic growth, the threat of discrimination is ever-present. For example, in 2014 Amazon began to build AI programs to automate its recruitment.²³ Just one year later, however, the company realized that the AI was ranking female applicants’ resumes lower than for male applicants, because the models were trained on resumes predominantly submitted by men.²⁴

Facial recognition technologies are also likely to result in discriminatory employment decisions. For instance, women with darker skin tones are 32 times more likely to be misclassified than are men with lighter skin tones.²⁵ Additionally, Microsoft AI and Face++ assigned more negative emotions to Black NBA basketball players than they did for white NBA players.²⁶ Because facial recognition AI is not race-neutral, this technology might become “a license to discriminate.”²⁷

²¹ Conn. Fair Hous. Ctr. v. CoreLogic Rental Prop. Sols., LLC, 369 F.Supp.3d 362, 367 (D. Conn. 2019).

²² Eric Dunn & Marina Grabchuk, *Background Checks and Social Effects: Contemporary Residential Tenant-Screening Problems in Washington State*, 9 SEATTLE J. SOC. JUST. 319, 323 (2010).

²³ Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, REUTERS (Oct. 10, 2018), <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.

²⁴ *Id.*

²⁵ Joy Buolamwini, *Gender Shades*, MIT MEDIA LAB (Aug. 10, 2017), <https://www.media.mit.edu/publications/full-gender-shades-thesis-17/>.

²⁶ Lauren Rhue, *Racial Influence on Automated Perceptions of Emotions*, SSRN 6 (2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3281765.

²⁷ Drew Harwell, *A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job*, WASH. POST (Nov. 6, 2019),

Predictably, plaintiffs have begun to litigate algorithmic discrimination issues in the courts. For instance, Meta had been employing its “Special Ad Audience” tool which, using characteristics protected by the Fair Housing Act (“FHA”), determined eligibility to receive advertisements for its users.²⁸ Because this tool revealed advertisements only to certain users, thereby excluding certain groups from applying for housing, Meta and the Department of Justice reached a historic settlement in which Meta agreed to stop using the algorithm.²⁹

Additionally, in a 2018 lawsuit against CoreLogic, plaintiff Carmen Arroyo contended that CoreLogic’s tenant-screening algorithm violated the FHA.³⁰ CoreLogic offers a tenant-screening product called “CrimSAFE,” which determines whether an applicant should be accepted or rejected after an algorithm analyzes an applicant’s criminal record.³¹

Ms. Arroyo tried to move her disabled son, Mikhail Arroyo, into her apartment but was unable to do so after CrimSAFE stated that Mikhail had a disqualifying criminal record.³² Mikhail’s only criminal record, however, was a withdrawn charge in 2014.³³ The Arroyos, a Hispanic family, alleged that because African-Americans and Latinos are arrested, convicted, and incarcerated at rates disproportionate to their share of the general population, CrimSAFE disproportionately denies housing opportunities to applicants of color.³⁴ *CoreLogic* held that Ms. Arroyo has standing to bring a disparate impact claim under the FHA and demonstrates how plaintiffs may bring claims for algorithmic discrimination. Importantly, the plaintiffs also brought a claim of

washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/.

²⁸ United States Attorney’s Office, Southern District of New York, *United States Attorney Resolves Groundbreaking Suit Against Meta Platforms, Inc., Formerly Known As Facebook, To Address Discriminatory Advertising For Housing*, DEP’T OF JUST. (June 21, 2022), <https://www.justice.gov/usao-sdny/pr/united-states-attorney-resolves-groundbreaking-suit-against-meta-platforms-inc-formerly>.

²⁹ Office of Public Affairs, *Justice Department Secures Groundbreaking Settlement Agreement with Meta Platforms, Formerly Known as Facebook, to Resolve Allegations of Discriminatory Advertising*, DEP’T OF JUST. (June 21, 2022), <https://www.justice.gov/opa/pr/justice-department-secures-groundbreaking-settlement-agreement-meta-platforms-formerly-known>.

³⁰ Complaint at 22, *Conn. Fair Hous. Ctr. v. CoreLogic Rental Prop. Sols., LLC*, 369 F.Supp.3d 362 (D. Conn. 2019) (No. 3:18-CV-705).

³¹ *Id.* at 2.

³² *Id.* at 3.

³³ *Id.* at 4.

³⁴ *Id.*

disability discrimination under the FHA, potentially broadening standing for disparate impact claims based on characteristics other than race.³⁵

In the employment sector, too, algorithmic discrimination has the potential to harm applicants with disabilities. The Equal Employment Opportunity Commission (“EEOC”) has even outlined three ways by which algorithmic discrimination can arise: 1) an employer does not provide a reasonable accommodation for the job applicant to be rated fairly and accurately by the algorithm, 2) when an algorithm that “screens out” an individual with a disability and 3) when the algorithm violates restrictions on disability-related inquiries and medical examinations.³⁶ Although the plaintiffs in *CoreLogic* brought a claim of disability discrimination under the FHA rather than under the Americans with Disabilities Act (“ADA”), the case represents how disability discrimination might result in damaging consequences, regardless of statutory offense.

Algorithmic discrimination is guaranteed to become more pervasive as AI gains more prominence. Thus, the courts should bolster its disparate impact jurisprudence so that it remains an effective tool to combat unintentional discrimination. As it stands, the doctrine is unprepared to analyze AI’s complexity, leaving plaintiffs with little recourse.

II. INTRODUCTION TO RELEVANT SUBSTANTIVE LAW

A. Common Law: Disparate Impact Analysis

Although employment and housing discrimination are endemic throughout U.S. history, it was not until the 1970s that the Supreme Court developed a coherent discrimination jurisprudence. Disparate impact refers to practices in housing and employment in which a facially neutral policy adversely affects one group of people of a protected class more than another.³⁷ Disparate impact was first articulated by the Supreme Court in *Griggs v. Duke Power Co.*³⁸ In *Griggs*, the Court held that because the employment practice in question did not reasonably relate to the

³⁵ *Id.* at 6.

³⁶ *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*, U.S. EQUAL EMP. OPPORTUNITY COMM’N (May 12, 2022), <https://www.eeoc.gov/laws/guidance/americans-disabilities-act-and-use-software-algorithms-and-artificial-intelligence>.

³⁷ *What are disparate impact and disparate treatment?*, SOCIETY FOR HUMAN RESOURCE MANAGEMENT, <https://www.shrm.org/resourcesandtools/tools-and-samples/hr-qa/pages/disperateimpactdisparatetreatment.aspx>.

³⁸ *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).

applicant's ability to perform the job but resulted in fewer Black employees being hired, Duke Power had discriminated against a protected class in violation of federal law.

It was incontrovertible after *Griggs* that disparate impact claims could be brought in employment discrimination cases. But, the courts still faced the question whether such claims were cognizable under the FHA. In 2013, the Department of Housing and Urban Development ("HUD") formalized the three-step disparate impact analysis.³⁹ The Supreme Court then doubled down and formally ruled that disparate impact claims are cognizable under the FHA in *Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc.*⁴⁰

To prevail on a disparate impact claim, a plaintiff first must establish an adverse disparate impact.⁴¹ Importantly, the plaintiff must demonstrate causation between the practice and adverse outcome. *Inclusive Communities* recently heightened this mandate by stressing the importance of a "robust causality requirement"⁴² and the Department of Justice ("DOJ") has noted that causation is frequently shown with statistics.⁴³ Because AI is inherently complex,⁴⁴ demonstrating causation is already a high threshold requirement in bringing a successful case. Significantly, too, *Inclusive Communities* mandated that a causal connection must be made at the pleading stage,⁴⁵ further complicating a plaintiff's evidentiary burden.

If the plaintiff does meet this initially steep burden, the defendant must prove the challenged practice is necessary to achieve a substantial and legitimate interest.⁴⁶

Lastly, if a substantial, legitimate, justification is provided for the discriminatory practice, the burden shifts back to the plaintiff to identify a

³⁹ 24 C.F.R. § 100.500(b)(4) (2013).

⁴⁰ *Tex. Dep't of Hous. & Cmty. Affairs v. Inclusive Cmty. Project, Inc.*, 576 U.S. 519 (2015).

⁴¹ *Proving Discrimination*, DEPARTMENT OF JUSTICE, <https://www.justice.gov/crt/fcs/T6Manual7#D>.

⁴² 576 U.S. at 521.

⁴³ *Proving Discrimination*, *supra* note 41.

⁴⁴ *See* AI Liability Directive, *supra* note 13, at 16 ("[T]he specific characteristics of certain AI systems, such as opacity, autonomous behavior and complexity, may make it excessively difficult, if not impossible, for the injured person to meet this burden of proof.").

⁴⁵ *See Inclusive Cmty.*, 576 U.S. at 543 ("A plaintiff who fails to allege facts at the pleading stage ... cannot make out a prima facie case of disparate impact.").

⁴⁶ *Id.* at 527.

“less discriminatory alternative.”⁴⁷ Importantly, this alternative must be feasible and meet the defendant’s proffered legitimate objective.⁴⁸ There are two ways for a plaintiff to fail on this step. First, the identified alternative is not sufficiently specific. Second, the alternative does not meet all the defendant’s goals that prompted the use of the practice in the first place. Because the deployment of AI is often grounded in unprecedented efficiency, this second failure will likely be implicated often.

B. Statutory Law

1. State and Local Artificial Intelligence Statutes

Unlike Congress, states and localities have enacted legislation against algorithmic bias. In what seems to be a direct response to the negative implications of facial recognition technology, Illinois became the first state to enact legislation protecting consumers from bias. Illinois enacted the Artificial Intelligence Video Act in 2020, which applies to all employers that use AI to analyze applicants’ video interviews of employment positions in Illinois.⁴⁹ Consistent with the OECD Principles of inclusive growth and transparency,⁵⁰ Illinois explicitly requires employers’ disclosure that they are employing AI in their video interviewing and deletion of videos if requested to do so. The Illinois statute does not directly address algorithmic bias, but still represents a monumental first step in regulating artificial intelligence.

New York City recently followed with Local Law Int. No. 1894-A. Unlike Illinois’ legislation, the NYC statute regulates algorithmic bias directly, requiring audits on employment decision tools.⁵¹ The statute contains a notice requirement and requires that an employer may use the tool only if it was audited within one year of its deployment.⁵² Due to New York City’s financial and employment prominence, the statute could be a bellwether for similar legislation in other large cities.

2. Federal Civil Rights Statutes

⁴⁷ *Id.*

⁴⁸ *Proving Discrimination*, *supra* note 41.

⁴⁹ 820 Ill. Comp. Stat. Ann. 42/5

⁵⁰ See ORG. FOR ECON. COOP. AND DEV., *Recommendation of the Council on Artificial Intelligence*, (May 21, 2019).

⁵¹ Anthony A. Mingione & Mara B. Levin, *NYC Employers Using AI for Screening Beware*, BLANK ROME (Sept. 30, 2022), <https://www.blankrome.com/publications/nyc-employers-using-ai-screening-beware>.

⁵² *Id.*

Although the landmark federal civil rights statutes do not directly cover algorithmic bias, they are still likely to be implicated in the courts. The Civil Rights Act of 1964 is a landmark piece of legislation that bans employment discrimination based on race, color, religion, or sex, in various facets of public life.⁵³

The Civil Rights Act of 1968 contains Title VII, known as the FHA. Title VII permits claims of both disparate impact and disparate treatment⁵⁴ for several protected classes, including race, religion, national origin, sex, and people with disabilities.⁵⁵ The FHA bans several forms of housing discrimination, including advertisements that indicate a preference based on the Act's protected classes.

Lastly, the ADA prohibits employers from discriminating on the basis of disability. The Supreme Court has overtly held that disparate impact claims are cognizable under the ADA.⁵⁶ And although disparate impact has not been overtly endorsed by the EEOC, its recent guidance invites plaintiffs to employ such an analysis in a claim.⁵⁷ Importantly, an employer can violate the ADA if the artificial intelligence technology utilized was secured from an outside vendor.⁵⁸

Although our body of federal anti-discrimination law is the result of an unfortunate history, it has been helpful in allowing plaintiffs to seek redress over the decades. However, its already-limited efficacy is undermined by current artificial intelligence practices.

III. THE DIFFICULTY IN PROVING DISPARATE IMPACT OF ARTIFICIAL INTELLIGENCE PRACTICES

Even before the advent of AI, plaintiffs in disparate impact cases only had on average a 19.2% success rate between 1984 and 2001 in Court of Appeals decisions.⁵⁹ This section will detail each burden of proof, showing how AI further frustrates each step of a plaintiff's case.

⁵³ Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241 (codified as amended at 42 U.S.C.).

⁵⁴ See 576 U.S. at 545 (“The Court holds that disparate-impact claims are cognizable under the Fair Housing Act...”).

⁵⁵ Fair Housing Act of 1968, Pub. L. No. 90-284, §§ 801-819, 82 Stat. 73, 8189 (codified as amended at 42 U.S.C.).

⁵⁶ See *Raytheon Co. v. Hernandez*, 540 U.S. 44, 53 (2003) (“Both disparate-treatment and disparate-impact claims are cognizable under the ADA.”).

⁵⁷ *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*, *supra* note 36.

⁵⁸ *Id.*

⁵⁹ Selmi, *supra* note 2, at 738.

A. Establishing an Adverse Impact

The first step in disparate impact analysis lies with the plaintiff to establish an adverse impact, which itself has four elements: 1) identifying the facially neutral policy, 2) establishing adversity or harm, 3) establishing a disparity, and 4) demonstrating causation.⁶⁰

Identifying the facially neutral policy is typically an easy task for a plaintiff, as they likely know the policy against which they are complaining, but this threshold requirement is complicated by AI. For example, if a hypothetical employer required a written examination for promotion in which candidates of a certain race were excluded much more often than white candidates, one can obviously identify the facially neutral practice as the written examination. But, when a job candidate with disabilities is “screened out” after a video interview,⁶¹ the information asymmetry between applicant and employee renders it difficult to identify exactly which practice resulted in the adverse decision. Some possibilities include the AI’s interpretation of their speech cadence, facial movements, and even recognition of necessary medical equipment.

The DOJ has also stated that to identify the facially neutral policy, a plaintiff must “accurately and completely define” the practice in question.⁶² AI has yet to enjoy a single definition, and many companies and computer scientists cannot even “completely define” this novel technology. A plaintiff with presumably very little technical knowledge in this field is then left with a very high threshold standard to make out a *prima facie* case.

Second, a plaintiff must establish whether this practice constitutes a harm unto a protected class to become actionable. Because many courts impliedly assume that the alleged impact was adverse,⁶³ this element will not be discussed in detail.

Third, a plaintiff must establish a disparity, meaning that a disproportionate share of the adversity is borne by a protected class. In so doing, a plaintiff will determine the need for statistical evidence to support their case.⁶⁴ This is problematic for several reasons.

From the outset, a plaintiff who employs statistics to support their claim would be confronted with incomprehensible data and the black box

⁶⁰ *Proving Discrimination*, *supra* note 41.

⁶¹ *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*, *supra* note 36.

⁶² *Proving Discrimination*, *supra* note 41.

⁶³ *See id.* (“Rather, courts frequently assume that the impacts alleged were adverse.”).

⁶⁴ *See id.* (explaining that statistical evidence is often necessary).

problem. Courts, too, might find interpreting such tortuous data challenging. Although the EEOC has stated that employers can be held liable for vendor technologies,⁶⁵ the recent demand of *Inclusive Communities* is that statistical evidence must be presented at the pleading stage of a disparate impact claim.⁶⁶ Because pleading happens before discovery, a plaintiff might not be able to glean crucial information about the vendor's technology to include in their complaint, rendering the EEOC guidance moot. So, not only does a plaintiff face difficulty in understanding the data, but the data might not even be available.

The necessity of statistics is especially troublesome in cases of disability discrimination. The EEOC has noted that each disability is unique,⁶⁷ with the natural corollary being that not all disability discrimination cases are alike. Unlike cases of racial or sex discrimination, where there is likely more available data for employment or rental decisions based on that single demographic, disability discrimination can involve any disability from deafness, blindness, or autism. Therefore, it is likely that there is less data available for a disabled plaintiff, rendering it extremely difficult for such a complainant to meet this standard. The Fourth Circuit, however, has stated that statistics are not needed if there is an inference of this adverse decision being a common experience among similarly situated individuals,⁶⁸ which might be helpful for plaintiffs. The DOJ has even recognized this possibility in stating that statistics are unnecessary when the disparate effect of the defendant's policy is so "obvious or predictable."⁶⁹ While this is a helpful reprieve for defendants whom statistics are unavailable, AI is hardly "obvious or predictable." In any other instance, this safety valve might be mitigative.

The fourth and final criterion is to demonstrate causation between the practice and the adverse effect. The opacity of AI, accompanied by the heightened standard from *Inclusive Communities*, now makes causation the most difficult part of making out a prima facie case. Before *Inclusive Communities*, causation was shown on a fairly liberal basis by showing overrepresentation in adverse circumstances⁷⁰ or underrepresentation in

⁶⁵ *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*, *supra* note 36.

⁶⁶ *Tex. Dep't of Hous. & Cmty. Affairs v. Inclusive Cmty. Project, Inc*, 576 U.S. 519, 543 (2015).

⁶⁷ *The Americans with Disabilities Act and the Use of Software, Algorithms, and Artificial Intelligence to Assess Job Applicants and Employees*, *supra* note 36.

⁶⁸ *Mitchell v. Bd. of Trs.*, 599 F.2d 582, 585–86 (4th Cir. 1979).

⁶⁹ *Proving Discrimination*, *supra* note 41.

⁷⁰ *Huntington Branch, NAACP v. Town of Huntington*, 844 F.2d 926 (2d. Cir. 1988).

public benefits.⁷¹ And although *Inclusive Communities* provided relief to plaintiffs by affirming that disparate impact theory is cognizable under the FHA,⁷² it simultaneously made it more difficult to prove causation, for two reasons.

First, while the owner of the proprietary technology might be ultimately compelled to disclose how the technology operates in the discovery or over the course the litigation, *Inclusive Communities* mandated that plaintiffs demonstrate causation at the pleading stage.⁷³ So even if a complainant had the technical capacity to determine the causal effect between the practice and adverse decision, she might never have the chance to do so.

Second, Justice Kennedy stressed the importance of a “robust causality requirement.”⁷⁴ While the Court’s new heightened standard is rooted in an understandable policy concern for frivolous disparate impact claims against housing providers,⁷⁵ it is unduly narrow. The Supreme Court doubled down on this new standard in 2017 by going as far to say that foreseeability alone is not sufficient to establish proximate cause.⁷⁶ Instead, there must now be must “direct relation” between the injury and alleged conduct.⁷⁷ This is a sharp deviation from the common notion that “foreseeability is the touchstone of proximate cause.”⁷⁸ When contrasted with the EU’s approach of reducing evidentiary hurdles for victims harmed by AI,⁷⁹ the US approach is sorely outmoded. Recent Supreme Court demands pay no mind to the advent of complex technologies and their unique challenges to plaintiffs.

The “black-box problem” is one such challenge. A “black-box” is a system that allows one to see the input and output but provides no information on the processes or workings between.⁸⁰ Naturally, this

⁷¹ Comm. Concerning Cmty. Improvement v. City of Modesto, 583 F.3d 690 (9th Cir. 2011).

⁷² See generally *Tex. Dep’t of Hous. & Cmty. Affairs v. Inclusive Cmty. Project, Inc.*, 576 U.S. 519 (2015).

⁷³ *Id.* at 543.

⁷⁴ *Id.* at 542.

⁷⁵ *Id.* at 543–44 (“These limitations are also necessary to protect defendants against abusive disparate-impact claims.”).

⁷⁶ See *Bank of Am. Corp. v. City of Miami*, 581 U.S. 189, 201 (2017) (“We conclude that foreseeability alone is not sufficient to establish proximate cause under the FHA, and therefore vacate the judgment below.”).

⁷⁷ *Id.* at 202.

⁷⁸ FORESEEABILITY AS AN ELEMENT OF NEGL. AND PROXIMATE CAUSE (Am. L. Rep.).

⁷⁹ See generally *Commission Proposal*, *supra* note 13.

⁸⁰ *The AI Black Box Problem*, *supra* note 3.

implies that the system's creator cannot establish a causal link between the input and the output. If experts have trouble establishing this causal link, how can a plaintiff with presumably minimal technical knowledge be expected to do so? The plaintiff has the nearly impossible burden of proving causation at the pleading stage, at which point she likely has little insight into the system's workings. Further complicating the analysis is Justice Kennedy's assertion that "a statistical disparity must fail if the plaintiff cannot point to defendant's policy or policies causing that disparity."⁸¹ The Supreme Court's absolute language compounds the black-box problem by requiring that a plaintiff identify specifically how the practice caused their adverse decision, disregarding the unique issues that AI poses to plaintiffs.

In summary, causation is now significantly harder to prove after the Court mandated that a plaintiff demonstrates causation by showing a direct relation between the injury and practice, notably at the pleading stage. This difficulty is further augmented by the black-box problem, suggesting that the Supreme Court is not aware of recent technological developments.

B. Substantial Legitimate Justification

If the plaintiff does make out a prima facie case, the defendant then has the burden to show that the practice in question was necessary to achieve a legitimate and substantial goal related to the employment or housing goal. The obvious justification offered by employers and housing providers will be monetary benefits and enhanced efficiencies. Although several lower courts have taken the position that monetary justifications may fail for lack of evidence,⁸² defendants will almost always succeed since the unprecedented efficiencies and monetary benefits brought about by AI practices are well-documented.⁸³ Therefore, I predict that the identified justifications will almost always be persuasive to the courts.

C. Less Discriminatory Alternative

If the defendant provides a substantial, legitimate justification for its AI practice, then the burden will shift back to the complainant to

⁸¹ Tex. Dep't of Hous. & Cmty. Affairs v. Inclusive Cmty. Project, Inc, 576 U.S. 519, 542 (2015).

⁸² *Proving Discrimination*, *supra* note 41.

⁸³ Kelly Main, *How a Brilliant Artificial Intelligence-Inspired Solution Can Cut Your Company's Costs by 20 Percent*, INC. (Jun. 15, 2022), <https://www.inc.com/kelly-main/artificial-intelligence-efficiency-productivity.html>.

identify a “less discriminatory alternative[.]”⁸⁴ There are two common ways plaintiffs fail this evidentiary burden: 1) lack of specificity and 2) inability to meet all the defendant’s needs.⁸⁵ Similar to the first burden of persuasion, the complex features of AI will make it likely that a plaintiff falls into one of these buckets.

To overcome a lack of specificity failure, a plaintiff must provide a sufficiently specific alternative to the current practice.⁸⁶ This is difficult to achieve for two reasons.

First, a plaintiff might seek to modify the AI system to reduce its discriminatory effects. To have made it to this third evidentiary burden, the plaintiff will have transcended the pleading stage by making out a prima facie case and will be in the discovery stage. Although the plaintiff presumably now has access to information about the technology’s features, the black-box problem still imposes complications. Even more, a plaintiff likely has no understanding of how AI functions and will be unable to “specific[ally]”⁸⁷ propose modifications to the system. Of course, they can make blanket recommendations to the defendant, but the complexity of AI makes offering the requisite specific alternative nearly impossible.

Second, a plaintiff might propose that employers and housing providers utilize a new algorithm. If a complainant seeks to modify an existing algorithm, they at least have some foundation on which to rest recommendations to make it less discriminatory. By contrast, a plaintiff proposing a brand-new algorithm has no basis or requisite understanding of the technology to build one from scratch. Outside of the realm of AI, this might be quite simple to achieve. Take the following hypothetical as an example. Company X relies on educational credentials to promote its employees from within. However, female employees at this company, on average, possess weaker educational credentials than do male employees. As a result, they are adversely impacted by this facially neutral policy. If an employee brings a suit, she can propose that promotions should instead be based on the number of widgets produced each year. This less discriminatory alternative is sufficiently specific and allows the company to continue its goal of promoting its employees from within. Promulgating a new algorithm with “specific[ity],”⁸⁸ by contrast, is a herculean task for a plaintiff alleging disparate impact.

⁸⁴ *Inclusive Cmty.*, 576 U.S. at 528.

⁸⁵ *Proving Discrimination*, *supra* note 41.

⁸⁶ *Id.*

⁸⁷ *See id.*

⁸⁸ *See id.*

The other way a plaintiff often fails on this element is if they offer a less discriminatory alternative that does not satisfy all the defendant's needs. DOJ guidance notes that a plaintiff need not propose exact substitutes but can instead offer policies of a different manner.⁸⁹ A plaintiff might take this guidance to mean that they can propose an alternative that does not utilize AI at all. In most other cases, this would likely not be an issue, as different practices are largely interchangeable, such as my hypothetical example in the preceding paragraph. But AI is unique in that it can cut costs and increase efficiencies at an unprecedented rate. Because of this, a court will likely find that any less discriminatory alternative offered by a complainant, that does not involve artificial intelligence, will likely not meet all the defendant's monetary goals.

To summarize, a plaintiff has three choices under this last burden: 1) try to modify an existing algorithm, 2) propose an entirely new algorithm, or 3) offer a specific practice that does not employ AI but still matches the economic benefits of the system. The specificity requirement renders the first two choices impracticable, and the unparalleled economic benefits garnered from employing AI makes the third option similarly difficult to prevail.

IV. POLICY REFORMS

In response to the challenges outlined in the foregoing discussion, I propose three reforms. First, I propose a new disparate impact theory that is similar to the EU model, placing more burden onto the defendant and easing the causality requirement put forth in *Inclusive Communities*. Second, to bolster our case law on algorithmic discrimination, we should enhance plaintiffs' remedies and ensure that administrative agencies are vigilant in litigating claims. Third, Congress should implement uniform artificial intelligence legislation.

A. A New Disparate Impact Theory

As it currently stands under *Griggs*, disparate impact analysis places two of the three burdens of persuasion onto the plaintiff. Notably, the plaintiff bears the first burden in proving that there was an adverse disparate impact. However, there are significant causation, statistical, and identification issues that a plaintiff faces in making out a prima facie case.

Due to these problems, courts should place this initial burden onto the defendant. Specifically, defendants would bear the burden on proving that their AI practice *did not* result in an adverse decision. Although this may seem like a radical proposition, consider the following two points.

⁸⁹ *Id.*

First, this would only apply in cases of algorithmic discrimination, where the defendant has significantly more insight into the causation and statistical underpinnings of the model. Second, this approach has already gained significant traction in the EU.

The EU's regulatory approach has been innovative in protecting consumers from AI, with some commentators even arguing that it is "[s]etting the [r]ules of the [f]uture."⁹⁰ The US should then look to the recent EU AI Liability Directive for a model in where to place burdens of proof. Expressly acknowledging that the complexity of AI "may make it excessively difficult, if not impossible, for the injured person to meet this burden of proof," the EU proposed significantly easing its "causal link" requirement.⁹¹ In fact, the Directive would implement a "rebuttable presumption of causality."⁹² This less stringent causation standard highlights a recognition that AI has drastically altered liability analysis and that policy reform is needed. The US, by contrast, has gone in the opposite direction and instead made it more difficult to demonstrate causation.⁹³

Critics will assert that this new framework might result in unnecessary and expensive litigation. I push against this in three ways. First, the information asymmetry between plaintiff and defendant suggests that the defendant could more cheaply produce all the relevant data necessary to rebut the causal link.

Second, the EU would relax the standards for obtaining relevant data by allowing disclosure of relevant evidence upon which a plaintiff can make a claim.⁹⁴ Doing this from the outset of litigation, the Directive asserts, "should lead to a reduction of unnecessary litigation and avoid costs for the possible litigants caused by claims which are unjustified or likely to be unsuccessful."⁹⁵ Logically, then, the Supreme Court should rethink its recent command that a plaintiff fails to make out a prima facie case if she does not demonstrate causation at the pleading stage. Permitting a plaintiff to reach discovery before the causation analysis would result in a more informed decision on the part of the plaintiff on whether to continue with the suit, thereby reducing unnecessary litigation.

Third, this approach aligns more closely with the GPAI and principles put forth by the OECD. Collectively, these agreements include principles such as inclusive growth, transparency, and accountability.

⁹⁰ Broadbent & Arrieta-Kenna, *supra* note 14.

⁹¹ *Commission Proposal*, *supra* note 13, at 14, 16.

⁹² *Id.* at 13, 26.

⁹³ *See* 576 U.S. at 542 (stressing the importance of a "robust causality requirement.").

⁹⁴ *13 Commission Proposal*, *supra* note 13, at 12, 25.

⁹⁵ *Id.* at 18.

Imposing this rebuttable presumption will incentivize companies to put these ideals at the forefront of their AI engineering. Although there is no legal command to do so, the US should actually follow through on these agreements by changing this incentive structure for such companies.

Justice Kennedy likely did not have algorithmic discrimination top of mind when he imposed this stricter causation requirement. For future policy, judges should incorporate the EU's thinking on the interaction between AI and liability into its disparate impact jurisprudence.

B. Bolstering Case Law and Precedent

There is currently very little case law and accompanying precedent for algorithmic discrimination. Therefore, plaintiffs face uncertainty for plaintiffs in bringing a claim under disparate impact theory, which can be mitigated two ways: 1) improving the remedies for discrimination claims and 2) ensuring that federal agencies are vigilant in litigating cases by allocating sufficient resources to such organizations.

If a complainant alleges algorithmic discrimination, there are two threshold barriers to bringing suit. First, the success rate of disparate impact cases is already quite low.⁹⁶ Second, the high rate of failure is likely to be compounded by AI, for the reasons outlined in the foregoing discussion. From the outset, there is already little incentive to filing a claim.

Next consider the modest remedies from winning on a claim. In the employment setting, the limit on punitive and compensatory damages on winning against an employer with more than 500 employees is \$300,000.⁹⁷ This figure, however, includes punitive damages, which are only awarded when an employer commits an “especially malicious or reckless act of discrimination.”⁹⁸ But, I predict that algorithmic discrimination claims will primarily rest on a disparate impact analysis, which are definitionally unintentional. So, it is unlikely punitive damages would ever be awarded. As a result, that figure is significantly lower in reality. In the housing sector, the maximum civil penalty is a mere \$16,000 for a first-time offense.⁹⁹

⁹⁶ Selmi, *supra* note 2, at 738.

⁹⁷ U.S. EQUAL EMP. OPPORTUNITY COMM'N, *Remedies for Employment Discrimination*, <https://www.eeoc.gov/remedies-employment-discrimination>. (Last visited Mar. 6, 2024)

⁹⁸ *Id.*

⁹⁹ U.S. DEP'T OF HOUS. AND URB. DEV., FAIR HOUSING – EQUAL OPPORTUNITY FOR ALL 13 (2011), https://www.hud.gov/sites/documents/FHEO_BOOKLET_ENG.PDF.

A plaintiff thus faces a very low success rate and little remedial benefit. Therefore, the implicitly calculated expected value of bringing a claim is quite low. To alleviate this, we must increase the damages available for plaintiffs, thereby raising the incentive to bring suit.

This proposal is especially significant for plaintiffs who were discriminated against on the basis of disability. As discussed in Section IV, each disability is unique,¹⁰⁰ so disability discrimination can take many forms. If a Black or female plaintiff wins on a discrimination claim, the precedent is established for that respective demographic. But if a blind plaintiff wins on a claim of disparate impact, a deaf or neurodivergent complainant might not have directly analogous precedent on which to rely. Therefore, it is imperative that the case law is bolstered for all forms of disability discrimination claims.

Additionally, administrative agencies, namely the EEOC and HUD, charged with enforcing federal civil rights statutes must be active in litigating wrongdoings. However, the EEOC and HUD are sorely understaffed,¹⁰¹ lacking the capacity to bring suits on behalf of plaintiffs. After the Trump Administration, the EEOC's staffing was reduced by more than 40% and has a budget smaller than that over 40 years ago.¹⁰² Similarly, HUD lost approximately 20% of staffing from 2012 to 2019.¹⁰³ HUD Secretary Marcia L. Fudge has even stated "...we are at risk of not doing some things that we should do to make sure that our mission is completed."¹⁰⁴ There must be more resources allocated to these agencies to ensure that plaintiffs' statutory rights are restored.

There is currently little incentive for individual plaintiffs to bring claims due to the difficulty in winning on a claim and meager monetary rewards. Even more, federal agencies who would otherwise pick up this slack simply lack the resources to litigate on a plaintiff's behalf. Until this current regime undergoes significant reform, members of protected classes will continue to be deprived of their federal statutory rights. Although

¹⁰⁰ U.S. EQUAL EMP. OPPORTUNITY COMM'N, *supra* note 36.

¹⁰¹ Pavithra Mohan, *For workers alleging discrimination, a convoluted bureaucracy awaits*, FAST COMPANY (Aug. 6, 2021), <https://www.fastcompany.com/90660766/for-workers-alleging-discrimination-a-convoluted-bureaucracy-awaits>; *see also* Katy O'Donnell, 'Understaffed and overworked': Fudge calls for \$100B boost in HUD funds, POLITICO (Mar. 18, 2021, 3:30 PM), <https://www.politico.com/news/2021/03/18/marcia-fudge-hud-funding-477018>.

¹⁰² Mohan, *supra* note 101.

¹⁰³ Caitlin Reilly, *Senators press Fudge to rebuild HUD's depleted staff*, ROLL CALL (Jun. 10, 2021, 2:02 PM), <https://rollcall.com/2021/06/10/senators-press-fudge-to-rebuild-huds-depleted-staff/>.

¹⁰⁴ *Id.*

fortified case law would be mitigative in the short term, legislation from Congress would be invaluable.

C. Uniform Federal Artificial Intelligence Legislation

States have been active in regulating AI. Congress should follow suit and implement blanket legislation to cover all entities operating in the US. The US has signed onto a multitude of international agreements outlining principles to which companies should prioritize when developing these technologies. But these international treaties and domestic frameworks are mere voluntary guidelines. It is time to give them teeth through the democratic process.

The Algorithmic Accountability Act of 2022 is a great starting point. Initially proposed in 2019, Senators Wyden, Booker, and Clarke updated the landmark bill in 2022 after meeting with various experts. The Bill requires companies to assess the impacts of automated decision-making.¹⁰⁵ Overtly rooting the Bill in a concern for algorithmic discrimination, Senator Booker has stated that “...we have a responsibility to ensure that they are adequately assessed for biases that may disadvantage minority or marginalized communities...”¹⁰⁶

The original 2019 Bill seemingly emulates the EU risk-based approach from the Draft AI Regulation, which distinguishes high-risk AI systems from low-risk AI systems.¹⁰⁷ The 2022 Bill, however, omits its original “High-Risk” language, which should be reimplemented. The EU bifurcates by risk to impose different regulatory obligations, which allows for innovation in lower-risk systems while ensuring that higher-risk technologies, such as medical devices and machinery, are properly regulated. This approach is also supported by several academics who note that the lack of a risk-based approach might unduly stifle innovation in

¹⁰⁵ Algorithmic Accountability Act of 2022, S. 3572, 117th Cong. § 3(b)(1)(A)(i) (2022).

¹⁰⁶ Press Release, Ron Wyden, United States Senator for Oregon, Wyden, Booker and Clarke Introduce Algorithmic Accountability Act of 2022 To Require New Transparency And Accountability For Automated Decision Systems (February 3, 2022), <https://www.wyden.senate.gov/news/press-releases/wyden-booker-and-clarke-introduce-algorithmic-accountability-act-of-2022-to-require-new-transparency-and-accountability-for-automated-decision-systems>.

¹⁰⁷ *Compare* Algorithmic Accountability Act of 2022, S. 3572, 117th Cong. § 4(a)(3)(B) (2022) with *Commission Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, at 10, COM (2021) 206 Final (Apr. 21, 2021).

domains where AI does not pose a significant threat.¹⁰⁸ I urge future policymakers to wake up¹⁰⁹ and enact a coherent, uniform legislative structure to ensure marginalized communities are protected from this technology.

CONCLUSION

This Note highlights the ways in which artificial intelligence renders a disparate impact claim futile. There are already prohibitively high causation and statistical standards in making out a prima facie case of disparate impact that are only made more difficult to prove when the practice involves AI. Additionally, AI's unparalleled economic potential means a defendant will typically be able to satisfy their burden of proof. By contrast, a plaintiff might never be able to proffer a satisfactory less discriminatory alternative due to their lack of understanding of AI. I then propose several policy reforms, including a new evidentiary burden for defendants, a more coherent case law upon which plaintiffs can rely, and the endorsement of federal legislation.

The cat is out of the bag.¹¹⁰ AI will only become more prominent, so the courts, legislature, and executive branch all must remain vigilant in ensuring that this technology does not result in further discrimination. The courts promulgated disparate impact theory with the recognition that historical discriminatory practices echo through unintentional discrimination. New technologies cannot become barriers in fulfilling the purpose of the doctrine.

¹⁰⁸ See FURKAN GURSOY ET AL., A CRITICAL ASSESSMENT OF THE ALGORITHMIC ACCOUNTABILITY ACT OF 2022 4 (2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4193199.

¹⁰⁹ Broadbent & Arrieta-Kenna, *supra* note 14.

¹¹⁰ Mina Kim & Sara Mohamad, 'The Cat Is Out of the Bag': As DALL-E Becomes Public, the Possibilities — and Pitfalls — of AI Imagery, KQED (Sep. 26, 2022), <https://www.kqed.org/news/11926565/the-cat-is-out-of-the-bag-the-possibilities-and-pitfalls-of-ai-imagery>.