

## 6

# *Understanding United States anti-discrimination law*

In this chapter, we hope to give you an appreciation of what United States anti-discrimination law is and isn't. We'll use the U.S. legal experience as a case study of how to regulate discrimination. Other countries take different approaches. We don't aim to describe U.S. law comprehensively but rather give a stylized description of the key concepts.

We'll start with a history of how the major civil rights statutes came to be, and draw lessons from this history that continue to be relevant today. Law represents one attempt to operationalize moral notions. It is an important and illustrative one. We will learn from the way in which the law navigates many tricky tradeoffs. But we will also study its limitations and explain why we think algorithmic fairness shouldn't stop at legal compliance.

The final section addresses the specifics of regulating machine learning. Although U.S. antidiscrimination law predates the widespread use of machine learning, it is just as applicable if a decision maker uses machine learning or other statistical techniques. That said, machine learning introduces many complications to the application of these laws, and existing law may be inadequate to address some types of discrimination that arise when machine learning is involved. At the same time, we believe that there is also an opportunity to exercise new regulatory tools to rein in algorithmic discrimination.

This chapter can be skipped on a first reading of the book, but a few connections are worth pointing out. The first section elaborates on a central viewpoint of the book, especially Chapter 4, which is that attributes like race and gender are salient because they have historically served as organizing principles of many societies. That section also sets up Chapter 8 that conceives of discrimination more broadly than in discrete moments of decision making. The section on the limitations of the law motivates another core theme of this book, which is using the debates on machine learning and discrimination as an opportunity to revisit the moral foundations of fairness.

## *History and overview of U.S. anti-discrimination law*

Every inch of civil rights protections built into law was fought and hard won through decades of activism. In this section, we briefly describe these histories of oppression and discrimination, the movements that arose in response to them, and the legal changes that they accomplished.

## *Black Civil Rights*

The Black civil rights movement, often simply called the civil rights movement, has its roots in slavery in the United States and the rampant racial discrimination that persisted after its abolition. The period immediately following the American civil war and the abolition of slavery (roughly 1865-1877) is called the Reconstruction era. It resulted in substantial progress in civil rights. Notably, the Constitution was amended to abolish slavery (13th amendment), require equal protection under the laws (14th amendment), and guarantee voting rights regardless of race (15th amendment).

However, these gains were rapidly undone as White supremacists gained political control in the Southern states, ushering in the so-called Jim Crow era, a roughly 75-year period in which the state orchestrated stark racial segregation, discrimination, and near-total disenfranchisement of Black people. Nearly every facet of life was racially segregated, including residential neighborhoods, schools, workplaces, and places of public accommodation such as restaurants and hotels. This segregation was blessed by the Supreme Court in 1896, when it ruled that laws mandating segregation did not violate the Equal Protection clause under the “separate but equal” doctrine.<sup>1</sup> But in practice, things were far from equal. The jobs available to Black people usually paid far less, schools were underfunded and subject to closure, and accommodations were fewer and of inferior quality. As late as the 1950s, a cross-country drive by a Black person would have involved great peril, i.e., showing up at a small town at night and being refused a place to stay.<sup>2</sup> Black people could not democratically challenge these laws as the states erected numerous practical barriers to voting — ostensibly race neutral, but with vastly different effects by race — and Black people at the polls were often met with violence. As a result, disenfranchisement was highly effective. For example, in Louisiana until the mid 1940s, less than 1% of African Americans were registered to vote.<sup>3</sup> (Data limitations preclude a nation-wide assessment of the effectiveness of disenfranchisement.)

Meanwhile, in the Northern states, racial discrimination operated in more indirect ways. Residential zoning laws that prohibited higher-density, lower-cost housing were used to keep poorer Black residents out of White neighborhoods. The practice of “redlining” by banks, orchestrated to some extent by federal regulators, limited the availability of credit, especially mortgages, in specific neighborhoods.<sup>4</sup> The justification proffered was the level of risk, but it had the effect of discrimination against Black communities. Another prevalent technique to achieve segregation was the use of racially restrictive covenants in which property owners in a neighborhood entered into a contract not to sell or rent to non-White people.<sup>1</sup>

The civil rights movement emerged in the late 1800s and the early 1900s to confront these widespread practices of racism. Broadly, the movement adopted two complementary strategies: one was to challenge unjust laws and the other was

---

<sup>1</sup>This practice continued until the Supreme Court struck it down in 1948 (*Shelley v. Kraemer*), arguing that even though these were private contracts, if the state were to enforce them, it would violate the constitution’s Equal Protection clause.

to advance Black society within the constraints of segregation and discrimination. A key moment in the first prong was the formation of the National Association for the Advancement of Colored People in 1909. In addition to lobbying and litigation against Jim Crow laws, it sought to fight against lynching. Prominent efforts under the second prong included the Black entrepreneurship movement — 1900-1930 has been called the golden age of Black businesses<sup>5</sup> — and notable achievements in education. Many of the Historically Black Colleges and Universities were founded during the Jim Crow era.

After decades of activism, an epochal moment was a Supreme Court ruling in 1954 that declared the segregation of public schools unconstitutional. This began the gradual dismantling of the Jim Crow system, a process that would take decades and whose effects we still feel today. The court victories further galvanized the movement, leading to more intense activism and mass protests. This led to major federal legislation in the following decade: the Civil Rights Act of 1960 and the Voting Rights Act of 1965, both of which targeted voter suppression efforts, and the Civil Rights Act of 1964 and the Fair Housing Act of 1968 which targeted private discrimination. We will discuss the latter two in detail throughout this chapter.

Antidiscrimination laws were clearly a product of history and decades or centuries-long trends – slavery, Jim Crow, and the civil rights movement. At the same time, their proximate causes were often specific, unpredictable events. For example, the assassination of Martin Luther King Jr. provided the impetus for the passage of the Fair Housing Act. They also reflect political compromises that were necessary to secure their passage. For example, Title VII of the Civil Rights Act of 1964 created the Equal Employment Opportunity Commission; was stripped of the enforcement powers that had been present in the original wording of the title.<sup>6</sup>

### *Gender Discrimination*

The struggle for gender equality also has a long and storied history of activism. In the 1800s, the law did not recognize basic rights of women, including voting and owning property. Changing this was the primary goal of *first-wave feminists* whose strategies included advocacy, civil disobedience, lobbying, and legal action. The culminating moment was the ratification of the 19th amendment in 1920, guaranteeing women the right to vote (yet, as discussed above, Black women's right to vote was still limited in the South). *Second-wave feminism* began in the 1960s. It targeted stereotypes about the role of women in society, private discrimination in education and employment, and bodily rights, including reproductive rights and domestic violence. In the early post-war years, gender norms regressed in some ways (e.g., women lost access to jobs that had been available to them because of the war) which was arguably an impetus for the movement.<sup>7</sup> Two early legislative victories were the Equal Pay Act of 1963 and Title VII of the Civil Rights Act of 1964 prohibiting employment discrimination.<sup>2</sup>

---

<sup>2</sup>Although the latter was primarily a response to the Black civil rights movement, sex was added as a protected category in a last-minute amendment.

However, these did not initially have much impact due to the aforementioned lack of enforcement, and the movement only intensified. An important milestone was the founding of the National Organization of Women in 1965. Borrowing strategies from the Black civil rights movement, the second-wave feminists adopted a plan to litigate in the courts to secure protections for women. A notable court victory in the following decade was the expansion of abortion rights by the Supreme Court.<sup>8</sup> On the legislative front, two major achievements for gender equality were Title IX of the Education Amendments Act of 1972 that prohibited sex discrimination in federally funded educational programs, and the Equal Credit Opportunity Act that prohibited sex discrimination in credit.

Education, especially higher education, and credit were both important sectors for women's rights. Historically, many elite colleges simply did not accept women. Even in the 1970s, women faced many barriers in academia: sexual harassment, higher bars for admission, outright exclusion from some high-status fields such as law and medicine, and limited athletic opportunities. Similarly, credit discrimination in the 1970s was also stark, such as requiring women to reapply for credit upon marriage, usually in the husband's name.<sup>9</sup> After this period, the focus of the feminist movement expanded beyond major legislative victories to include the questioning of gender as a social construct.

### *LGBTQ Civil Rights*

Discriminatory laws against LGBTQ people were historically numerous: prohibition of some sexual behavior (i.e., anti-sodomy laws<sup>10</sup>), lack of marriage rights, bans on military service and some other government positions, a failure to prohibit private discrimination and to treat hate crimes as such, and even a prohibition of literature advocating for gay rights under obscenity laws.

Tentative activism began in the 1950s with the first legal changes coming in the early 60s. A pivotal movement was the 1969 Stonewall riots, a series of protests in response to a police raid at a New York City gay bar. The aftermath of this event kickstarted the push for U.S. LGBTQ rights, including the gay pride movement for visibility and acceptance. In 1973, the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* dropped homosexuality as a disorder, signaling (and furthering) a major shift in attitudes. The list of legal changes is long and ongoing. They include state-by-state changes to laws involving sodomy, marriage equality, private discrimination, and hate crimes; a 2003 Supreme Court decision ruling anti-sodomy laws unconstitutional;<sup>11</sup> and a 2015 Supreme Court decision guaranteeing the right to marry for same-sex couples nationwide.<sup>12</sup> In parallel, the push for LGBTQ rights in the private sector has progressed in part by interpreting existing statutory prohibitions on sex discrimination, such as Title VII of the Civil Rights Act of 1964, to encompass sexual orientation and gender identity discrimination.<sup>13</sup>

## *Disability laws*

Another dimension of identity covered by anti-discrimination statutes is disability. Over a quarter of adults in the United States today have some type of disability, including mobility disabilities, blindness or other visual disability, deafness or other hearing disability, and cognitive disabilities.<sup>14</sup> These and other disabilities are distinct identities corresponding to different lived experiences and, sometimes, cultures.<sup>3</sup> Still, the emergence of a cross-disability coalition and identity enabled more effective advocacy for disability rights. This movement gained steam in the decades following World War II. Activists aimed to make disability visible, rather than stigmatized, pitied, and hidden, and sought to achieve independent living. Like other rights movements, disabled people faced multiple, mutually-reinforcing barriers: society's attitudes towards disability and disabled people, the lack of physical accommodations and assistive technologies, and discriminatory policies.<sup>16</sup> Attitudes that held back disabled people weren't just prejudice, but also mistaken views of disability as residing in the person (the *medical model*) instead of, or in addition to, being created by barriers in society (the *social model*). The first federal law protecting disability rights was the Rehabilitation Act of 1973 which prohibited disability discrimination in federally funded programs. Activism toward a broad civil rights statute continued, with the 1964 Civil Rights Act as a model. These efforts culminated in the Americans with Disabilities Act (ADA) of 1990. While the ADA has many similarities to the other civil rights statutes, it also has major differences due to its emphasis on accommodation in addition to formal nondiscrimination.

## *Lessons*

The histories of the various civil rights movements hold several lessons that continue to be relevant today. First of all, the law is a political instrument: it can be used to discriminate, to create the conditions under which discrimination can flourish, or to challenge discrimination. It can be a tool for subjugation or liberation. Laws may be facially neutral but they are created, interpreted, and enforced by actors that respond to the changing times and to activism. Court decisions are also influenced by contemporary activism and even scholarship.

Our brief historical discussion also helps explain why certain sectors are regulated, and not others. Education, employment, housing, credit, and public accommodation are domains that are both highly salient to people's life courses and have had histories of discrimination that were deliberately used to subordinate some groups.<sup>4</sup>

One consequence of this sector-specific approach is that the law can be tailored to the particularities of the sector in an attempt to avoid loopholes. For example, the Fair Housing Act encompasses the full range of practices related to housing

---

<sup>3</sup>Notably Deaf culture; for an introduction see<sup>15</sup>

<sup>4</sup>In addition, there are constitutional limitations on the ability of Congress to regulate private discrimination.

including sales, rentals, advertising, and financing. It lists (and prohibits) various ways in which housing agents may subtly mislead or discourage clients belonging to protected classes. Recognizing the importance of financing for securing housing, it prohibits discrimination in financing with respect to “purchase, construction, improvement, repair, or maintenance”. It even prohibits ads indicating a discriminatory preference. And that includes not just categorical statements such as “no children”, but also targeting of ads to certain geographic regions in a way that correlates with race, and the selection of actors used in advertising.

In many cases these attempts to avoid loopholes have held up well in the face of recent technological developments. The prohibition on discriminatory advertising has forced online ad platforms to avoid discriminatory targeting of housing ads.<sup>17</sup> But this is not always so. Ride hailing platforms are able to evade Title VII (employment discrimination) liability even though they terminate drivers based on the (potentially discriminatory) ratings given by passengers.<sup>18</sup>

Even though laws are sector-specific, it is hard to understand discrimination by looking at any one set of institutions (such as employment or education, much less a single organization) in isolation. History shows us that there tend to be multiple interlocking systems of oppression operating in tandem, such as federal housing policy and private-sector discrimination. Similarly, the line between state and private discrimination is not always clear.

History also shows that when disrupted, hierarchies tend to reassert themselves by other means. For example, the end of de-jure segregation accelerated the phenomenon of “White flight” from cities to the suburbs, exacerbating de-facto segregation. Not only is progress fitful, regression is possible. For example, Woodrow Wilson and his administration segregated large parts of the federal workforce in the 1910s, eroding some of the gains Black people had made in previous decades. And as we were writing this chapter, the Supreme Court reversed *Roe v. Wade*, ending federal protection of abortion rights and enabling severe restrictions on abortion in many states.

Another important point that is not apparent from the laws themselves is that the various protected dimensions of identity have complex and distinct histories of discrimination and activism, even if statutes attempt to treat them all in a uniform and formal way. Even within a single dimension like ethnicity, the oppression and struggles of different groups take drastically different forms. Native Americans endured a century of attempts at forced assimilation in which children were sent to boarding schools and asked to abandon their culture. The Chinese Exclusion Act of 1882 all but eliminated the immigration of Chinese people for over half a century and made conditions inhospitable for the Chinese immigrant community that already existed. During World War II, over 100,000 people of Japanese ancestry, the majority of whom were U.S. citizens, were interned in concentration camps under the pretense that they were disloyal to the country. These are just a few of the more gruesome episodes of discrimination on the basis of race, ethnicity, and national origin in U.S. history, focusing on the actions of the government. National origin discrimination was often a thinly veiled form of racial discrimination. Thus, although the list of protected attributes in the law may grow over time, it is not

arbitrary and is deeply informed by history.<sup>5</sup>

Equality under the law remains a contested and evolving notion. This is especially the case when antidiscrimination runs up against some countervailing value or principle, such as religious freedom or limiting state authority. And because the law is intertwined with our lives and livelihoods in so many ways, equality under the law, in a broad sense, requires far more than formal nondiscrimination. Consider gender equality. The range of legal interventions necessary to achieve it is long and growing. Beyond voting rights and prohibition of sex and gender discrimination, it includes prohibition of pregnancy and marital status discrimination, curbing sexual harassment and sexual violence, abortion rights, maternity leave laws, and childcare subsidies. Each one of these battles has many fronts. For example, the #MeToo movement brought to light the role of non-disparagement clauses by employers in settlements to silence victims of workplace sexual harassment, and there is an ongoing effort to prohibit such clauses.

Finally, legal change is not the end of the road but in some ways the beginning. The effects of past discrimination tend to leave a lasting imprint. The law itself, given political realities, can only do so much to erase the effects of that history.

Table 1: A summary of the major anti-discrimination statutes: Titles VI and VII of the Civil Rights act of 1964, the Fair Housing Act, Title IX of the Education Amendments Act of 1972, the Equal Credit Opportunity Act, and the Americans with Disabilities Act.

Law	Year	Covered entities and regulated activities	Protected categories (* = added later)
Title VII	1964	Employers, employment agencies, labor unions	Race, color, religion, sex, national origin, pregnancy*
Title VI	1964	Any organization receiving federal funding (due to breadth, doesn't list regulated activities)	Race, color, national origin
FHA	1968	sales, rentals, advertising, and financing of housing	Race, color, religion, national origin, sex(*?), handicap, familial status.
Title IX	1972	Educational programs receiving federal funding: hiring, pay, rank, sexual harassment, retaliation, segregation & same-sex education	Sex

<sup>5</sup>In the context of equal protection doctrine, the Supreme Court has explicitly listed the criteria that qualify a trait for protection ("heightened scrutiny"): a history of past discrimination, political powerlessness, the irrelevancy of a trait to an individual's ability to contribute to or participate in society, and immutability.<sup>19</sup>

Law	Year	Covered entities and regulated activities	Protected categories (* = added later)
ECOA	1974	Creditors (Banks, small loan and finance companies, retail and department stores, credit card companies, and credit unions.)	Race*, sex, age*, national origin*, marital status, receipt of public assistance*
ADA	1990	Employers, public services, public accommodation	Disability; record of disability; perception of disability

### *A few basics of the American legal system*

The U.S. Constitution is the ultimate law of the land. The constitution created the three branches of government: the legislature (Congress), the executive (the president, executive agencies, and others reporting to the president), and judiciary (the Supreme Court and other courts). All three branches have important roles when it comes to antidiscrimination law. State and local governments and laws also play important roles in antidiscrimination, but we will say less about them due to our pedagogical focus on federal law.

Before we get to the three branches, it is worth noting that the constitution itself contains two elements relevant to discrimination law: the right to the due process of law (fifth and fourteenth amendments) and the right to equal protection under the laws (fourteenth amendment). Both of these curtail the ability of the government to discriminate. Equal protection law has also sometimes been used to curtail private discrimination. Due process has been raised as a defense by *defendants* of discrimination lawsuits contending that laws that curtail their ability to discriminate violate their due process rights.

#### *The role of Congress (legislative branch)*

Laws passed by Congress are called statutory laws, as opposed to constitutional law, case law, and other types of law. We encountered some of the major antidiscrimination statutes earlier. But there are many practical and political barriers to congressional action, and statutes or amendments are relatively rare. Thus, to stay relevant in a changing world, laws are generally broadly worded policies and do not attempt to anticipate the nuances of every situation in which they might be applied. To interpret and enforce these policies, Congress delegates authority to federal agencies such as the Equal Employment Opportunity Commission (established by the Civil Rights Act of 1964). The courts also perform a vital interpretive function, as well as keeping a check on the power of Congress itself.

There are three main law making powers that Congress has used to enact antidiscrimination statutes within the limits of its constitutional authority. The first is the Commerce Clause, which allows Congress to regulate interstate commerce. The meaning of this clause has been interpreted expansively by the Supreme



Court.<sup>6</sup> Legislation pertaining to employment and credit antidiscrimination finds basis in the Commerce Clause.<sup>7</sup> The second power comes from the Fourteenth Amendment, which guarantees to all citizens the equal protection of the laws, and further empowers Congress to enforce it through appropriate legislation. While the extent of state involvement needed for the Fourteenth Amendment to apply is not settled, the Fair Housing Act and the Americans with Disabilities Act owe their constitutional basis in part to this power. Finally, Congress has the “power of the purse”: the ability to enact policy goals by controlling spending, and by threatening to withhold federal funding for entities that fail to meet certain obligations. Title VI of the Civil Rights Act of 1964 and Title IX of the Educational Amendments Act of 1972 fall under this category, which is why they only covered organizations receiving federal funding.

Congress has used its power to enact antidiscrimination statutes covering a broad swath of activities. Still, there are many gaps and limitations in federal antidiscrimination law, in part because of constitutional limitations and in part because Congress has failed to act. As a result, state laws sometimes fill these gaps.

### *The role of the courts (judicial branch)*

The United States adopts a common-law system, which means that courts have the power to make law that guides decisions in future cases. This is the concept of precedent. In disputes where the facts or principles are similar to previous cases decided by relevant courts, judges are bound to follow the reasoning used in the past decision (the precedent). Similarly, courts are tasked with interpreting the statutory laws and the Constitution. This body of precedent is referred to as case law and can be as binding as any other law. For example, the important concept of disparate impact, under which decision making practices may be unlawful if they have disproportionate effects even if facially neutral and without discriminatory intent, is the result of a Supreme Court decision interpreting the scope of a statutory law. Most of Europe, in contrast, adopts a civil law system which means that legislation is the primary source of law and judicial decisions have less value as precedent.

The hierarchical organization of courts determines which precedents are binding on a particular dispute. The federal courts are organized into three levels: the district courts at the bottom, thirteen appellate courts (also known as circuit courts) above them, and the Supreme Court at the top. Supreme Court decisions are binding on all lower courts and appellate court decisions are binding on the corresponding district courts.<sup>8</sup> The appellate courts only hear cases on appeal —

---

<sup>6</sup>Between 1937 and 1995, a period that includes all the statutes discussed above, not a single statute (whether or not pertaining to discrimination) was invalidated by the Court on the ground that Congress exceeded its power under the Commerce Clause.

<sup>7</sup>When the Civil Rights Act of 1964 was enacted, its Title II, which prohibits discrimination in public accommodation, was famously challenged as unconstitutional by the owner of motel in Atlanta, Georgia. The Supreme Court upheld the constitutionality of Title II, in a case that has since had tremendous precedential value.<sup>20</sup>

<sup>8</sup>All of these are part of the federal court system. The state courts are separate and in fact hear

that is, when one of the parties alleges a material error in a district court's decision. The Supreme Court, in turn, usually only hears appeals of circuit court decisions. The Supreme Court is not required to accept petitions for review; in fact, it only grants review in a small fraction of requests.

When interpreting statutes, courts adopt both textual and contextual methods. The former confines itself to the plain meaning of the statute itself, while the latter looks to sources outside the text of the law such as lawmakers' originally expressed intent and the law's purpose. The importance of contextual factors is a contested topic and judges differ in their approaches.

So far we've talked about the role of the courts in making law. Of course, the primary function of the courts is to adjudicate individual cases. So a few notes about court procedure are in order. Litigation can be *civil* or *criminal*. Civil cases involve wrongs against private individuals; most discrimination-related disputes fall into this category, with a few exceptions such as hate crimes. Criminal cases involve violations of criminal law and can only be brought by the government.<sup>9</sup>

A central feature of U.S. court procedure is the use of *adversary proceedings*. The two parties to a dispute are the *plaintiff* (who files a *complaint* alleging that they have been wronged) and the *defendant* (who is alleged to have committed the wrong). Both are typically represented by attorneys, who have a lot of power in determining how the case unfolds, with the judge having a relatively passive role as an arbiter and not an inquisitor.

An example will bring together the aspects of the court system that we've discussed so far. Consider the question of whether the websites of restaurants, retailers, etc. must be accessible to visually impaired people. The Americans with Disabilities Act prohibits excluding people with disabilities from availing of the services of a place of public accommodation. But does this include websites? Congress could not have anticipated this question in 1991, so the statute (despite being unusually detailed, running to over 20,000 words) does not address this question directly. One set of circuit courts has looked at Congress's intent and purpose and found that websites themselves can be considered places of public accommodation in keeping with the ADA's "broad mandate", "sweeping purpose", and "comprehensive character". Another set of circuit courts took a more textual approach, and considers it crucial that the statute applies to the services *of* a place of public accommodation, not services *in* a place of public accommodation. Thus, as long as there is a sufficient "nexus" between the physical place and the website, the accessibility requirement extends to the website. In April 2021, one circuit court ruled differently, reading the text of the statute to mean that only physical places can be places of public accommodation and also rejecting the "nexus" standard adopted by the second set of courts. When circuit courts are split in this way, it

---

the great majority of cases, but the federal courts have jurisdiction in disputes involving federal law, and are thus of more relevance to us.

<sup>9</sup>What makes an act a crime rather than a civil offense (other than being legally classified as such) is a deep question. One difference is the perceived seriousness of the wrong, deserving of retribution against the perpetrator and not just the victim being made whole. Another is the nature of the wronged party. Crimes can be thought of as an offense against the state or against society, which the state has an interest in preventing in order to avoid a breakdown of social order.

usually takes the Supreme Court stepping in to resolve the inconsistency, but this may take many years.

One reason for this state of affairs is that the Department of Justice, which is tasked with issuing regulations to implement the ADA, hasn't issued a final regulation on whether websites are places of public accommodation and, if so, what standard of accessibility they would need to satisfy. Therefore the courts had to exercise a greater degree of interpretive latitude than they otherwise would. It also gave rise to a concern in some courts that imposing accessibility requirements without setting a clear standard would deprive defendants of their constitutional right to due process. This highlights the importance of executive departments and federal agencies, to which we will turn next.

### *The role of the federal agencies (executive branch)*

The main anti-discrimination functions of the federal agencies are rulemaking, guidance, and law enforcement. For example, the Equal Credit Opportunity Act broadly makes credit discrimination unlawful, but leaves it to the Federal Reserve to draft and interpret regulations that implement this mandate (Congress later transferred this authority to the Consumer Financial Protection Bureau). This process is called *rulemaking*. The resulting regulations constitute *administrative law* and have the force of law alongside statutory law and case law.

Rules differ slightly from guidelines. A group of agencies led by the Equal Employment Opportunity Commission issued the *Uniform Guidelines for Employee Selection Procedures* in 1978 that spell out a framework for ensuring that tests and other employee selection procedures are compliant with VII of the Civil Rights Act of 1964. The Uniform Guidelines are widely relied upon by employers. But the Uniform Guidelines do not constitute law. They are often referenced in court opinions, and courts generally give significant deference to agency guidelines, but courts are not bound by them.

It's hard to overstate the practical importance of the agencies. Whether or not a statute has real teeth depends in large part on the implementing agency. The EEOC initially refused to take up gender discrimination despite being empowered to do so. In fact, Title VII had no teeth for even racial discrimination until amended in 1972 to empower the EEOC to take action (the Equal Employment Opportunity Act of 1972).<sup>21</sup>

Agencies differ in their level of political independence; some are housed within the executive (such as the Department of Housing and Urban Development and the Department of Labor) while others are more independent (such as the EEOC and the Federal Trade Commission (FTC)). The latter have enforcement powers in addition to rulemaking powers. They can conduct investigations and file suit in court; some even have their own judicial systems and are sometimes called a *de facto* fourth branch of government. In short, interpreting as well as enforcing the statutes are both tasks shared by the federal agencies and the courts. They generally work harmoniously together,<sup>10</sup> but the proliferation of sources of law

---

<sup>10</sup>However, this has been changing since the Trump administration.<sup>22</sup>

and methods of enforcement can lead to inefficiency and confusion.

It is worth mentioning two other important sources of policy: executive orders and rules internal to institutions. Executive orders are directives issued by the president of the United States. Originally intended as a way to manage the affairs of the government, the vast reach of the federal government has meant that they are a powerful de-facto for enacting policy. For example, precursors to Title VII in the form of executive orders go as far back as 1941.<sup>23</sup> Although much weaker in scope than the eventual legislation, they illustrate the ability of presidents to act quickly while Congress might be stalled.

Institutions, whether public or private, may set non-discrimination rules or guidelines for their employees that may go beyond the requirements of the law. For example, asking a job candidate about their marital status is not per-se unlawful. However, it would be construed as evidence of intent to discriminate in a legal dispute.<sup>24</sup> Considering this (and the fact that there is almost never a job-related reason for such an inquiry), many organizations prohibit their interviewers from asking such questions. On a day-to-day basis, these institutional guidelines are the most direct nondiscrimination rules that individuals are bound by.

### *Case study: the evolution of Title IX*

Title IX of the educational amendments act of 1972 prohibits sex discrimination in educational institutions that receive federal funds. In 1975, the Department of Health, Education and Welfare (HEW) published the final regulations detailing how Title IX would be enforced. Since 1975, the federal government has issued guidance clarifying how it interprets and enforces those regulations.

Two of the big questions surrounding Title IX were what constitutes receiving federal financial assistance and what constitutes sex discrimination. Each of these had the potential to vastly impact the reach of the law. In 1984, the Supreme Court ruled that Title IX was program specific: that is, only those programs and activities receiving direct federal funds needed to comply. This gutted the application of Title IX: for example, most athletic programs were no longer covered since they didn't directly receive federal funds. In response, Congress drafted a bill specifically intended to overturn this decision, restoring the broad scope of Title IX, which it retains to this day.<sup>25,26</sup><sup>11</sup>

Two other major disputes regarding the scope of Title IX are whether schools are responsible for sexual harassment that happens on campuses and whether discrimination on the basis of sexual orientation and gender identity is prohibited. Unlike the coverage question, these continue to be the topic of vociferous legal and political debate. On the sexual harassment front, the Supreme Court held that in the late '90s that schools are responsible for creating a safe environment, including

---

<sup>11</sup>Incidentally, the institution at the center of this controversy was Grove City college, a conservative Christian college which did not directly accept any federal money in an attempt to retain its autonomy. After 1988, it would have been covered under Title IX because its students were recipients of federal loans and grants, illustrating the broad reach of the state. To this day, the college escapes Title IX responsibilities by prohibiting its students from receiving any federal student loans or grants (such as Pell grants).

preventing harassment by other students, but “the student must show that an official of the school with authority to respond actually knew of and was deliberately indifferent to the harassment”. The Obama, Trump, and Biden administrations have all introduced guidelines or regulations on this question, in turn expanding and contracting the scope of Title IX.<sup>27</sup> A similar political seesaw has played out with respect to LGBTQ protections, with the latest move being an expansive interpretation in 2021 by the Department of Education, buttressed by a 2020 Supreme Court ruling also involving the relationship between sex discrimination and sexual orientation, but in the employment discrimination context.<sup>28</sup>

### *How the law conceives of discrimination*

There are many possible ways to define discrimination and attempt to achieve nondiscrimination. In this section, we will discuss how the law conceives of discrimination and how it tries to balance nondiscrimination with other ideals.

#### *Disparate treatment and disparate impact*

Imagine an employer turning down a job candidate and explicitly informing them that this decision was on account of a protected characteristic. Such a case would be relatively straightforward to adjudicate based on the text of the statute itself (“It shall be an unlawful employment practice for an employer to fail or refuse to hire an individual. . . because of such individual’s race, color, religion, sex, or national origin.”) However, in most cases of discrimination, the decision maker’s behavior is less explicit and the evidence is more circumstantial. To deal with these, courts have created two main doctrines called disparate treatment and disparate impact.

Disparate treatment refers to intentional discrimination and roughly matches the average person’s conception of discriminatory behavior. It subsumes the straightforward case described in the previous paragraph. For more circumstantial cases, the Supreme Court has established a so-called burden shifting framework under Title VII (employment law). First, the plaintiff must establish a “prima facie” case of discrimination by showing that they are a member of a protected class, was qualified for a position, was denied it, and the position then remained open or was given to someone not in the protected class. If the plaintiff is successful at this, the employer must produce a legitimate, non-discriminatory reason for the adverse decision. The plaintiff then has the burden of proving that the proffered reason is merely a pretext for discrimination.<sup>12</sup>

Disparate treatment usually involves reasoning about what action the defendant would have taken if the plaintiff’s protected characteristic had been different, with all other facts of the case unchanged. Elsewhere in this book we argue why, from a technical perspective, these “attribute-flipping” counterfactuals are at odds with a nuanced understanding of causality and result in brittle tests for discrimination. In

---

<sup>12</sup>More precisely, this is only one of the possible frameworks; we describe it for illustrative purposes. The full picture—including the game of which framework to choose—is a vast morass.<sup>29</sup>

any event, the importance of causality in disparate treatment, especially so-called but-for causation, has increased following a 2020 Supreme Court decision<sup>13</sup> which held that it is impossible to engage in sexual orientation discrimination without engaging in sex discrimination by imagining a counterfactual in which the victim's sex is changed without affecting anything else, including gender preference. While celebrated from a civil rights perspective because of its implications for LGBTQ rights, we should keep in mind that this represents a narrow understanding of causality and its application in other scenarios may yield conclusions not so favorable to civil rights.

In contrast to disparate treatment, disparate impact is about practices that have a disproportionate effect on a protected class, even if unintentional. At a high level, disparate impact must be both unjustified and avoidable. This is again operationalized through a burden-shifting framework. First, the plaintiff must establish that there is a disproportionate difference in selection rates between different groups. If that can be shown, then the employer has the opportunity to explain if the reason for the different selection rates has a business justification. The burden then reverts to the plaintiff to show that there is an "alternative employment practice" that would have achieved the employer's aims while being less discriminatory.

One way to think about disparate impact is as a way to "sniff out" well-concealed intentional discrimination by putting the focus on its impacts, which are more readily observable. Indeed, the case that led to the doctrine involved an employer that introduced aptitude tests for promotion on the very day that the Civil Rights Act of 1964, that prohibited employment discrimination based on race, took effect.<sup>30</sup>

But disparate impact is also thought to be motivated by a consideration of distributive justice, that is, minimizing unjustified inequalities in outcomes. In this sense, disparate impact roughly corresponds to the middle view of equality of opportunity that we discussed in Chapter 4. Disparate impact tries to force decision-makers to treat seemingly dissimilar people similarly on the belief that their current dissimilarity is a result of past injustice. It aims to compensate for at least some of the disadvantages suffered for unjust reasons. Indeed, in the case mentioned above, the Supreme Court pointed out that the racial performance disparity on aptitude tests could be explained by inequalities in the educational system. But disparate impact doctrine has evolved over the years and the extent to which it reflects distributive justice, as opposed to a device for illuminating well-concealed discrimination, is thought to have waned over time.

While we have discussed these two doctrines in the context of employment law, they are found in each of the six domains we discussed in the first section. Disparate impact has been so central to the legal understanding of discrimination that it was later incorporated into statutes, notably the ADA (disability law), but also into Title VII (equal employment law) itself through a 1991 amendment. But the Supreme Court has not extended the doctrine to situations where the laws or procedures of the state (rather than private actors) violate the Equal Protection Clause if they have a discriminatory purpose. In other words, there is no equivalent



of disparate impact doctrine for state actors, only disparate treatment.

An important general observation about antidiscrimination law—especially for readers who may be accustomed to thinking about fairness in terms of the statistical properties of the outputs of decision making processes—is that the law is primarily concerned with the processes themselves. Relatedly, the way that courts go about weighing evidence is also highly procedural, to the point where it may seem tangential to the substantive question of whether discrimination took place.<sup>31</sup> Even disparate impact, despite being motivated in part by distributive notions of justice, is treated in a formal and procedural manner. As an illustration of the centrality of the procedural element, the DOJ’s legal manual for proving Title VI disparate impact claims is over 20,000 words long.<sup>32</sup>

There are many possible reasons for the law’s focus on process. One is historical: the statutes were primarily responding to blunt discrimination and formal denials of opportunity as opposed to more subtle statistical phenomena. It is also a better match for how the law works: the definition of discrimination cannot be divorced from the procedure for proving it in court. A third reason is political: it is easier to achieve consensus about fair processes than what the right distributive outcome is. Finally, at a pragmatic level, it reflects the law’s attention to the nuances of workplaces and other institutions, compared to which statistical fairness criteria seem crude and oversimplified.

### *Avoiding excessive burdens on decision makers*

A recurring theme is how much burden on decision makers is justified in pursuit of fairness goals. For example, making accommodations for disabled employees results in some cost to a firm.

In general, the law gives substantial deference to the interests of the decision maker. This has been repeatedly made clear by lawmakers and the courts at various points in time. For example, the House Judiciary Committee said on the role of the EEOC at the moment of the agency’s inception: “management prerogatives, and union freedoms are to be left undisturbed to the greatest extent possible.”<sup>33</sup> The Supreme Court clarified in 2015 that “The FHA (Fair Housing Act) is not an instrument to force housing authorities to reorder their priorities.”<sup>34</sup>

One exception is the Americans with Disabilities Act which imposes substantial compliance requirements on a large set of firms and governments. This shouldn’t be a surprise since the law sought to create structural changes in society, especially to the built environment. The ADA does have an “undue hardship” defense to the requirement that employers provide “reasonable accommodations” to qualified employees with disabilities, but courts appear to tolerate a higher burden than under, say, Title VII.<sup>13</sup> As an illustration, blind employees sued their employers for failing to provide paid readers for four hours of the workday; the court sided with

---

<sup>13</sup>The ADA originally called for a higher bar: an accommodation would have to threaten the continued existence of the employer’s business.<sup>35</sup> In fact, the part of the law that applies to public services does have a higher bar for the defendant: the accommodation must fundamentally alter the nature of the service or program.

the plaintiffs, implying that a roughly 50% increase in the cost of the employees to the employer did not constitute an undue hardship.<sup>35</sup> We hasten to add that undue hardship is a multi-factor test and there is no clear or uniform cost threshold; cost is rarely the determinative factor.

There are a range of potential justifications of the burdens on decision makers in the academic scholarship and legislative history. Often the responsibilities of decision makers are justified by appeal to the human rights of those being harmed, rather than an economic analysis. For example, the ADA sought to intervene in discrimination against disabled people that often affected their livelihoods and sometimes cost lives.<sup>35</sup> Alternatively, burdens are sometimes justified because they pose only a “de minimis” cost. For example, Title VII requires employers to make accommodations for employees’ religious beliefs, but not if it would pose more than a minimal cost.

In between these two types of cases lie a variety of others where a more careful balance between benefits and costs is necessary.<sup>14</sup> We give a few brief examples here.

- *Positive externality.* A hope behind the Americans with Disabilities Act was that it would make it easier for disabled people to enter the workforce and contribute to the overall economy.<sup>15</sup>
- *Regulation as collective action.* Title II of the Civil Rights Act of 1964 prohibits discrimination in places of public accommodation (e.g., restaurants). A major reason why such establishments discriminated against minorities is due to the prejudices of their White customers. Title II enabled them to stop discriminating, gaining business from minority customers without incurring lost business from their White customers ; thus, the law did not impose a burden on them but rather created an opportunity.<sup>36</sup> Similarly, consider insurance. In the absence of regulation, if an insurer avoids calibrating premiums to risk in the interest of fairness, it might go out of business. But if all firms in the market have their behavior constrained by antidiscrimination law in the same way, they can no longer claim to be at a competitive disadvantage.
- *Cheapest cost avoider.* The cheapest cost avoider or least cost avoider principle assigns liability from a harm to the party that can avoid the harm at the lowest cost. It is the reason why firms, to some extent, bear liability for discrimination or harassment committed by their employees. If an employer is forced to internalize the costs of discriminatory harassment committed by its employees, it will, on standard economic theory, invest in precautions up to the point where they are no longer cost justified.<sup>37</sup>
- *Correcting irrationality.* Some commentators suggest that the rampant discrimination against women before the passage of the ECOA was irrational behavior by creditors, and women were in fact good credit risks.<sup>38</sup> In this

---

<sup>14</sup>These types of questions are studied in the field of law and economics, which applies microeconomic theory to explain the effects of laws.

<sup>15</sup>In fact, the fraction of disabled individuals employed has decreased since the passage of the ADA, although the causal effect is far from clear.



view, ECOA can be seen as correcting this irrationality rather than imposing a burden on creditors.

### *Limits of the law in curbing discrimination*

How effective has United States antidiscrimination law been? The best-case scenario is that the possibility of penalties has sufficiently deterred would-be discriminators that rates of discrimination have plummeted, and, in the few remaining cases of discrimination, victims manage to obtain redress through the courts. The worst-case scenario is that the laws have had virtually no effect, and any reductions in disparities since their passage can be attributed to other factors such as discriminators being less successful in the market.

The reality is somewhere in between. Rigorously evaluating the effect of laws is a tricky counterfactual problem and is subject to much uncertainty and debate. However, there is much evidence suggestive of a positive effect. For example, one study used a natural experiment to evaluate the impact of Title VII on job opportunities for African Americans relative to White Americans. It showed that the relative employment of African Americans increased more in industries and regions with a greater proportion of firms that were newly covered under Title VII by the Equal Employment Opportunity Act of 1972.<sup>39, 21</sup>

While the gains have been non-negligible, the effectiveness of anti-discrimination law is blunted for many reasons which we now discuss. This motivates our view that work on algorithmic fairness should not treat the approach adopted in antidiscrimination law as a given, but should instead reconnect with the moral foundations of fairness.

### *Burdens on victims of discrimination*

The law places an array of burdens on victims of discrimination if they wish to seek legal recourse. We will use the labor market as our running example, but our observations apply to other contexts as well.

To begin with, legal intervention is initiated by the victim, not the government,<sup>16</sup> and cannot begin until after victims have already suffered discrimination. Regulators do not prospectively review employment practices, in contrast to other areas of law such as pharmaceutical regulation where drugs must be thoroughly tested before being allowed on the market. Further, there is a fundamental information asymmetry between firms and employees (or job candidates). Victims may not even be aware that they have faced discrimination. After all, job candidates and employees have no direct visibility into employers' decision making process and

---

<sup>16</sup>There are limited exceptions to this general principle, such as the ability of enforcement agencies to bring "pattern-or-practice" cases against repeat discriminators. A prominent example is the Department of Justice settlement against the Pennsylvania Police Department. U.S. v. Pennsylvania & Pennsylvania State Police, no.1:14-cv-01474-SHR (M.D. Pa. July 29, 2014), <https://www.justice.gov/sites/default/files/crt/legacy/2014/07/31/pennsylvaniapdcomp.pdf>.

firms need not provide a justification for an adverse hiring or promotion decision. Only in the domain of credit does some form of transparency requirement exist.<sup>17</sup>

Even if a victim becomes aware of discrimination, they face barriers that may deter them from suing. Litigation involves additional mental anguish. Victims may also be deterred by the high financial costs of litigation.<sup>18</sup> Lawsuits typically take several years to reach a conclusion, by which time the victim's career may suffer a significant and irreparable setback. If the victim remains at the firm after filing suit, they face an uncomfortable situation in the best case, and potentially retaliation from the employer (even though laws specifically prohibit retaliation, it remains a common result of discrimination lawsuits). And if the victim seeks employment elsewhere, future employers may weigh negatively the fact that the candidate sued a previous employer for discrimination.

Victims who decide to sue face a battery of procedural hurdles. If the employer has internal grievance procedures, the victim may be required to try those before suing (or risk losing her claims). Another prerequisite to filing suit is to file an administrative complaint with the Equal Employment Opportunity Commission promptly after the discrimination starts. The timeliness requirement often puts victims in a double bind because of the need to exhaust internal channels. It also makes it difficult to collect the evidence necessary to prevail in court.<sup>41,42</sup>

That brings us to the final and most serious difficulty that plaintiffs face, which is the burden of proof. To be sure, the standard of proof that the plaintiff must meet in discrimination cases is "preponderance of the evidence", which means more likely than not, which is lower than the standard in criminal cases. But even this standard has proved daunting. According to Katie Eyer, "anti-discrimination law is a highly rigid technical area of the law, in which any of a myriad of technical doctrines can lead to dismissal. Courts approach the question of discrimination as if it were a complex legal puzzle, in which any piece out of place must result in the dismissal of the plaintiffs' claims."<sup>31</sup>

Specifically, in disparate treatment cases, courts have created numerous defendant-friendly doctrines. Under the "stray remarks" doctrine, discriminatory comments made by the employer about the plaintiff do not constitute evidence of discriminatory intent unless there is a sufficiently clear causal nexus to the decision itself. Under the "same actor" defense, if the employer was willing to hire the plaintiff at a previous time, it is taken as evidence that the employer bears no discriminatory intent against the plaintiff. Under the "honest belief" rule, a case can be summarily dismissed if the employer "honestly believed" in the reasons for the decision, even if they can later be shown to be "mistaken, foolish, trivial, or baseless."

In disparate impact, an overlapping set of factors is arrayed against the plaintiff.

---

<sup>17</sup>However, the turn to algorithmic tools in hiring has opened the possibility of transparency requirements and ex-ante review. A coalition of civil rights organizations have advocated for such practices in a document laying out a set of civil rights principles for hiring assessment technologies.<sup>40</sup> We discuss emerging ideas such as Algorithmic Impact Assessments in the final section of this chapter.

<sup>18</sup>This can be mitigated in some cases if law firms are willing to be paid on a contingency fee basis—wherein they are paid only in the event of a favorable result as a fixed percentage of the damages recovered—or if the statute provides for collective or class actions.

While there is no need to establish intent, there is a new set of requirements: identifying a specific policy or practice that caused the adverse employment decision; compiling the requisite statistics to show that the policy has a disparate impact; and rebutting the employer's defense that the policy is justified by job-relatedness. The third prong is a particularly severe hurdle for plaintiffs as they are structurally poorly positioned to identify an alternative employment practice, lacking the knowledge of the internals of the business that the employer does.<sup>43,35</sup>

The net result of these barriers to plaintiffs is that their odds of success at trial are exceedingly low. Katie Eyer summarizes data from the Uncertain Justice project: "of every 100 discrimination plaintiffs who litigate their claims to conclusion (i.e., do not settle or voluntarily dismiss their claims), only 4 achieve any form (de minimis or not) of relief. . . . These odds can properly be characterized as shockingly bad, and extend (with minor differences) to every category of discrimination plaintiff, including race, sex, age, and disability."<sup>31</sup>

We should note that there is a widespread view that employment discrimination lawsuits are too *easy* to file and too favorable to plaintiffs, a position we reject. Selmi critically examines this perception and notes that it is prevalent among judges; correcting this perceived imbalance may in fact be one reason for the creation of numerous hurdles for plaintiffs.<sup>44</sup> Whether or not one subscribes to the view that many "nuisance lawsuits" are filed by plaintiffs alleging discrimination, it is true that courts are highly strained and judges are wary of decisions that might open the "floodgates" to lawsuits. This suggests that the burdens we have discussed above are unlikely to go away.<sup>45,42</sup>

### *The difficulty of substantive and structural reform through procedural intervention*

Even if compliance with antidiscrimination law is high, and legal remedies are readily attained, there may be even more fundamental limits to the effectiveness of the law. To what extent do the formal limits that the law imposes on individuals and organizations lead to a just society? How big is the gap between legal and moral notions of unfairness?

Stephen Halpern frames the issue thus:<sup>46</sup>

In translating a social problem into the "language" of the law, lawyers must frame their analysis in terms of contrived concepts, issues, questions, and remedies that the legal system recognizes and deems legitimate. In that translation, as in any translation, there are constrictions and distortions. Framing a social problem as a legal issue produces a transformation of the issue itself—a reconceptualization of the problem, yielding unique questions and concerns that first become the focus of the legal debate and subsequently tend to dominate public discussion. When racial problems are reformulated as questions of legal rights, the resulting dialogue does not capture the complexity and subtlety of those problems or permit consideration of the fullest range of remedies for them. Inevitably, the demands and limits of the legal process alter the public discourse about and understanding of vital racial issues.

Halpern's book is about racial inequality in education; his main example of his thesis is the effort that was put into ending segregation of public schools without much attention paid to the quality of education received by Black students in integrated schools. Similarly, Title VI litigation focused on procedures for processing complaints of discrimination filed with the federal government, rather than mechanisms to vindicate substantive rights to an education of comparable quality. He argues that "[f]ew, if any, of the factors that have an impact on educational achievement are governed by 'legal rights' or are readily translatable into an issue of 'racial discrimination'." He gives two reasons why inequalities persist despite the law's formal remedies: the de-facto segregation of American cities and the fact that academic differences often arise from instability in the home and other social, economic, and health disparities.

While the effect of school desegregation in the United States is a vast topic, the broader point is that limitations of the legal process restrict what is achievable—and even shape our understanding of the issues themselves. Another example of this comes from Richard Rothstein's book *Color of Law*:<sup>4</sup>

Although most African Americans have suffered under [historically racist government housing policies], they cannot identify, with the specificity a court case requires, the particular point at which they were victimized. For example, many African American World War II veterans did not apply for government guaranteed mortgages for suburban purchases because they knew that the Veterans Administration would reject them on account of their race, so applications were pointless. Those veterans then did not gain wealth from home equity appreciation as did white veterans, and their descendants could not then inherit that wealth as did white veterans' descendants. With less inherited wealth, African Americans today are generally less able than their white peers to afford to attend good colleges. If one of those African American descendants now learned that the reason his or her grandparents were forced to rent apartments in overcrowded urban areas was that the federal government unconstitutionally and unlawfully prohibited banks from lending to African Americans, the grandchild would not have the standing to file a lawsuit; nor would he or she be able to name a particular party from whom damages could be recovered.

Another impetus toward proceduralism comes from the interaction of the court system with the internal procedures of organizations. Under the theory of legal endogeneity, organizations enact procedural protections, such as diversity training programs, with the putative aim of curbing discrimination; over time, courts gradually come to mistake these procedural and symbolic compliance-oriented activities for substantive measures; but once these symbolic measures themselves attain legal significance, substantive concerns have been pushed outside the scope of legitimate debate.<sup>47</sup>

Further, even substantive change at individual organizations may not imply structural change—that is, a change to the underlying factors in society that

produce disparities in the first place. Even if an employer achieved statistical parity in hiring and promotion rates, the application rates might themselves reflect unequal opportunity in society and/or discrimination at previous levels or stages of the system, and there is little the law can do to compel individual decision makers to remedy these inequalities.

Legal interventions whose effects are both substantive and structural are rare. One notable example is the impact of Title IX on women's athletics. The law has been interpreted to not only prohibit discrimination in a narrow sense but also require equity in a number of areas such as scholarships, coaching, and facilities. Arguably as a result of these interventions, women's athletics in the United States has gradually risen in prestige, weakening the gender hierarchy in athletics, leading to greater parity in athletics even outside the collegiate context.<sup>19</sup> In general, however, these types of substantive interventions have so far proved less feasible than formal ones in part because of the funding they require.

Although we have contrasted procedural interventions with substantive and structural interventions above, the line between them can be murky, and the former can at least function as a toehold to the latter. To the extent that inequality persists because of entrenched policies that maintain an unequal distribution of resources, procedural interventions that allow members of historically oppressed groups to rise to positions of authority might allow them to more effectively alter these policies. Procedural interventions can help reduce the capacity of already advantaged groups from usurping full control over the policy-making process. Still, this is far from an ideal route to change, as it places the burden of advancing the interests of specific groups on individuals who belong to those groups.

Another seeming contrast is between discrimination law and redistributive policies, i.e., the government directly taxing certain actors and reallocating those funds to the disadvantaged group. But discrimination law can be understood to be a mechanism that places the economic burden of rectifying past injustice to some extent on employers, lenders, etc. In some ways, this might be similar to a policy of taxing employers and using those funds to support groups that have been subject to discrimination in the past.

Affirmative action policies, in particular, occupy a space that is squarely in between formal nondiscrimination and redistributive policies. An example of such a policy would be a job training program offered by an employer that favored groups with lower access to opportunities.<sup>48</sup> However, except in rare cases, the law does not *compel* affirmative action by private entities but merely *allows* it. More commonly seen are affirmative requirements for governments. The Fair Housing Act, in addition to nondiscrimination mandates, requires HUD and recipients of federal funds from HUD to "affirmatively further" the policies and purposes of the act. This might enable, for instance, subsidized housing in high-income communities that opens up access to higher-quality schools and amenities. However, this part of the FHA has largely lain dormant. Thus, at least some of the limitation of the law in creating meaningful change can be attributed to the lack of

---

<sup>19</sup>This is not to say that fairness has been achieved in women's athletics or athletic programs in general: horrific sexual abuse scandals remind us that there is a long way to go.

political will to fully act upon existing laws, rather than an inherent limitation of the legal system.

## *Regulating machine learning*

Although U.S. antidiscrimination law predates the widespread use of machine learning, it is just as applicable if a decision maker uses machine learning or other statistical techniques. That said, machine learning introduces many complications to the application of these laws. These complications are being vigorously debated in the legal scholarship, and many scholars are concerned that existing law may be inadequate to address the types of discrimination that arise when machine learning is involved. At the same time, there is also an opportunity to exercise new regulatory tools to rein in algorithmic discrimination. There is little case law on this topic, so our discussion of these issues will be based on legal scholarship. As before, our discussion is U.S.-centric but we touch upon the EU's General Data Protection Regulation (GDPR) in a few places.

### *Disparate treatment*

Recall that the two main anti-discrimination doctrines are disparate treatment and disparate impact. Disparate treatment is principally concerned with the explicit intent to discriminate on the basis of legally protected characteristics; in contrast, disparate impact focuses on decision-making where there is no explicit intent to discriminate, but where even decisions made on the basis of seemingly benign characteristics nevertheless results in unjustified disparities along characteristics that are legally protected.

Most reports of discrimination in machine learning have been cases of unintentional rather than intentional discrimination. Besides, developers of machine learning systems who intend to discriminate are unlikely to rely explicitly on protected attributes due to the easy availability of proxies. When this happens, it can be hard to prove that there was an intent to mask discrimination. For these reasons, disparate treatment is rarely invoked and disparate impact is seen as much more relevant. We will return to disparate impact shortly. But one important question involving disparate treatment relates to systems that explicitly rely on the protected attribute to correct data biases or mitigate the effects of past discrimination. Does this constitute disparate treatment? In other words, does law impose limits on algorithmic fairness interventions?

The answer is nuanced. One relatively bright line in the law is that selection quotas are unconstitutional. In machine learning terms, this roughly maps to the difference between techniques that aim to enforce parity and those that merely penalize disparity during the optimization step. The latter type of technique is analogous to a process that is race conscious and values diversity but still allows the final distribution to vary depending on the set of candidates. It is helpful, as always, to remember that technical distinctions rarely map cleanly to legal determinations.

There is also a major difference between an individual decision made on the basis of a protected attribute and an overall policy that takes the interests of protected groups into account. Disparate treatment primarily applies to the former type of decision. This is similar to the distinction between the use of a protected attribute at training versus test time, although, again, this distinction by itself is far from legally determinative.

A Supreme Court case that is often cited as an example of the tension between disparate treatment and disparate impact (and the disparate-treatment pitfalls of race-conscious decision making) is *Ricci v. DeStefano*. The case arose because the New Haven fire department scrapped a promotional exam after finding that Black firefighters had a lower passage rate than White firefighters. The department worried that it would open itself to disparate impact liability. But it was then sued by the White and Hispanic firefighters who would have qualified for promotion based on the exam. The court agreed with the plaintiffs that the department had engaged in disparate treatment against them.

Pauline Kim notes a crucial distinguishing feature of the *Ricci* case: the plaintiffs had already invested considerable time and expense in studying for the exam, and thus the department's actions resulted in concrete harm to specific individuals. In Kim's view, the court's logic wouldn't apply when an employer *prospectively* makes a change to its hiring practices in order to avoid the potential for disparate impact.<sup>49</sup>

Finally, even if a practice constitutes *prima-facie* disparate treatment, it may be legal if it is part of a valid affirmative action program, i.e., one that aims to remedy past discrimination. In employment, the Supreme Court has ruled that race- or gender-based affirmative action programs are valid if they seek to eliminate "manifest imbalances" in "traditionally segregated job categories" and do not "unnecessarily trammel" the interests of other candidates. Some scholars have argued that this should hold for voluntary algorithmic affirmative action as well.<sup>50</sup>

### *Disparate impact*

To understand how disparate impact applies to statistical decision making, we must unpack the legal doctrine. The burden-shifting framework established by the Supreme Court for Title VII employment discrimination claims works as follows.<sup>51</sup> First, the plaintiff must establish a *prima-facie* case by showing a sufficient difference in selection rates between different groups. What constitutes a sufficient difference is unclear. The EEOC has established a threshold of four-fifths (i.e. a difference of 20%) as a guideline, but this is not a strict rule. In a big-data world, some commentators have argued that the criterion should be based on statistical significance of the difference rather than the magnitude.<sup>20</sup>

If the plaintiff is successful at showing a sufficient difference, burden shifts to the defendant, who must then establish that the challenged practice is "job related"

---

<sup>20</sup>In fact, statistical significance has always been a part of the EEOC criteria alongside substantive significance.

and consistent with “business necessity”. If the defendant can show this, then the plaintiff can still win by showing that there is an “alternative employment practice” that would have achieved the employer’s aims while being less discriminatory.

The critical step from the perspective of statistical decision making is the question of business necessity. One way the employer can show this is through “empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance.” Since machine learning is a technique for establishing predictive validity, commentators such as Barocas and Selbst suggest that this represents an exceedingly low bar for employers.<sup>52</sup> As long as the target variable used in a predictive model is putatively job-related, the requirement is satisfied.

On the other hand, Pauline Kim argues that Title VII can in fact effectively address discriminatory effects of machine learning, based on a close reading of the statute.<sup>53</sup> However, the doctrine that has developed since its passage is a poor fit for addressing discriminatory machine learning. For example, the requirement for the plaintiff to identify a specific employment practice that caused the disparity developed in an era when written tests were the primary vehicle for disparate impact. But when a statistical model is at play, especially an uninterpretable one that uses a large number of features, it is not clear what the plaintiff is supposed to identify. Thus, the doctrine will have to evolve if Title VII is to address discriminatory machine learning.

Another issue that’s specific to automated decision making arises from the fact that the software is usually not developed in-house by the decision maker but rather by specialized external firms. For example, companies such as Hirevue and Pymetrics offer tools to automate part of the hiring process and Upstart provides a predictive model for loan underwriting. In such cases, who should bear liability? In employment law, employers, not vendors, bear legal liability.<sup>54</sup> But employers (and other clients of these tools) resist this since they usually lack the expertise to conduct statistical validation. Shifting some or all of the liability from clients to vendors would have pros and cons from an anti-discrimination perspective. It might mean that vendors become much more careful at testing their offerings. On the other hand, even if a tool has been broadly tested for disparate impact, it may perform differently in the context of a particular employer’s applicant pool. Further, plaintiffs may face even more difficulty in showing an alternative employment practice.

While disparate treatment and disparate impact are the two main prongs of anti-discrimination law, when it comes to data-driven decisions the anti-discrimination toolbox is wider and includes privacy law, explanation, and potentially consumer protection law. We discuss these in turn.

### *Privacy*

When we worry about privacy, the underlying concern is often that data about us could be used to discriminate or might result in adverse treatment. For example, if a job interviewer inquired about religion, it may be considered a privacy violation.



The harm that animates this worry is the denial of a job. As another example, reports that the retailer Target uses shopping records to identify pregnant customers sparked outrage.<sup>55</sup> The potential for harm arises because pregnancy is a time when individuals are particularly susceptible to manipulation through marketing (which is the reason that marketers are interested in pregnancy in the first place).

Yet, data privacy law and anti-discrimination law have largely been separate in the United States. Returning to the above example, it is not privacy law that forbids interviewers from asking about religion. Rather, since employment law forbids discriminating on the basis of religion, interpretive guidance from the EEOC and, often, from institutions themselves discourages such questions during interviews.<sup>21</sup>

Still, given the normative alignment, it is natural to wonder whether privacy law can be adapted to serve antidiscrimination ends. There is a lot of intuitive appeal to this idea, especially when it comes to machine learning. If a decision making system relies on data, why not put restrictions on the flow of data to prevent unjustified discrimination?

But when we examine this argument in more detail, difficulties emerge. The most obvious is the issue of proxies. As we discussed in Chapter 3, prohibiting access to sensitive attributes such as race or gender typically has a negligible impact on a classifier when rich datasets are available. It isn't just that the decision maker may train a model to predict the sensitive attribute from innocuous attributes, as in the Target example above. It may instead directly use the innocuous attributes to predict the outcome of interest, such as the susceptibility of a particular person to a particular marketing message. This is in fact exactly what has been shown to happen on Facebook-scale advertising platforms.<sup>57</sup>

If proxies are the problem, another approach is to prohibit the collection of proxies. This is the idea behind "ban the box" laws in U.S. states that prohibit employers from inquiring about criminal history. Ban-the-box has two motives. One is to make it easier for formerly incarcerated people to be rehabilitated into society through employment. In this view, criminal history itself can be seen as the sensitive attribute. The other motive is to combat the racially disparate impact of discrimination against formerly incarcerated people. Here, criminal history can be seen as a proxy for race. It is this view that is of interest to us.

An influential study by Agan and Starr found that employers increased racial discrimination when they were subject to ban-the-box laws.<sup>58</sup> What does this mean for the prospect of preventing discrimination by prohibiting information flows? One view is that in the light of this finding, ban-the-box laws clearly harm more than they help. But another perspective is that racial discrimination is already unlawful, so what the study really reveals is the need to step up auditing and enforcement. If that were to happen, ban-the-box laws might be able to achieve their intended effects.

---

<sup>21</sup> An early attempt to mingle privacy and accountability (but not antidiscrimination) goals is seen in the Fair Information Practice Principles. The FIPPS contain the seeds of comprehensive data protection laws enacted around the world. In the U.S., they do not have the force of law, except in some sectoral laws such as the Fair Credit Reporting Act. A watered-down version of FIPPS focusing on "notice and choice" governs U.S. online commerce.<sup>56</sup>

Going beyond protected attributes and their proxies, privacy law may make it harder to amass dossiers on individuals (for example, containing shopping or browsing records), and we might hope that this would make discrimination harder. While a full discussion of this is beyond our scope, U.S. privacy law is often criticized for failing to accomplish this effectively, for several reasons. There is no general federal privacy law analogous to the GDPR in the E.U. Only a few sectoral privacy laws exist, such as the Health Insurance Portability and Accountability Act (HIPAA). Privacy in most commercial transactions or interactions comes down to “notice and choice”, which is ineffective for many reasons including the power asymmetry and information asymmetry between firms and individuals. In the machine learning context, the notice and choice approach to privacy is particularly ineffective as a barrier to firms building models that may infer sensitive attributes or make adverse decisions based on innocuous attributes. That’s because of the “tyranny of the minority”: it takes only a small number of individuals to consent to collection to be able to uncover the statistical patterns that make such inferences possible.<sup>59</sup>

While privacy laws have not so far helped to address discrimination, discrimination law has sometimes helped to preserve privacy. The Genetic Information Nondiscrimination Act is an anti-discrimination law that mutated into a privacy law through expansive court decisions and EEOC interpretation.<sup>60</sup> Genetic information is an exception to the ubiquity of proxies, as it cannot readily be inferred with any degree of completeness or accuracy from observable characteristics.

Broader senses of the word privacy go beyond information flow and encompass transparency, explanation, and redress. We turn to those next.

### *Explanation*

In the context of automated decision making, explanation could have one of two goals. The first is an explanation of the overall system. In a rule-based system this might be the set of decision rules. In a machine learning system it’s less obvious what form this explanation should take, and it is a subject of active research in the field of interpretable machine learning.

An explanation of the overall system promotes fairness objectives because it allows regulators, users, and developers to check whether the system adheres to normative requirements. In many cases, explanation allows us to immediately spot potential unfairness. For example, if we know that a system used for detecting fake accounts on social media relies on an uncommon name as a signal of inauthenticity, it is easy to see why it may be more likely to incorrectly flag users who are from minority cultures, as we discussed in Chapter 1.

The second goal of explanation of how a particular decision was made given the characteristics of the decision subject. This goal can also promote fairness objectives. It satisfies a powerful innate need to understand how consequential decisions about us are made. The dread that arises when a decision system denies us such explanation is visceral enough that it has a name: Kafkaesque. Explanation of individual decisions also serves more instrumental purposes. It allows us to

contest decisions that may have been made on the basis of erroneous information. Even if the decision was accurate, explanation allows recourse, that is, actions that decision subjects might take to alter the decision in the future. For example, a loan applicant who was denied because of a low credit score may make attempts to improve that score.

Taking a step back, decision systems can be analyzed at three levels. The highest level is that of values, goals, and normative constraints (for example, maximizing predictive accuracy while ensuring fairness). The second is the design of the system and its rules. The third is the level of individual decisions. Justification is needed at all three levels. In traditional decision making systems, values and goals derive legitimacy through stakeholder participation, deliberation, and democratic debate. It is often merged with the next step, rulemaking or policymaking, which is the process of going from the first level to the second—designing a decision system based on values and goals. If the first step was skipped, tensions between different values or between different stakeholders’ objectives become apparent in this process. In administrative bureaucracies, they are resolved through a process of public participation.<sup>61</sup> In contrast, the process of *adjudication* bridges the second and the third levels.<sup>62</sup> For example, in the United States, bureaucrats periodically assess the value of homes and other real estate based on an elaborate policy in order to determine how much property tax should be levied. If the owner disagrees with the assessment, they can appeal, and have a right to a hearing.

Automated systems erode the procedural protections involved in rulemaking and adjudication: public participation and appeals respectively.<sup>63</sup> These issues are exacerbated when machine learning is involved, due to its inscrutability and non-intuitiveness.<sup>64</sup> The two goals of explanation might help mitigate these concerns: by allowing us to understand how the overall system and policy conform to normative constraints and how individual decisions conform to the policy. These roughly correspond to the distinction between “global” and “local” interpretability in the technical literature.

Requirements for both flavors of explanation can be seen in existing laws. The FCRA and ECOA contain an “adverse action notice” requirement. This is an example of the second goal, as it pertains only to the individual decision and does not require transparency about the overall model. In contrast, the GDPR requires “meaningful information about the logic involved” if an individual is subject to a consequential decision by an automated system. This is generally understood as requiring some degree of explanation of both the overall model and the specific decision.

Table 2: Comparison of the two flavors of model explanation

	Explanation of overall system	Explanation of specific decisions
Goal	Justify policy based on goals and values	Justify decision based on policy

	Explanation of overall system	Explanation of specific decisions
Bureaucratic analog	Rulemaking or policymaking	Adjudication
Technical tool	Global interpretability	Local explanation
Example legal requirement	GDPR: “meaningful information about the logic involved”	FCRA and ECOA: adverse action notice

Selbst and Barocas describe several limitations to the usefulness of explanations. We highlight two main ones. The first is the difficulty of producing explanations that are simultaneously faithful to the model and understandable to a nonexpert. If a credit model combines dozens of variables in nonlinear ways, a reason such as “length of employment” or “insufficient income” might fall far short of fully explaining a decision; yet this is all that is required of adverse action notices. Conversely, an explanation of a decision that is fully faithful to a statistical model may be incomprehensible to most decision subjects.

There is an important distinction between explanations given willingly and those demanded by law of a decision maker who has no other incentive to provide them. So far, it has proven challenging for regulators to set legal requirements for what constitutes a good explanation and assess whether they are working as intended. Empirical evidence supports the difficulty of compelling unwilling decision makers to provide meaningful explanations. For example, a 2018 study found that Facebook’s “Why am I seeing this?” ad explanations are vague, incomplete, misleading, and generally useless.<sup>65</sup> The research literature shows that if Facebook *wanted* to provide good explanations, it is possible to do far better.

A more fundamental limitation described by Selbst and Barocas is that even explanations that are faithful and understandable may not enable normative assessment. If an employer uses a screening model that computes a score based on some keywords (a faithful and understandable explanation), it is normatively important to know whether those keywords represent job-related skills, or act as proxies (for example, hobbies) that signal social class, or something else. We may be able to make such an assessment given the keywords, but it is not straightforward. Modern methods that provide explanations based on high-level concepts rather than low-level features hold promise in this regard, but the gap between explanations and a full normative justification is likely to remain.

Because of these limitations, there has been a gradual turn from explanations to *algorithmic impact assessments* (AIAs). A full discussion of AIAs is beyond our scope, but we point out how AIAs, at least in an idealized version, differ from explanations. First, AIAs go beyond explaining the model itself and focus on how it was created, how it will be used, and what impacts it is likely to have. Second, the primary consumers of AIAs are not decision subjects but rather regulators and other experts, which alleviates the faithfulness-comprehensibility tradeoff to some degree. Third, AIAs must be performed before the model is deployed and must be

updated periodically. Some visions of AIAs call for the involvement of impartial external parties in producing them.

The GDPR incorporates one version of AIAs, namely Data Protection Impact Assessments (DPIAs). DPIAs must include a description of the algorithm, and the purpose of the processing, an assessment of the necessity of processing in relation to the purpose, an assessment of the risks to individual rights and freedoms, and the measures a company will use to address these risks. It requires consultation “where appropriate” with impacted individuals. But DPIAs are not required to be released to the public. It is too early to tell how effective they will be in practice; much will rely on the behavior of regulators.<sup>66,67</sup>

Algorithmic Impact Assessments are closely related to audits and the terms are sometimes used interchangeably. Nonetheless, there are several types that are worth distinguishing. A 2020 report classifies them into four categories:<sup>68</sup>

- *Bias audits* conducted by researchers, journalists, or civil society organizations (inspired by social science audits, as we saw in the chapter on testing discrimination in practice).
- *Regulatory audits* conducted by regulators with statutory powers to examine internal data and systems, modeled on financial audits.
- *Algorithmic risk assessments* conducted by the developer or procurer of a tool, modeled on environmental impact assessments, to assess possible risks and mitigation strategies before deploying a system.
- *Algorithmic impact evaluations*, which are retrospective and modeled on policy evaluations, conducted typically by public sector agencies w.r.t. algorithms which implement a policy.

Algorithmic impact assessments and audits are a burgeoning area,<sup>69</sup> and their potential is still being explored. For example, Ifeoma Ajunwa ambitiously argues for reading in existing employment law a duty of care that would obligate employers to conduct audits of automated hiring systems.<sup>70</sup> Malgieri and Pasquale propose an ex-ante model of regulation where developers of consequential AI systems must perform a risk assessment before deployment and, in some cases, be required to get it approved by an authority[i@malgieri2022transparency]. These developments illustrate our point that the turn to machine learning, while indeed creating challenges for antidiscrimination law, also creates opportunities alongside it. The software tool and data records involved in automated systems provide a leverage point for regulators.

### *Consumer protection*

Consumer protection law has completely different roots from either antidiscrimination law or privacy law. Consumer movements first gained ground in the United States in the early 20th century, initially due to food safety issues.<sup>71</sup> The Federal Trade Commission was established in 1914. Although it initially focused on antitrust, consumer protection gradually became an equally important prong of its activities. It has been the primary agency responsible for consumer protection,

and has the statutory authority to challenge “unfair or deceptive” practices in commerce. It is this authority that the agency uses to carry out the activities that it is well known for, such as policing false advertising and fraud, especially identity theft.<sup>72</sup> Many states have consumer protection laws with similar import, enforced by attorneys general.

Credit regulation is one area of consumer protection law that also serves fairness purposes, understood in a broad sense. The Fair Credit Reporting Act of 1974 narrows the permissible uses of credit reports so that they are not used for arbitrary purposes. It gives consumers ways to contest inaccuracies in the data, considering that they are used to make consequential decisions. And it requires notifying the consumer when adverse action is taken against them. FCRA does not address discrimination in the sense of disparate treatment or disparate impact; that would come later, in the Equal Credit Opportunity Act of 1976. In other areas such as employment law, consumer protection does not currently play a role, although scholars have speculatively advocated for treating job candidates as consumers.<sup>43</sup>

Outside the traditional sectors of antidiscrimination law, there is a vast swath of everyday digital products in which machine learning biases manifest, and this is where consumer protection law is potentially highly relevant. For example, if a face unlock feature on a smartphone is substantially less accurate for some groups of users, this is not a violation of any of the sector-specific statutes we’ve discussed so far, but it may fall under FTC authority. Even in domains such as employment discrimination, there are peculiar gaps such as the fact that vendors of algorithmic screening tools are not covered entities, and consumer protection law can potentially help fill this gap.

As of this writing, this is all speculative. So far, the FTC hasn’t gone after discriminatory practices except when the company also violates an anti-discrimination statute such as ECOA, which the FTC has authority to enforce.<sup>73</sup> The term “unfairness” in the FTC act has traditionally meant something quite different: taking unjustified advantage of consumers that they cannot avoid. The prototypical example of an unfair commercial practice would be selling snake oil, and a more modern example would be lax data security leading to a data breach. But note that unlike the anti-discrimination statutes, the FTC has substantially more power to determine what constitutes deceptive and unfair. It is quite possible that the agency will take a broad view of unfairness and that the courts will permit it. The statute allows the FTC to look to “established public policies” in determining what is unfair. It has been suggested that the FTC can thus look to antidiscrimination statutes and rules as scaffolding to build a framework for making determinations about algorithmic discrimination.<sup>74</sup>

The FTC’s deception authority offers a clearer but more circumscribed option. Companies often make affirmative claims about their products being unbiased. If those claims turn out to be false, that’s deception. The same goes for false claims about products being *effective*. This is relevant since many predictive decision making tools on the market lack evidence of predictive validity, which means that they may subject people to arbitrary decisions. Yet, unless those arbitrary decisions are also systematically biased, they are difficult to challenge

under antidiscrimination law. Finally, a lack of transparency may also constitute a deceptive practice. Indeed, the FTC took action against a company that trained a face recognition model on its users' photos while falsely telling them that the feature was opt-in.<sup>75</sup>

Historically, the FTC has had a roller-coaster ride in terms of how broadly it treats its authority and how much it flexes its muscles. After being ineffective in the '60s and reinvigorated in the '70s,<sup>76</sup> Congress rebuked it in the early '80s and limited its authority due to lobbying by powerful business interests.<sup>72</sup> It has remained cautious since, and was further caught off guard in the technology era due to limitations of in-house technical expertise. This led to withering criticism for failures such as allowing Cambridge Analytica's exfiltration of Facebook users' data, despite the FTC long being aware of similar previous events and supposedly closely monitoring Facebook under a "consent decree". In the 2020s, the agency has shown some signs of being invigorated. Specifically on algorithmic discrimination, it published a blog post containing surprisingly strong language.<sup>77</sup> A whitepaper co-authored by a sitting commissioner also lays out an ambitious agenda.<sup>78</sup> All this is to say: the relevance of consumer protection law to algorithmic discrimination remains a wild card.

Beyond the traditional conception of consumer protection, there are emerging ideas such as a duty of loyalty for companies who are entrusted with customers' data.<sup>79</sup> Such companies would be obligated to act in the best interests of people exposing their data and online experiences. The duty of loyalty is a common obligation in fiduciary relationships (for example, a lawyer owes such a duty to her client). But its application to the holders of personal data is a relatively new idea. Although it has been proposed mainly with the aim of improving privacy and minimizing manipulative practices such as "dark patterns", it would have some implications for nondiscrimination as well.

## *Concluding thoughts*

We've covered a lot of ground in this chapter. We reviewed how the various civil rights movements together gave rise to a relatively robust body of anti-discrimination law in the U.S. Generally, this law aims to strike a balance between preventing (and remedying) discrimination on the one hand, and avoiding excessive burdens on decision makers on the other hand. It has been refined, contested, and implemented over decades by the push and pull of court decisions, regulatory agencies, institutional bureaucrats, continued civil rights activism, and shifts in public opinion. It has important limitations: practically, private plaintiffs find it difficult to find legal recourse; more fundamentally, the law itself is far from an ideal route to bring about structural changes.

Turning to the novel challenges raised by automated decision making, there is a risk that discriminatory machine learning might slip through the gaps in how the law conceives of discrimination. In our view, this risk is counterbalanced by the expanded legal toolkit available: privacy law, requirements regarding explanation

and impact assessment, and consumer protection law. So far, this potential has lain mostly dormant for various reasons: a narrow conception of privacy, a lack of broad legislation in the U.S. requiring explanation of consequential decisions, and the timidity of consumer protection agencies. This could yet change; it is possible that the law and enforcement agencies could be reformed to effectively address the new problems. At a minimum, even if not enshrined into law, the tools for intervention that we've discussed offer a blueprint for public interest advocates seeking to hold companies accountable.



## Bibliography

- <sup>1</sup> Plessy v. ferguson, 1896.
- <sup>2</sup> Isabel Wilkerson. *The Warmth of Other Suns: The Epic Story of America's Great Migration*. Penguin Random House, New York, NY, 10 2011.
- <sup>3</sup> Luke Keele, William Cubbison, and Ismail White. Suppressing black votes: A historical case study of voting restrictions in louisiana. *American Political Science Review*, 115(2):694–700, 2021.
- <sup>4</sup> Richard Rothstein. *The Color of Law: A Forgotten History of How Our Government Segregated America*. W. W. Norton & Company, New York, NY, 5 2018.
- <sup>5</sup> Juliet E. K. Walker. *History of Black Business in America: Capitalism, Race, Entrepreneurship*. Evolution of Modern Business. Twayne Publishers, New York, NY, 1998.
- <sup>6</sup> Francis J Vaas. Title vii: Legislative history. *BC Indus. & Com. L. Rev.*, 7:431, 1965.
- <sup>7</sup> Betty Friedan. *The Feminine Mystique*. W. W. Norton & Company, New York, NY, 9 2001.
- <sup>8</sup> Roe v. wade, 1973.
- <sup>9</sup> Senate report 589 of the 94th congress, 1976.
- <sup>10</sup> William N. Eskridge. *Dishonorable Passions: Sodomy Laws in America, 1861-2003*. Viking, New York, NY, 5 2008.
- <sup>11</sup> Lawrence v. texas, 2003.
- <sup>12</sup> Obergefell v. hodes, 2015.
- <sup>13</sup> Bostock v. clayton county, georgia, 2020.
- <sup>14</sup> Catherine A. Okoro, NaTasha D. Hollis, Alissa C. Cyrus, and Shannon Griffin-Blake. Prevalence of Disabilities and Health Care Access by Disability Status and Type Among Adults — United States, 2016. *Morbidity and Mortality Weekly Report*, 67(32):882–887, 2018.
- <sup>15</sup> Carol Padden and Tom Humphries. *Inside Deaf Culture*. Harvard University Press, Cambridge, MA, 10 2006.

- <sup>16</sup> Doris Zames Fleischer and Frieda Zames. *The disability rights Movement: From charity to confrontation*. JSTOR, 2013.
- <sup>17</sup> Facebook, inc., 2019.
- <sup>18</sup> Alex Rosenblat, Karen EC Levy, Solon Barocas, and Tim Hwang. Discriminating tastes: Customer ratings as vehicles for bias. *Data & Society*, pages 1–21, 2016.
- <sup>19</sup> Lauren Sudeall Lucas. Identity as a Proxy. *Columbia Law Review*, 115(6):1605–1674, 2015.
- <sup>20</sup> Heart of atlanta motel, inc. v. united states, 1964.
- <sup>21</sup> Kenneth Glenn Dau-Schmidt and Ryland Sherman. The Employment and Economic Advancement of African-Americans in the Twentieth Century. *Jindal Journal of Public Policy*, 1(2):95–116, 2013.
- <sup>22</sup> Christopher J. Walker. Attacking Auer and Chevron Deference: A Literature Review. *The Georgetown Journal of Law & Public Policy*, 16(1):103–122, 2018.
- <sup>23</sup> Executive order no. 8802, 1941.
- <sup>24</sup> Questions and answers about the EEOC’s enforcement guidance on pregnancy discrimination and related issues, 06 2015.
- <sup>25</sup> Iram Valentin. Title IX: A brief history. Technical report, Equity Resource Center, Newton, MA, 1997.
- <sup>26</sup> Hugh Davis Graham. The storm over Grove City College: Civil rights regulation, higher education, and the Reagan Administration. *History of Education Quarterly*, 38(4):407–429, 1998.
- <sup>27</sup> R. Shep Melnick. Analyzing the Department of Education’s final Title IX rules on sexual misconduct. Technical report, The Brookings Institution, Washington, DC, 6 2020.
- <sup>28</sup> Department of Education. Enforcement of title ix of the education amendments of 1972 with respect to discrimination based on sexual orientation and gender identity in light of bostock v. clayton county, 2021.
- <sup>29</sup> Martin J Katz. Unifying disparate treatment (really). *Hastings Law Journal*, 59(3):643, 2008.
- <sup>30</sup> Robert N. Price. Griggs v. Duke Power Co.: The first landmark under Title VII of the Civil Rights Act of 1964. *Southwestern Law Journal*, 25(3):484–493, 2016.
- <sup>31</sup> Katie Eyer. The but-for theory of anti-discrimination law. *Va. L. Rev.*, 107:1621, 2021.
- <sup>32</sup> Title VI Legal Manual. Technical report, Civil Rights Division U.S. Department of Justice, Washington, DC, 4 2021.

- <sup>33</sup> House of Representatives Report No. 914 Pt 2, 88th Congress, 1st Session, 1963.
- <sup>34</sup> Texas department of housing and community affairs v. inclusive communities project, inc., 2015.
- <sup>35</sup> Nicole Buonocore Porter. A new look at the ada's undue hardship defense. *Mo. L. Rev.*, 84:121, 2019.
- <sup>36</sup> Louis Menand. The changing meaning of affirmative action. *The New Yorker*, 1 2020.
- <sup>37</sup> Samuel R. Bagenstos. Formalism and employer liability under Title VII. *University of Chicago Legal Forum*, 2014(1):145–176, 2014.
- <sup>38</sup> John W. Cairns. Credit equality comes to women: An analysis of the Equal Credit Opportunity Act. *San Diego Law Review*, 13(4):960–977, 1976.
- <sup>39</sup> Kenneth Y. Chay. The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Employment Opportunity Act of 1972. *Industrial & Labor Relations Review*, 51(4):608–632, 1998.
- <sup>40</sup> Civil rights principles for hiring assessment technologies, 2020.
- <sup>41</sup> Scott A. Moss. Fighting discrimination while fighting litigation: A tale of two supreme courts. *Fordham Law Review*, 76(2):981–1013, 2007.
- <sup>42</sup> Sandra F. Sperino and Suja A. Thomas. *Unequal: How America's Courts Undermine Discrimination Law*. Oxford University Press, New York, NY, 2017.
- <sup>43</sup> Ifeoma Ajunwa. The paradox of automation as anti-bias intervention. *Cardozo Law Review*, 41(5):1671–1742, 2020.
- <sup>44</sup> Michael Selmi. Why are Employment Discrimination Cases So Hard to Win? *Louisiana Law Review*, 61(3):555–575, 2001.
- <sup>45</sup> Katie R. Eyer. That's not discrimination: American beliefs and the limits of anti-discrimination law. *Minnesota Law Review*, 96:1275–1362, 2012.
- <sup>46</sup> Stephen C. Halpern. *On the Limits of the Law: The Ironic Legacy of Title VI of the 1964 Civil Rights Act*. Johns Hopkins University Press, Baltimore, MD, 1995.
- <sup>47</sup> Lauren B Edelman, Linda H Krieger, Scott R Eliason, Catherine R Albiston, and Virginia Mellema. When Organizations Rule: Judicial Deference to Institutionalized Employment Structures. *American Journal of Sociology*, 117(3):888–954, 2011.
- <sup>48</sup> *Steelworkers v. weber*, 1979.
- <sup>49</sup> Pauline Kim. Race-aware algorithms: Fairness, nondiscrimination and affirmative action. *California Law Review*, 2022.

- <sup>50</sup> Jason R. Bent. Is algorithmic affirmative action legal? *The Georgetown Law Journal*, 108(4):803–853, 2020.
- <sup>51</sup> *Griggs v. duke power co.*, 1971.
- <sup>52</sup> Solon Barocas and Andrew D. Selbst. Big data’s disparate impact. *California Law Review*, 104, 2016.
- <sup>53</sup> Pauline T. Kim. Data-driven discrimination at work. *William & Mary Law Review*, 58(3):857–936, 2017.
- <sup>54</sup> Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Conference on Fairness, Accountability, and Transparency*, pages 469–481, 2020.
- <sup>55</sup> Charles Duhigg. *The Power of Habit: Why We Do What We Do in Life and Business Paperback*. Random House, New York, NY, 2014.
- <sup>56</sup> Martha K. Landesberg, Toby Milgrom Levin, Caroline G. Curtin, and Ori Lev. Privacy online: A report to Congress. Technical report, Federal Trade Commission, Washington, DC, 1998.
- <sup>57</sup> Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, and Aaron Rieke. Discrimination through optimization: How facebook’s ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):199, 2019.
- <sup>58</sup> Amanda Agan and Sonja Starr. Ban the Box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics*, 133(1):191–235, 2017.
- <sup>59</sup> Solon Barocas and Helen Nissenbaum. Big data’s end run around anonymity and consent. In Julia Lane, Victoria Stodden, Stefan Bender, and Helen Nissenbaum, editors, *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, pages 44–75. Cambridge University Press, New York, NY, 2014.
- <sup>60</sup> Bradley A. Areheart and Jessica L. Roberts. GINA, big data, and the future of employee privacy. *Yale Law Journal*, 128(3):710–790, 2019.
- <sup>61</sup> Katherine J. Strandburg. Rulemaking and inscrutable automated decision tools. *Columbia Law Review*, 119(7):1851–1886, 1999.
- <sup>62</sup> Katherine J. Strandburg. Adjudicating with inscrutable decision tools. In Marcello Pelillo and Teresa Scantamburlo, editors, *Machines We Trust: Perspectives on Dependable AI*. MIT Press, Cambridge, MA, 2021.
- <sup>63</sup> Danielle Keats Citron. Technological due process. *Wash. UL Rev.*, 85:1249, 2007.
- <sup>64</sup> Andrew D Selbst and Solon Barocas. The intuitive appeal of explainable machines. *Fordham Law Review*, 87(3):1085, 2018.

- <sup>65</sup> Athanasios Andreou, Giridhari Venkatadri, Oana Goga, Krishna Gummadi, Patrick Loiseau, and Alan Mislove. Investigating ad transparency mechanisms in social media: A case study of Facebook’s explanations. In *NDSS 2018 - Network and Distributed System Security Symposium*, 2018.
- <sup>66</sup> Margot E Kaminski and Gianclaudio Malgieri. Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, 11(2):125–144, 2020.
- <sup>67</sup> Andrew D. Selbst. An institutional view of algorithmic impact assessments. *Harvard Journal of Law & Technology*, 35(1):117–191, 2021.
- <sup>68</sup> Ada Lovelace and UK DataKind. Examining the black box: Tools for assessing algorithmic systems. Technical report, Technical report, Ada Lovelace Institute, 2020.
- <sup>69</sup> Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. Who audits the auditors? recommendations from a field scan of the algorithmic auditing ecosystem. In *Conference on Fairness, Accountability, and Transparency*, pages 1571–1583, 2022.
- <sup>70</sup> Ifeoma Ajunwa. An auditing imperative for automated hiring. *Harvard Journal of Law & Technology*, 34(2):621–299, 2021.
- <sup>71</sup> Upton Sinclair. *The Jungle*. Penguin, New York, NY, 2006.
- <sup>72</sup> Chris Jay Hoofnagle. *Federal Trade Commission Privacy Law and Policy*. Cambridge University Press, New York, NY, 2016.
- <sup>73</sup> Statement of Commissioner Rebecca Kelly Slaughter In the Matter of Liberty Chevrolet, Inc. d/b/a Bronx Honda. Office of Commissioner Rebecca Kelly Slaughter, Federal Trade Commission, 2020.
- <sup>74</sup> Andrew D Selbst and Solon Barocas. Unfair artificial intelligence: How ftc intervention can overcome the limitations of discrimination law. *University of Pennsylvania Law Review*, 171, 2023.
- <sup>75</sup> Decision and Order, In re Everalbum, Inc., File No. 192-3172. Federal Trade Commission, 2021.
- <sup>76</sup> Robert Pitofsky. Past, present, and future of antitrust enforcement at the Federal Trade Commission. *The University of Chicago Law Review*, 72(1):209–227, 2005.
- <sup>77</sup> Elisa Jillson. Aiming for truth, fairness, and equity in your company’s use of AI, 4 2021.
- <sup>78</sup> Rebecca Kelly Slaughter, Janice Kopec, and Mohamad Batal. Algorithms and economic justice: A taxonomy of harms and a path forward for the federal trade commission. *Yale JL & Tech.*, 23:1, 2020.
- <sup>79</sup> Neil Richards and Woodrow Hartzog. A duty of loyalty for privacy law. *Washington University Law Review*, 99(3):961–1021, 2021.