HARVARD LAW SCHOOL | The Case Studies

CSP057
MARCH 2020

Machine Learning in the Underwriting of Consumer Loans

YINAN LIU AND TALIA GILLIS

Memorandum

TO: Senior Counsel
FROM: Director for the Office of Innovation at Consumer Financial Protection Bureau
DATE: March 2020
RE: Disparate Impact Analysis for Algorithmic Decision Making

Introduction

On February 11, 2019, President Trump issued an executive order[1] launching the American AI Initiative following up on the recommendations of the 2018 White House Summit on Artificial Intelligence for American Industry.[2] The initiative will focus the resources of the Federal government to establish America's place as the global leader in artificial intelligence (AI). One of the key areas of emphasis of this initiative is on setting AI governance standards, and ensuring that advances in AI will improve quality of life for the American people.[3]

The ability to access credit is a critical component for the quality of life for American families and individuals so that they have the opportunity to climb the economic ladder, build wealth, and achieve economic stability. In September 2018, our Bureau, the Consumer Financial Protection Bureau ("CFPB" or "Bureau") held a symposium, *Building a Bridge to Credit Visibility,*[4] aiming at expanding access to credit for consumers who face barriers to accessing credit. The Bureau estimates that 26 million Americans are "credit invisible" and that another 19 million people lack sufficient credit history, *i.e.*, "unscorable," which imposes on those consumers to substantial barriers to accessing credit or higher costs for credit. This burden falls upon almost 20% of the entire U.S. adult population. Credit invisibility affects some groups more than others. According to the Bureau's report, about 27-28% of minority populations are either

[1] Executive Office of the President. *Executive Order on Maintaining American Leadership in Artificial Intelligence* (February 11, 2019). https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

[2] The White House Office Of Science And Technology Policy. *Summary Of The 2018 White House Summit On Artificial Intelligence For American Industry* (May 10, 2018). https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf

[3] The White House Office Of Science And Technology Policy. *Accelerating America's Leadership in Artificial Intelligence* (February 11, 2019). https://www.whitehouse.gov/articles/accelerating-americas-leadership-in-artificial-intelligence/

[4] Patrice Ficklin and J. Frank Vespa-Papaleo, *Building a bridge to credit visibility - a report on the CFPB's credit visibility symposium*, CFPB (July 19, 2019), https://www.consumerfinance.gov/about-us/blog/report-credit-visibility-symposium/.

Written by Yinan Liu and Talia Gillis under the supervision of Howell E. Jackson, James S. Reid, Jr., Professor of Law at Harvard Law School. Case development at Harvard Law School is partially funded by a grant from Dechert LLP. Cases are developed solely as the basis for class discussion. They are not intended to serve as endorsements, sources of primary data, legal advice, or illustrations of effective or ineffective management.

credit invisible or have unscorable credit records. Young consumers and new immigrants are also more likely to be credit invisible.

These differences in credit access are sometimes the result of credit market discrimination. Specifically, lenders may have denied loans to some borrowers solely on the basis of race, ethnicity, and other personal traits.

More recently, the increased use of machine learning in the underwriting of consumer loans raises concerns about whether possible discriminatory use of big data is resulting in denial of equal credit access protection.[5] More specifically, alarms have been raised about whether data from historical discriminatory practice may be used to train a biased algorithm.

The Bureau is interested in identifying and regulating potential discrimination caused by machine-learning algorithms. At the same time, the Bureau also wants to encourage new strategies and innovation, including using machine-learning algorithms, to explore alternative data to improve the credit risk assessment process, and help provide affordable and sustainable credit to the "invisible" population.

The Bureau requests that a team of staff attorneys brief the Board on potential legal frameworks for supervising and enforcing algorithmic accountability in credit lending practice. The team has suggested two possible frameworks to regulate the use of machine learning algorithms. The first option is to adopt an *ex post* approach, following a new five-element burden-shifting framework proposed by the United States Department of Housing and Urban Development ("HUD") in 2019. The second option is to adopt an *ex ante* approach, a preclearing framework which evaluates lenders' algorithms before their deployment.

There is an internal debate within the legal team about which framework might be better for regulating discriminatory lending. We have been asked to evaluate the two proposals. To that end, please analyze the potential legal and policy issues that could arise from the two frameworks. Please carefully weigh each against the merits and demerits of maintaining the status quo.

Background

[5] Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values*, EXECUTIVE OFFICE OF THE PRESIDENT (May 1, 2014), https://bigdatawg.nist.gov/pdf/big_data_privacy_report_may_1_2014.pdf; Standing Committee of the One Hundred Year Study of Artificial Intelligence, *Artificial Intelligence and Life in 2030, One Hundred Year Study on Artificial Intelligence, Report of the 2015 Study Panel* 7 (Sept. 2016), https://ai100.stanford.edu/sites/default/files/ai_100_report_0831fnl.pdf.

The Federal Fair Lending Laws

The federal fair lending laws—the ECOA[6] and the FHA[7]—prohibit discrimination in credit transactions, including but not limited to transactions related to residential real estate.

Fair Housing Act (FHA) of 1968

Title VIII of the Civil Rights Act of 1968, the Fair Housing Act (FHA), prohibits discrimination in the sale, rental, or financing of dwellings and in other housing-related activities on the basis of race, color, religion, sex, disability, familial status, or national origin.

Congress gave HUD the authority and responsibility for administering the FHA and the power to make rules to carry out the Act.

Equal Credit Opportunity Act (ECOA) of 1974

The ECOA prohibits creditors from discriminating against credit applicants on the basis of race, color, religion, national origin, sex, marital status, age, an applicant's receipt of income from a public assistance program, or an applicant's good-faith exercise of any right under the Consumer Credit Protection Act.

It is worth to note that the language used in FHA and ECOA are different. Under Section 804(a) of the FHA, it includes a result-oriented language "or otherwise make unavailable." However, ECOA does not have such language. This difference is important in understanding the Supreme Court's decision in *Housing and Community Affairs v. Inclusive Communities Project* in 2015, which will be discussed in more details later.

The Home Mortgage Disclosure Act (HMDA)

The Home Mortgage Disclosure Act (HMDA) requires many financial institutions to maintain, report, and publicly disclose loan-level information about mortgages. The public data are modified to protect applicants' and borrowers' privacy[8]. The data helps shed light on lending patterns that could be discriminatory.

Traditional Credit Pricing Practice

Traditional risk-based pricing was introduced in the late 1980s to expand credit access to American consumers. Each borrower has unique characteristics that influence the probability of default on a loan. Risk-based pricing in consumer finance tailors the price and terms of a loan based on each borrower's likelihood of repayment.

Two tools were developed to help lenders pierce the fog of uncertainty surrounding each new loan applicant, which made the widespread of risk-based pricing possible. First, credit reports issued by a third-party credit bureau "institutionalize the sharing of consumer payment data," and "reduced the cost of

[6] Equal Credit Opportunity Act, 15 U.S.C 1691, *et seq*. The CFPB has issued regulations under ECOA, known as Regulation B, 12 C.F.R. Part 1002 (2019).
[7] The Fair Housing Act, 42 U.S.C. 3601, *et seq*.
[8] 12 U.S.C. Chapter 29.

assessing borrower risk."[9] Second, statistical credit scoring evaluates default risks, generates specific predictions and summarizes it to be a numerical score. It gave lenders a powerful tool for rapidly and consistently evaluating loan applications while reducing processing costs. Instead of rejecting applicants with high risk, lenders can accept them and charge an appropriately higher price for the loan to cover the extra risk.

Studies from Stanford and the University of Pennsylvania (Wharton) illustrate how risk-based pricing helped a lender mitigate both adverse selection and moral hazard through the adjustment of both interest rates and loan terms based on borrower risk.[10] The use of credit report data and credit scoring to prescreen borrowers brought in more lenders and the competition between lenders helped reduced finance charge rate and annual fees. This in turn helps expanding credit availability across all consumer loans.

However, despite its role in improving credit availability and affordability, critics of credit scoring have alleged that scoring models actually have an adverse effect on certain demographic groups. In a 2007 Report to Congress on the impact of credit scoring, the Federal Reserve concluded that (1) the increase in credit availability appears to hold for the population overall as well as for major demographic groups, including those of different races and ethnicities, and (2) it is true that different demographic groups have substantially different credit scores, on average.[11]

However, variables that have predictive value for credit risk will often also be correlated with demographic characteristics. Barring use of all such variables in credit scoring would undermine completely the use of credit scoring. "Risk-based pricing, by its very nature, leads to disparities based on credit characteristics, and if those discrepancies are deemed impermissible, and lenders are pushed to flatter pricing, the very consumers the government seeks to protect–high-risk borrowers–stand to lose the most."[12]

Thus, in this case study, we will focus on changes in disparities as a result of the adoption of a new pricing method—for example, whether the move from traditional credit pricing to algorithmic credit pricing decreased or increased disparities or whether the use of social media data increases disparities, rather than an absolute level of disparity.[13] For example, In a recent update, the CFPB described how an algorithmic lender, Upstart, had demonstrated that it provides cheaper loans than under traditional lending models and accepted applicants who would have been rejected under those models.[14]

The Role of the Consumer Financial Protection Bureau (CFPB)

[9] Center for Capital Markets. Risk-based Pricing in Consumer Lending. https://www.centerforcapitalmarkets.com/wp-content/uploads/2013/08/CCMC_RiskBasedPricing_FINAL_to_post_10_24_2014.pdf

[10] Adams, W., L. Einav, and J. Levin (2009): "Liquidity Constraints and Imperfect Information in Subprime Lending." American Economic Review, 99(1): 49–84; Einav, L., M. Jenkins, and J. Levin (2012): "Contract Pricing in Consumer Credit Markets." Econometrica, 80(4): 1387–1432; Einav, L., M. Jenkins, and J. Levin (2013): "The Impact of Credit Scoring on Consumer Lending." RAND Journal of Economics, 44(2): 249–274.

[11] Board of Governors of the Federal Reserve System (2007): Report to Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit. Submitted to the Congress pursuant to section 215 of the Fair and Accurate Credit Transactions Act of 2003; August 2007.

[12] Center for Capital Markets. Risk-based Pricing in Consumer Lending. https://www.centerforcapitalmarkets.com/wp-content/uploads/2013/08/CCMC_RiskBasedPricing_FINAL_to_post_10_24_2014.pdf

[13] Talia B. Gillis, *Discrimination Stress-Testing*. Draft

[14] Patrice Ficklin and Paul Watkins, An update on credit access and the Bureau's first No-Action Letter. https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/.

The Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 ("Dodd-Frank Act") transferred the authority to implement the ECOA to the CFPB. The CFPB prohibits discriminatory practice under the ECOA. In 2012, the CFPB announced that "[c]onsistent with other federal supervisory and law enforcement agencies, the CFPB reaffirms that the legal doctrine of disparate impact remains applicable . . . to enforce compliance with the ECOA . . . "[15] On May 21, 2018, the CFPB issued a statement indicating its intent to reexamine requirements of the ECOA in light of recent Supreme Court case law addressing the availability of disparate-impact legal theory under the Fair Housing Act ("FHA").[16]

Congress in section 1021(a) of the Dodd-Frank Act established the CFPB's statutory purpose as ensuring that all consumers have access to markets for consumer financial products and services and that markets for consumer financial products and services are fair, transparent, and competitive.[17] The CFPB has jurisdiction over banks, credit unions, securities firms, mortgage servicing operations, and other financial companies operating in the United States. The Bureau is responsible for monitoring markets for consumer financial products and services to identify risks to consumers and the proper functioning of such markets.

In particular, the Bureau has focused on fair lending, by providing oversight and enforcement of federal laws intended to ensure "fair, equitable, and nondiscriminatory access to credit for both individuals and communities."[18] The CFPB has issued regulations under the ECOA, known as Regulation B[19] and regulations under the HMDA, known as Regulation C.[20] These regulations provide the substantive and procedural framework for fair lending.

Interagency Cooperation

As required by the Dodd-Frank Act, the Bureau's Office of Fair Lending coordinated fair lending regulatory, supervisory, and enforcement activities with other Federal agencies and State regulators to promote consistent, efficient, and effective enforcement of federal lending laws, including but not limited to the ECOA.[21] The CFPB is authorized to bring public enforcement actions against any person, subject to the CFPB's supervisory or enforcement authority, for violations of the ECOA. The Dodd-Frank Act expressly authorizes the CFPB to conduct joint investigations with the U.S. Department of Justice ("DOJ") in matters relating to fair lending. The CFPB is also required to refer certain violations of the ECOA to the DOJ for possible enforcement actions.[22]

Unlike the ECOA, the CFPB does not have supervisory authority over the FHA under the Dodd-Frank Act.[23] The CFPB cooperates with the U.S. Department of Housing and Urban Development ("HUD") to further the purposes of the FHA.

Under the Dodd-Frank Act, Congress also transferred HMDA rulemaking authority and other functions to CFPB. Regulation C, 12 C.F.R. part 1003, implements the Home Mortgage Disclosure Act. HMDA agencies

[15] *CFPB Bulletin 2012-04 (Fair Lending)*, CFPB, (Apr. 18, 2012), http://files.consumerfinance.gov/f/201404_cfpb_bulletin_lending_discrimination.pdf.
[16] CFPB, *Statement of the Bureau of Consumer Financial Protection on enactment of S.J. Res. 57*, CONS. FIN. PROT. BUR. (May 21, 2018), https://www.consumerfinance.gov/about-us/newsroom/statement-bureau-consumer-financial-protection-enactment-sj-res-57/; *see* CFPB, *Fall 2018 Regulatory Agenda Preamble*, CONS. FIN. PROT. BUR. (Aug. 30, 2018), https://www.reginfo.gov/public/jsp/eAgenda/StaticContent/201810/Preamble_3170.html.
[17] 12 U.S.C. 5511(a).
[18] 12 U.S.C. § 5493(c)(2)(A).
[19] CFPB Equal Credit Opportunity Act (Regulation B) (2019), 12 CFR Part 1002.
[20] CFPB Home Mortgage Disclosure (Regulation C) (2020), 12 CFR Part 1003.
[21] Dodd-Frank Act § 1013(c)(2)(B) (codified at 12 U.S.C. § 5493(c)(2)(B)).
[22] CFPB, *Consumer Financial Protection Bureau and Justice Department pledge to work together to protect consumers from credit discrimination*, CONS. FIN. PROT. BUR. (Dec. 6, 2012). https://www.consumerfinance.gov/about-us/newsroom/consumer-financial-protection-bureau-and-justice-department-pledge-to-work-together-to-protect-consumers-from-credit-discrimination/.
[23] H.R. 4173 § 1027(s) ("No provision of this title shall be construed as affecting any authority arising under the Fair Housing Act.").

include both the CFPB and HUD, as well as other federal agencies, such as the Office of the Comptroller of the Currency ("OCC"), the Federal Deposit Insurance Corporation ("FDIC"), the Federal Reserve System, the National Credit Union Administration ("NCUA").[24]

Regulation B

The ECOA is implemented by Regulation B. It contains two basic and comprehensive prohibitions against discriminatory lending practices:[25] "(1) A creditor shall not discriminate against an applicant on a prohibited basis regarding any aspect of a credit transaction; (2) A creditor shall not make any oral or written statement, in advertising or otherwise, to applicants or prospective applicants that would discourage, on a prohibited basis, a reasonable person from making or pursuing an application." Note that the regulation is concerned not only with the treatment of persons who have initiated the application process, but also with lender behavior before the application is even taken. For example, a creditor may not advertise its credit services and practices in ways that would tend to encourage some types of borrowers and discourage others on a prohibited basis. In addition, a creditor may not use prescreening tactics likely to discourage potential applicants on a prohibited basis.

Discrimination Doctrines in Lending

There are two principles of discrimination doctrines: disparate treatment and disparate impact.[26] Disparate treatment occurs when a creditor treats an applicant differently based on a prohibited basis such as race or national origin.[27] Disparate impact occurs when a creditor employs facially neutral policies or practices that have an adverse effect or impact on a member of a protected class, unless those policies or practices meet a legitimate business need that cannot reasonably be achieved by means that are less disparate in their impact.[28]

There are ongoing disputes with respect to the foundations and scope of the different doctrines. First, is disparate treatment intended to focus on lenders with animus toward protected groups? Does disparate treatment also include mere direct consideration of a protected characteristic, even when it has a rational explanation or basis? Second, is disparate impact doctrine meant to "address only covert intentional discrimination, or more broadly address the furthering of preexisting disadvantage into the credit context"?[29]

While the ECOA and the FHA do not explicitly recognize the two discrimination doctrines in the language of the law itself, the disparate impact doctrine has been recognized with respect to credit pricing by courts and agencies in charge of enforcing the laws. In *Texas Department of Housing and Community Affairs v. Inclusive Communities Project*, the Supreme Court recognized disparate impact under the FHA in 2015.[30]

In *Inclusive Communities*, Justice Kennedy held that "antidiscrimination laws should be construed to encompass disparate-impact claims when their text refers to *the consequences of actions* and not just to the mindset of actors, and where that interpretation is consistent with statutory purpose."

[24] CFPB, *supra* note 11, at § 1003.1.
[25] 12 C.F.R. § 1002.4.
[26] Federal Fair Lending Regulations and Statutes. https://www.federalreserve.gov/boarddocs/supmanual/cch/fair_lend_over.pdf.
[27] 12 C.F.R. § 1002.4(a)-(b). *See also* official interpretation of paragraph 4(a) and 4(b) at https://www.consumerfinance.gov/policy-compliance/rulemaking/regulations/1002/4/.
[28] 12 C.F.R. Part 1002 Supp. I Sec. 1002.6(a)-2.
[29] Talia Gillis. *Discrimination Stress-Testing* (March 2020). Draft.
[30] *Texas Dep't of Hous. & Cmty. Affairs v. Inclusive Communities Project, Inc.,* 135 S. Ct. 2507 (2015).

In its analysis, the court reasoned that the phrase "or otherwise make unavailable" in Section 804(a) of the FHA is equivalent in function and purpose to "or otherwise adversely affect" in Title VII and ADEA. Justice Kennedy writes, "Congress's use of the phrase 'otherwise make unavailable' refers to the consequences of an action rather than the actor's intent." The decision states that in all "three statutes the operative text looks to results," and "this results-oriented language counsels in favor of recognizing disparate impact liability."

In addition, the majority opinion also points to four actions that confirm Congress's understanding that disparate impact liability exists under the FHA: (1) Congress passed the FHA only four years after passing Title VII and only four months after passing ADEA, all three with similar text and structure; (2) Congress passed the FHA amendments of 1988 knowing that all nine Courts of Appeals that had addressed the question concluded the FHA encompassed disparate impact claims; (3) The 1988 amendments included three exemptions to disparate impact that assume the existence of disparate impact claims under the FHA as enacted in 1968; and (4) also in 1988, Congress rejected a proposed amendment that would have eliminated disparate impact liability for certain zoning decisions. Thus, the court concluded that "This results-oriented language counsels in favor of recognizing disparate-impact liability" under FHA.

However, there are some ambiguities about disparate treatment and disparate impact under ECOA, because it includes no such "results-oriented language."

The U.S. Supreme Court also ruled that institutional policies are not contrary to the discrimination laws unless they are "artificial, arbitrary, and unnecessary barriers." The Court cautioned there must be adequate safeguards around application of disparate impact analysis to avoid setting "our Nation back in its quest to reduce the salience of race in our social and economic system," and that the "[c]ourt should avoid interpreting disparate-impact liability to be so expansive as to inject racial considerations into every housing decision" or "to second-guess" between "two reasonable approaches." In addition, the language of the FHA should not be construed to force defendants to "resort to the use of racial quotas." Disparate-impact liability must be limited so employers and other regulated entities are able to make the "practical business choices and profit related decisions that sustain a vibrant and dynamic free-enterprise system."

Although there is not an equivalent Supreme Court case with respect to ECOA, the CFPB and courts have found that the relevant statutes allow for a claim of disparate impact.[31]

Existing Discriminatory Framework

Under the existing framework, a defendant may be liable for practice with a discriminatory effect unless there is a legally sufficient justification.[32]

Discriminatory Effect

First, the government and private plaintiffs "bears the burden of proving its prima facie case that a practice" actually, or predictably: "(1) results in a disparate impact on a group of persons on the basis of race, color,

[31] *See, e.g., Ramirez v. GreenPoint Mortgage Funding, Inc,* 633 F. Supp. 2d 922, 926–27 (N.D. Cal. 2008)
[32] 24 C.F.R. Part 100, Implementation of the Fair Housing Act's Discriminatory Effects Standard; Final Rule.

religion, sex, handicap, familial status, or national origin; or (2) has the effect of creating, perpetuating, or increasing segregated housing patterns on the basis of race, color, religion, sex, handicap, familial status, or national origin."[33] Thus, plaintiffs need not allege any *intent* to discriminate against borrowers.

Legally Sufficient Justification

When the plaintiffs prove a prima facie case, the burden of proof shifts to the respondent or defendant to prove that there is a legally sufficient justification for the alleged practice. A legally sufficient justification exists where the challenged practice "(i) is necessary to achieve one or more of its substantial, legitimate, nondiscriminatory interests" of the respondent or the defendant; and "(ii) those interests could not be served by another practice that has a less discriminatory effect." A legally sufficient justification "must be supported by evidence and may not be hypothetical or speculative."[34]

Less Discriminatory Alternatives

If the defendant satisfies this burden, then the plaintiff may still establish liability by proving that the substantial, legitimate, nondiscriminatory interests could be served by a practice that has a less discriminatory effect.

Machine Learning in Lending

Basics of Machine Learning

Fueled by exponentially increasing data and compute power, machine learning is having an unprecedented impact on our everyday lives. Credit risk assessment represents one of the earliest uses of machine learning. For example, in 1992, the Bank of Scotland led an early effort to collaborate with academic researchers to apply machine learning algorithms in scoring credit-card applications and compare their performance.[35]

Machine learning is a method of using a computer to parse data, learn from it, and then make a prediction/determination based on the analysis. For example, the learning process for Random Forest,[36] a popular supervised machine learning technique, includes: (1) data gathering and cleansing; (2) splitting the data into a training dataset and a testing dataset; (3) training the model with the training dataset based on various machine learning algorithms; and (4) validating the model with the testing dataset.

Random Forest,[37] Artificial Neural Networks,[38] and Boosting[39] are three of the most popular machine learning techniques that have been adopted to credit risk assessment. Machine learning algorithms can analyze large volumes of data, and quickly and accurately identify patterns and trends that might not be apparent to a human. In addition, its performance can improve over time because of the increasing

---

33 *Inclusive Communities, supra* note 20, at 2523.
34 *Id*.
35 R. H. Davis, D.B. Edelman and A.J. Gammerman, *Machine-learning algorithms for credit-card applications*, 4 IMA JOURNAL OF MATHEMATICS 43-51 (1992).
36 Tony Yiu, *Understanding Random Forest: How the Algorithm Works and Why it Is So Effective*, TOWARDS DATA SCIENCE, (Jun.12, 2019), https://towardsdatascience.com/understanding-random-forest-58381e0602d2.
37 Deloitte, *Point of View: Using Random Forest for credit risk models*, Deloitte Risk Advisory, (Sept. 2018), https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/financial-services/deloitte-nl-fsi-using-random-forest-for-credit-risk-models.pdf.
38 Manish Bhoge, *Using the Artificial Neural Network for Credit Risk Management*, ORACLE DATA SCIENCE BLOG, (Jan. 23, 2019), https://blogs.oracle.com/datascience/using-the-artificial-neural-network-for-credit-risk-management.
39 Jocelyn D'Souza, *A quick guide to Boosting in ML*, MEDIUM (Mar. 21, 2018), https://medium.com/greyatom/a-quick-guide-to-boosting-in-ml-acf7c1585cb5.

amounts of data that are processed. As more data comes in, the algorithm becomes more experienced and can then make better predictions.

Potential Benefits of Machine Learning in Lending

In 2018, Fannie Mae, one of the federally sponsored agencies that purchases and guarantees mortgages, conducted a survey of mortgage lenders and their use of machine learning ("ML") in lending.[40] The survey found that over the next two years, the use of these technologies in the mortgage industry was going to boom. Some of the key drivers of such business change include rising customer expectations, cost-saving, and operational challenges.[41]

In particular, machine learning and AI solutions are helping banks and credit lenders use alternative data to evaluate creditworthiness to increase access to credit or decrease the cost of the credit. Mortgage lenders, generally, have a large amount of data regarding the applicant at the time of underwriting. The appeal for machine learning is that it manages, from the data, to uncover complex patterns and relevant variables that were not apparent in advance. This increased insight allows lenders to view borrowers as more than a handful of numbers, as occurs with traditional lending models. Such information can be used to more accurately assess credit risk than traditional rules-based underwriting. By considering alternative data, such as payment of bills or rents, loans will be available to a wider variety of people who, although they are good candidates to repay a mortgage, would be rejected today using traditional lending models because, for example, for whatever reason they lack a traditional credit history.

Results provided from CFPB's No-Action Letter Program show that machine learning with alternative data approves 23-29% more applicants and, compared with traditional models, lowers the average annual percentage rate by 15-17% for approved loans.[42] Such expansion of credit access occurs across all tested race, ethnicity, and sex segments. In particular, consumers with incomes under $50 thousand are 13% more likely to be approved.

Potential Challenges of Machine Learning in Lending

Despite the efficiency and accuracy gained via machine learning, it also presents certain significant potential risks.[43]

Yongqianbao (meaning, "Use Money, my pal") is an AI-based financial service app that provides small amount, short-term loans to the underbanked population in China. A Public Broadcasting Service documentary film shows that it only takes eight seconds for the Yongqianbao App to approve a loan by analyzing 5,000 features of the applicant while traditional banks only assess 10 features.[44] The 5,000 features include not only traditional credit information, but also unconventional features, such as smart phone behaviors, including what phone the borrower is using, how many calls go unanswered, and the battery level of the phone. It also uncovers some unexpected correlation: *e.g.*, how confident a borrower types in the loan application and whether the borrower keeps the cell phone charged are highly correlated to delinquency.

---

[40] Fannie Mae, *Mortgage Lender Sentiment Survey: How will Artificial Intelligence Shape Mortgage Lending? Q3 2018 Topic Analysis*, FANNIE MAE (Oct. 4, 2018), https://www.fanniemae.com/resources/file/research/mlss/pdf/mlss-artificial-intelligence-100418.pdf.

[41] Aseem Mital and Atul Varshneya (Tavant Technologies), *Reshaping Consumer Lending with Artificial Intelligence*, MORT. BANKERS ASSOC., https://www.tavant.com/sites/default/files/download-center/Tavant_Consumer_Lending_Artificial_Intelligence_Whitepaper.pdf

[42] Patrice Fickin and Paul Watkins, *An update on credit access and the Bureau's first No-Action Letter*, CONS. FIN. PROT. BUR. (Aug. 6, 2019), https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/.

[43] *See, e.g.*, Mikella Hurley and Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE L. J. & TECH. 148, 168 (2016); Matthew Adam Bruckner, *The Promise and Perils of Algorithmic Lenders' Use of Big Data*, 93 CHI. -KENT L. REV. 3, 25-29 (2018).

[44] *In the Age of AI, Frontline*, PUBLIC BROADCASTING SYSTEM (Nov. 5, 2019), https://www.pbs.org/wgbh/frontline/film/in-the-age-of-ai/.

While such use of alternative information may help expand credit access to under-represented groups, it raises privacy concerns because of the use of invasive personal information without the borrower's awareness. Notably, Yongqianbao has also been accused multiple times for predatory lending where the annual interest rate was as high as 109.5% and for abusing its AI technology to harass borrowers' social media and phone contacts for debt collection.[45]

Even when companies do not intend to discriminate and deliberately avoid using suspect classifications like race and gender, the output of an analytical process can have a disparate impact on a protected class. An algorithm could be trained with historically biased data and simply reflect preexisting discrimination or disparities. Analytics relying on existing data could reinforce and worsen past discriminatory practices. Fair lending law has traditionally addressed this concern by limiting the consideration of protected characteristics, such as race, gender, etc. However, other correlated input information, for example, zip code, may serve as a proxy for protected characteristics, and it is difficult to exclude all such correlated information.

It is also less feasible to apply the legal doctrine of disparate treatment to the practice of algorithmic credit pricing. The ubiquity of correlations in big data, combined with the flexibility and complexity of machine learning models, means that one cannot rule out the consideration of a protected characteristic, even when formally excluded.[46]

One recent study, "False Dreams of Algorithmic Fairness: the Case of Credit Pricing,"[47] has conducted simulations highlighting concerns over approaches focusing on input scrutiny. More specifically, the author argues that because information about a person's protected characteristic is embedded in other information about the individual, it is unlikely that the formal exclusion of the protected characteristic as an input guarantees that the characteristic was not used in forming a decision. Machine learning algorithms are particularly adept at uncovering some previously unknown relationship between alternative information and the protected characteristics because of the ubiquity of correlations in big data. On the other hand, machine learning algorithms can be unstable and may find spurious correlations even when there are none.[48] It is not enough to expand the prohibited inputs to also include "proxies" for protected characteristics. It is also not clear whether a showing of correlation should be admitted as evidence to demonstrate disparate treatment.

A disparate impact assessment seeks to assess the extent of any disproportionate, adverse impact on a protected class, identifies the extent to which the use of the algorithm contributes to a legitimate organizational or social need, and explores whether there are equally effective models with lesser disparate impact.

CFPB's Initiative with Machine Learning in Lending

45 *Repeatedly Banned? Exploited with Quinbao illegal collection of annualized interest rate of nearly 110%*, TOUTIAO, http://toutiao.manqian.cn/wz_9PnoVRZltf.html.

46 *Evaluating the Fair Lending Risk of Credit Scoring Models*, CHARLES RIVER ASSOC. Feb. 2014), http://www.crai.com/sites/default/files/publications/FE-Insights-Fair-lending-risk-credit-scoring-models-0214.pdf. The concern that big data analysis might discriminate inadvertently is explicitly recognized: "Ostensibly neutral variables that predict credit risk may nevertheless present disparate impact risk on a prohibited basis if they are so highly correlated with a legally protected demographic characteristic that they effectively act as a substitute for that characteristic."

47 Talia Gillis, *False Dreams of Algorithmic Fairness: the Case of Credit Pricing*, talk at Harvard Empirical Legal Studies, Harvard Law School (Sept. 20, 2019), available at https://scholar.harvard.edu/files/gillis/files/gillis_jmp_191101.pdf.

48 Robin Wigglesworth, *Spurious correlations are kryptonite of Wall St's AI rush*, FIN. TIMES (Mar. 14, 2018), https://www.ft.com/content/f14db820-26cd-11e8-b27e-cc62a39d57a0.

Regulatory uncertainty can substantially hinder the development of innovative products and services that may ultimately benefit consumers. The CFPB has been working towards improving its policies to promote innovation and facilitate compliance.

In 2017, the Bureau issued a Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process (RFI).[49] The RFI sought public comments regarding the potential positive and negative consequences associated with the use of alternative data to "encourage responsible use of alternative data and lower unnecessary barrier impeding its use."

In September 2017, the CFPB issued a Non-Action Letter[50] to Upstart Network Inc., an artificial intelligence lending platform.[51] In addition to using traditional factors such as credit score and income, Upstart also evaluates consumer loan applications incorporating non-traditional sources of information, such as education and employment history. Learning from its experience with Upstart, in 2019, the CFPB updated its NAL policy as it found its prior policy "not an adequate response to the extent of innovation occurring in markets for consumer financial products and services."[52] The CFPB also issued new Trial Disclosure Program (TDP) Policies,[53] and the new Compliance Assistance Sandbox (CAS) Policy.[54]

The new policy provides a more streamlined review process for products and services.[55] It also no longer requires applicants to show that they are likely to provide "substantial" benefit to consumers but rather a "potential" benefit.[56] Since its adoption of the new policy, the CFPB has provided NALs with respect to housing counseling agreements by Housing Counseling Agencies (HCA), which are regulated by the HUD.[57]

Under the CFPB's new TDP Policy, entities seeking to improve consumer disclosures may conduct in-market testing of alternative disclosures for a limited time upon receiving permission from the CFPB.

Under the CAS Policy, after the CFPB evaluates the product or service for compliance with relevant law, an approved applicant that complies in good faith with the terms of the approval will have a "safe harbor" from liability for specified conduct during the testing period. Approvals under the CAS Policy will provide

[49] CFPB, Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process, 82 Fed. Reg. 11183 (Feb. 21, 2017), https://www.govinfo.gov/content/pkg/FR-2017-02-21/pdf/2017-03361.pdf.

[50] The authorities given to the CFPB under title X of Dodd-Frank form the basis for the CFPB's NAL policy. 81 FR 8686 (Feb. 22, 2016) [the original NAL policy]. The SEC also has a mechanism for No Action letters, which may be considered as a more case-by-case adjudication on the part of the regulation. Although Dodd-Frank does not explicitly provide the authority for NAL, they are considered under the CFPB's supervisory and enforcement discretion. For further background on the differences between adjudication versus advance rule making, see David L. Shapiro, *The Choice of Rulemaking or Adjudication in the Development of Administrative Policy*, 78 Harv. L. Rev. 921 (1964–1965). For a more recent discussion, see Yehonatan Givati, *An Incomplete Contracting Approach to Administrative Law*, 18 Am Law Econ Rev 176 (2016).

[51] CFPB Announces First No-Action Letter to Upstart Network, CFPB (Sept. 14, 2017), https://www.consumerfinance.gov/about-us/newsroom/cfpb-announces-first-no-action-letter-upstart-network/ Upstart is an artificial intelligence lending platform. In addition to using traditional factors such as credit score and income, Upstart also evaluates consumer loan applications incorporating non-traditional sources of information such as education and employment history.

[52] Bureau of Consumer Financial Protection, Policy on No-Action Letters, Vo. 84 No. 178, page 48229 (September 13, 2019) [hereinafter New NAL Policy]. The Final Rule can be found at: https://files.consumerfinance.gov/f/documents/cfpb_final-policy-on-no-action-letters.pdf and https://www.govinfo.gov/content/pkg/FR-2019-09-13/pdf/2019-19763.pdf

[53] Policy to Encourage Trial Disclosure Programs, CFPB, https://files.consumerfinance.gov/f/documents/cfpb_final-policy-to-encourage-tdp.pdf.

[54] Policy on the Compliance Assistance Sandbox, CFPB, https://files.consumerfinance.gov/f/documents/cfpb_final-policy-on-cas.pdf.

[55] For example, the new policy states the intention of the CFPB to grant or deny an application within 60 days and includes a policy for obtaining NAL modifications.

[56] See New NAL Policy.

[57] The first NAL was issued in September 2019 and covered HUD supervised HCAs that enter agreements with consumers, as long as they comply with certain requirements such as providing consumers with a memorandum of understanding with the consumer reflecting the terms of the housing counselling. It provided a similar NAL to Bank of America in January 2020. See: https://www.consumerfinance.gov/about-us/newsroom/cfpb-issues-no-action-letter-to-facilitate-housing-counseling-services/

protection from liability under ECOA and other fair lending laws.[58] The regulatory sandbox policy can be helpful for fostering incentive to innovate when the legal uncertainty as to how algorithms can comply with fair lending laws.

However, there are various legal challenges in creating sandboxes.[59] The authority to regulate financial institutions and activities is fragmentation and overlapping in the United States and there is a complex relationship between federal and state regulation.[60] The possibility that financial activity trigger enforcement from regulators other than the CFPB could potentially undermine the benefit of the sandbox, as it does not provide innovators with added certainty.

For example, the Office of the Comptroller of the Currency (OCC) has opened its own "sandbox" through a Proposed Innovation Pilot Program designed to promote its innovation initiatives, add value through proactive supervision, and continue its objective to lead fintech innovation expansion.[61] Unlike the CFPB Sandbox, entities accepted under the OCC Program will receive no immunity from complying with applicable laws and regulations.

Moreover, following the proposal for the CFPB sandbox, twenty-two state attorney generals wrote a letter to the CFPB stating that the bureau could not provide a safe harbor protecting them from enforcement action.[62] The AGs also questioned whether it would be sufficient to adopt a sandbox as a policy rather than adopting the proposal through the formal rulemaking process. The CFPB, on the other hand, considers its authority to create a sandbox policy under its supervisory and enforcement discretion provided by Dodd-Frank.

However, the CFPB has adopted several approaches to mitigate this risk by committing to engage in outreach to other federal and state regulators and to coordinate when challenges arise.[63] Moreover, the creation of the American Consumer Financial Innovation Network (ACFIN), a network of federal and state regulators to facilitate innovation, is also expected to assist in coordination.[64]

[58] U.S. DEP'T OF TREASURY, A FINANCIAL SYSTEM THAT CREATES ECONOMIC OPPORTUNITIES (2018), https://home.treasury.gov/sites/default/files/2018-07/A-Financial-System-that-Creates-Economic-Opportunities---Nonbank-Financi....pdf. The Treasury recommended creating a regulatory sandbox to promote innovation.

[59] Talia Gillis, Discrimination Stress Testing. Draft.

[60] For example, one issue that can arise is the preemption of state laws, such as in the case of licensing or usury. The CFPB's approach in addressing these concerns that arise as a result of activity triggering several other regulatory bodies, is that the CFPB will coordinate with other regulators to resolve potential issues. Another initiative of the CFPB has been the creation

[61] OCC Solicits Public Comment on Proposed Innovation Pilot Program. https://www.occ.treas.gov/news-issuances/news-releases/2019/nr-occ-2019-42.html

[62] *See* Kate Berry, State AGs assail CFPB plan to build fintech sandbox, Am. Banker (Feb. 12, 2019, 4:14pm EST), available at: https://www.americanbanker.com/news/state-ags-assail-cfpb-plan-to-build-fintech-sandbox.

[63] https://www.jdsupra.com/legalnews/paul-watkins-director-of-the-cfpb-s-86267/

[64] See announcement by CFPB, at https://www.consumerfinance.gov/about-us/newsroom/bureau-state-regulators-launch-american-consumer-financial-innovation-network/. Allen suggests a different approach in which a committee of regulators approve sandbox applications. *See* Allen, *supra* note 138, at 622. While this approach would allow for better coordination it is likely to be highly unfeasible.

Framework Choices for CFPB

Option 1: 2019 HUD's Five-Element Burden-Shifting Framework

In August 2019, HUD proposed a new burden-shifting framework to establish discriminatory liability under the FHA.[65] This new framework replaced the previous three-step burden-shifting framework with a five-element approach. This rule is also one of the first attempts in the U.S. to determine whether an algorithm violates the FHA.

More specifically, a plaintiff's allegations that a specific, identifiable, policy or practice has a discriminatory effect must plead facts supporting five elements to establish a prima facie case: a plaintiff is required to allege that (1) "the challenged policy or practice is *arbitrary, artificial, and unnecessary* to achieve a valid interest or legitimate objective such as a practical business, profit, policy consideration, or requirement of law"; (2) there is "*a robust causal link* between the challenged policy or practice and a disparate impact on members of a protected class"; (3) the challenged policy or practice has "an adverse effect on *members of a protected class"*; (4) "the disparity caused by the policy or practice is *significant"*; and (5) "the *complaining party's alleged injury* is directly caused by the challenged policy or practice"[66] [Emphasis ours].

The defendant may rebut a claim at the pleading stage by asserting that a plaintiff has not alleged facts to support their prima facie claim. In a case when a challenged policy relies on an algorithmic model, three defenses are available to the defendant to defeat the claim.

The first defense allows a defendant to provide an analysis showing that "the model is not the actual cause of the disparate impact alleged by the plaintiff." The defendant may "break down the model piece-by-piece and demonstrate how each factor considered could not be the cause of the disparate impact and show how each factor advances a valid objective." The plaintiff may defeat this defense by showing that "the defendant's analysis is somehow flawed, such as by showing that a factor used in the model is correlated with a protected class despite the defendant's assertion."

There are cases where defendants may not have access to the model developed by a third party. For example, government-sponsored enterprises Fannie Mae and Freddie Mac require lenders to evaluate credit risk pursuant to automated underwriting systems containing models proprietary to those enterprises. Yet, lenders have no ability to alter the models used by Fannie Mae and Freddie Mac, and lenders are not in a position to justify each element of such a model, much less the relationships among all the variables.

In this case, a second defense provides that a defendant can show that "use of the model is standard in the industry." A recognized third party, not the defendant, is responsible for creating or maintaining the model. It is being used for the intended purpose of the third party. A plaintiff may rebut this allegation by showing that the plaintiff is not challenging the standard model alone, but that the defendant's unique use or misuse of the model is the cause of the disparate impact.

[65] HUD's Implementation of the Fair Housing Act's Disparate Impact Standard. A Proposed Rule by the Housing and Urban Development Department (Aug. 19, 2019), https://www.federalregister.gov/documents/2019/08/19/2019-17542/huds-implementation-of-the-fair-housing-acts-disparate-impact-standard.

[66] *City of Miami v. Bank of America Corp.*, 137 S. Ct. 1296, 1305-06 (2017). Decided two years after *Inclusive Communities*, *supra* note 20, the U.S. Supreme Court held that Fair Housing Act plaintiffs must satisfy a "requirement" of pleading "direct" proximate causation. Under this requirement, a lender cannot be held liable under the FHA for a disparate impact where the independent actions of independent third parties break the proximate causal chain.

A third defense allows a defendant to "prove through the use of a qualified expert that the model is not the cause of a disparate impact." A plaintiff may rebut this defense by showing that "the expert is not neutral, that the analysis is incomplete, or that there is some other reason why the expert's analysis is insufficient evidence that the defendant's use of the model is justified."

Once a plaintiff adequately alleges facts to support the prima facie claim, the defendant then has the burden to identify a valid interest/interests that the challenged policy or practice serves.

Finally, having articulated a legitimate business goal, the defendant should prevail unless the plaintiff can prove that "other tests or selection devices, without a similarly undesirable racial effect, would also serve the legitimate business interests in an equally effective manner."

Even if a policy or practice that has a disparate impact on a prohibited basis can be justified by business necessity, it still may be in violation if an alternative policy or practice could serve the same purpose with less discriminatory effect. For example, refusing to lend to all people in a heavily minority neighborhood has a disproportionate adverse impact on a protected class, but it might have a legitimate purpose, namely, to avoid making loans that could not be repaid. On average, people in red-lined neighborhoods are more likely to have bad credit risks. Thus, avoiding those neighborhoods controls credit risk, a business necessity, but may still be in violation because there are other ways of achieving the same purpose of controlling credit risk without disparate impact.

Option 2: Pre-clearing Framework: Discrimination Stress-Testing

Option 1 focus on auditing the algorithms in an *ex post* manner to ensure that it is operating properly.

However, there have been proposals that the *ex post* analysis should be supplemented by testing algorithms at the development stage for potential bias.[67] An outcome-based framework is proposed by "False Dreams of Algorithmic Fairness: the Case of Credit Pricing" and described in more details in "Discrimination Stress Testing." The framework shifts away from input scrutiny. Instead, a regulator will apply a pricing rule to a designated dataset to analyze the properties of the pricing rules. This option gives regulators more flexibility.

There are two questions the regulator should address using this framework—first, whether borrowers who are "similarly situated" are treated the same; and second, whether the pricing rule increases or decreases disparities relative to some baseline. The framework requires regulators to build a neutral database with real or hypothetical people and their characteristics, then apply a lender's pricing rule to a dataset of hypothetical borrowers, and then examine the properties of the outcome. The goal is for the regulator to examine the algorithm in an *ex ante* manner.

The testing involves three stages. At the first stage, the lender determines the input and algorithms for the pricing rule to use for credit risk assessment. This stage does not involve a regulator. At the second stage, the regulator takes the algorithm and applies it to the neutral database. The rationale for using a database is that it is difficult to analyze a prediction function in the abstract.[68] At the third stage, the

[67] Terrell McSweeny, Comm'r, Fed. Trade Comm., Keynote Remarks at Google Tech Talk: *Tech for Good: Data for Social Empowerment*, (Sept. 10, 2015), https://www.ftc.gov/system/files/documents/public_statements/800981/150909googletechroundtable.pdf; Gillis, *supra* note 38.

[68] Talia B. Gillis and Jann L. Spiess, *Big Data and Discrimination*, 86 U. CHI. L. REV. 459 (2019).

regulator evaluates the outcome to determine whether the disparities created by the pricing rule amount to discriminatory conduct. The outcome metric can be credit price, or other metrics, *e.g.*, error rates.[69] The exact criteria to be used in outcome analysis will depend on what the discrimination laws aim to achieve which, unfortunately, has not been clear and agreed upon.[70]

The discrimination stress-testing can be helpful in quantifying and understanding incremental changes in access to credit as a result of AI algorithms, i.e., whether the move from traditional credit pricing to algorithmic credit pricing decreased or increased disparities.

However, it is unclear whether the lenders or the regulators should run the outcome-based analysis. There may be proprietary concern for the lenders to give their algorithms to regulators to test. It is also possible for third parties to game the algorithm by analyzing the output. On the other hand, regulators may not have the resources or manpower to run such analyses, and, if they do, it would also discourage the lender's autonomy. For lenders serving small communities, requirement of model validation may discourage them from adopting innovative products or services.

An alternative approach is to create a voluntary regime. Firms are not required to receive permission from the regulator to implement algorithmic credit pricing but can apply for regulatory approval or temporary protection from regulatory enforcement.

Challenges also exist in defining the safe harbor provided to firms when participating in the sandbox, as doing so requires the delicate balancing of competing concerns.[71] On the one hand, for a sandbox to be effective, the safe harbor provided must be significant, otherwise there is only a weak incentive to participate in the sandbox. On the other hand, safe harbors could limit the regulator's ability to act when a product or service indeed proves to be harmful or have unintended consequences.[72] Beyond the requirements for and restrictions in entering sandbox programs, such as demonstrating the potential benefit of the innovation to consumers, safe harbors can be adjusted to balance the competing policy goals. For example, limiting the time for which the safe harbor is provided means that firms do not have an ongoing exemption from complying with regulation.[73] In the United Kingdom, the continuing communication with the designated officer means that the FCA has greater visibility of the potential harm of a product and engages in softer forms of regulation even when other regulatory action is not possible. One of the most significant potential benefits of a sandbox approach is that it may lead to more formal rulemaking with respect to algorithmic pricing lending. There is currently scant guidance on how to analyze an algorithm for the purpose of fair lending, and while algorithmic pricing is an outlier in how credit is currently priced, this is likely to change in the long term. Therefore, while in the short-term regulators may need to provide a safe policy space in which to innovate, in the long-term, regulators will need to create formal rule making.[74] Regulatory sandboxes may allow for regulators to develop the understanding and expertise needed for new regulations.[75]

[69] Sam Corbett-Davies and Sharad Goel, *The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning*, available at CORNELL UNIV. arXiv:1808.00023 [cs:CS]v2 6 (Aug. 14, 2018), https://arxiv.org/pdf/1808.00023.pdf.

[70] Gillis, *supra* note 38.

[71] Talia Gillis. Discrimination Stress Testing. Draft.

[72] For a general discussion of the types of relief sandboxes could provide, see *Id.* at 623

[73] Other countries have chosen a more sweeping approach. For example, in Australia, ASIC provides a full exemption from licensing requirements. See *Id.*

[74] One of the concerns with sandboxes is that it may lead to regulators avoiding formal rule-making, which may be more beneficial in the long term. See Matthew J. Razzano, *AN UNSAFE SANDBOX: FINTECH INNOVATION AT THE EXPENSE OF CONSUMER PROTECTION?*, 2019, 139 8 (2019).

[75] Rule-making that lacks understanding and expertise could result in confusing and incoherent regulation. See *False Dreams*, Section 3.1, for a discussion of HUD's Proposed Rule on the Implementation of the Disparate Impact standard under the Fair Housing Act. For further analysis in the risks of regulating innovation before it is adequately understood, see Wu, *supra* note 124.

Legal and Policy Considerations

Whether to choose either of the two frameworks depends on the objective and motivations of CFPB. Designing would need to take into consideration a variety of issues, as set forth below.

Ex Ante Approach or Ex Post Approach

Timing is an important consideration in regulatory design. One issue to consider in designing the framework is whether we should adopt an *ex ante* approach, which focuses on preventing financial harm, or an *ex post* approach, which focuses on mitigating the harm.

In an ideal world, with perfect government information, perfectly rational actors, and complete liquidity, *ex post* is equivalent to *ex ante*.[76] However, the reality is that the regulators will not have perfect information, and the actors are not always rational. The choice of an *ex ante*, as compared to an *ex post* approach, is essentially a problem of balancing factors such as the policing capacity of *ex ante* penalties, the cost of consequences to the lender resulting from its conduct compared to the cost to the lender of preventing the conduct, and the ability to know what conduct leads to adverse consequences, etc.[77]

Ex ante regulation, such as requiring disclosure, works by reducing asymmetric information among the regulated parties and regulators and making the risks transparent to all. With an *ex ante* regulatory-based approach, there is more certainty beforehand. In addition, the regulators will have more flexibility in designing the framework. This is particularly beneficial for regulating the algorithm-based lending industry where the industry is continually upgrading their techniques with alternative data, new machine-learning algorithms, etc. Moreover, instead of focusing on the narrow question of whether a particular lender discriminated or not, *ex ante* regulation improves market analysis and the understanding in general whether credit markets are well-functioning and serve the protected groups adequately by considering the effect big data and machine learning have on market-wide access to credit.[78]

However, *ex ante* regulation is inherently limited. Complex financial markets innovate more quickly than regulators can adapt. Thus, the regulators may not have perfect information, and *ex ante* regulation cannot anticipate every cause and prevent every financial failure. This may lead to inefficient under-regulation of some, and over-regulations of others. *Ex ante* regulation that attempts to prevent all financial harms may end up impeding economic growth or undermine the market experimentation and innovation on which market growth depends.[79]

Ex post remedies initiate only after the harm has occurred, but will help prevent financial harm from spreading and becoming systemic. [80] The threat of liability forces the market players to take appropriate

[76] Christine Jolls, *On Law Enforcement with Boundedly Rational Actors*, THE LAW AND ECONOMICS OF IRRATIONAL BEHAVIOR 268, 272 (Francesco Parisi and Vernon L. Smith eds.) (2005).
[77] STEVEN SHAVELL, Foundations of Economic Analysis of Law 87-91, 428-30, 479-82, 492-520, 572-78 (2004).
[78] Talia Gillis, *Discrimination Stress-Testing*. Draft.
[79] Steven L. Schwarcz, *Systemic Risk*, 97 Georgetown L.J. 193, 217 (2008).
[80] Steven L. Schwarcz, *Ex Ante Versus Ex Post Approaches to Financial Regulation*, 15 Chapman L. Rev. 257 (Jan. 31, 2011).

precautions and to internalize the expected social damages.[81] *Ex post* regulation also provides useful additional information when regulated parties are heterogeneous.[82]

However, when *ex post* liability is used exclusively, the uncertainty of legal standards will lead to inefficiencies. *Ex post* lawsuits may not always be brought against regulated parties due to the uncertainty of the legal standard,[83] or the remedy may come too late and the harm no longer can be prevented. Moreover, biased perceptions of the legal standards may also cause the regulated parties to take under- or over-precautions.

Ex post approaches and *ex ante* approaches may supplement each other as part of a comprehensive regulatory framework. For example, *ex ante* regulation can correct the inefficiencies associated with the use of *ex post* liability alone.[84]

Plaintiff's Burden or Defendant's Burden

Another issue to consider in designing the framework is how to properly allocate the burden of proof between plaintiffs and defendants in assessing disparate impact.

The five-element framework shifts more of the burden to the plaintiffs. It is insufficient to identify a program as a whole without explaining how the program itself causes the disparate impact, as opposed to a particular element of the program. Plaintiffs must identify the particular policy or practice rather than a one-time decision that causes the disparate impact. Without access to the algorithm and dataset, it would be extremely difficult for a plaintiff to establish a prima facie case for algorithmic decision-making. Even with access to the algorithm and dataset, the element of robust causality may be challenging to plead and prove because there is no clear distinction between correlation and causation. It is also not clear how to benchmark "an adverse effect on *members of a protected class*" of the third element according to the previous discussion on group fairness versus individual fairness.

The Court has placed special emphasis on the importance of the plaintiff's prima facie burden, warning that, ''[w]ithout adequate safeguards at the prima facie stage, disparate-impact liability might cause race to be used and considered in a pervasive way and would almost inexorably lead governmental or private entities to use numerical quotas, and serious constitutional questions then could arise.''[85]

Is it fair to ask either plaintiff or defendant to bear the burden of proof in each of these stages? For example, finding statistical disparity may be complicated by the difficulty of obtaining information about protected class status. Law often prevents the collection of protected characteristics to prevent the organization from engaging in disparate treatment. For example, ECOA and Regulation B generally

[81] Charles D. Kolstad, Thomas S. Ulen and Gary V. Johnson, *Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements? Working Papers Series on the Political Economy of Institutions No.4*, Vol. 80, Col. of Com. and Bus. Admin., Bur. of Econ. and Bus. Res., Univ. of Ill., 888-901 (Dec., 1987), https://www.ideals.illinois.edu/bitstream/handle/2142/29359/expostliabilityf1419kols.pdf?sequence%3D1.

[82] Brian Galle, *In Praise of Ex Ante Regulation*, 68 Vanderbilt L. Rev. 1715 (March 13, 2015).

[83] In the U.S. Supreme Court's landmark decision in *Wal-Mart Stores, Inc. v. Dukes*, 131 S.Ct. 2541 (2011), the Court ruled that a group of roughly 1.5 million women could not be certified as a valid class of plaintiffs in a class-action lawsuit for employment discrimination against Walmart. In this case, the plaintiffs alleged systemic pay and promotion discrimination against women. The plaintiff's sociology expert offered statistical analysis to show a corporate culture vulnerable to gender bias failed to meet this burden. In dismissing the expert's analysis, Justice Scalia's opinion focused on the expert's inability to ascertain whether 0.5 percent or 95 percent of employment decisions at Wal-Mart were determined by stereotyped thinking. The decision substantially raised the bar for plaintiffs to obtain class certification in all types of class actions, and is likely to have a substantial impact on class action law beyond the employment realm. This particularly poses challenges for future anti-discrimination class-action cases, because the Court concluded that, where Wal-Mart's policies allowed local supervisors substantial discretion over pay and promotion matters, the plaintiffs failed to identify "a common mode of exercising discretion that pervades the entire company." Wal-Mart Stores, Inc. v. Dukes

[84] *Ex ante* regulation can correct cases of under-precaution resulting from exposure to *ex post* liability alone, but it can also exacerbate over-precaution. *See* Kolstad, *supra* note 56.

[85] *Inclusive Communities*, supra note 20.

prohibits a creditor from inquiring "about the race, color, religion, national origin, or sex of an applicant or any other person in connection with a credit transaction."[86] There are a few exceptions, including exceptions for applications for home mortgages covered under the HMDA, and for the purpose of "self-test."[87] This creates a Catch-22 situation that while the lenders may be prevented from engaging in disparate treatment by seeking statistical data about disparity, the plaintiffs are also prevented from establishing a prima facie case by showing that statistical disparity.[88]

On the other hand, if defendants bear the burden of proof, would it possible that such a burden would be ultimately transferred to consumers?

[86] 12 CFR § 1002.5(b).
[87] 12 C.F.R. § 1002.5(a)-(d) and 12 C.F.R. § 1002.13.
[88] *See* CFPB, *Using publicly available information to proxy for unidentified race and ethnicity: A methodology and assessment* (Summer, 2014). The CFPB defined an approach using surname and geographic probabilities, Bayesian Improved Surname Geocoding (BISG), to estimate the probability that each applicant belonged to a particular ethnicity and race group. These probabilities are used to assign an ethnicity and race proxy as a substitute for missing data in assessing Fair Lending risk. It is demonstrated that the BISG proxy probability is more accurate than a geography-only or surname-only proxy in its ability to predict individual applicants' reported race and ethnicity and is, generally, more accurate than a geography-only or surname-only proxy at approximating the overall reported distribution of race and ethnicity., https://files.consumerfinance.gov/f/201409_cfpb_report_proxy-methodology.pdf.

Conclusion

Machine learning is a significant development that will continue to have increasing applications within the mortgage lending industry. Within machine learning, mortgage lenders and underwriters are going to be more efficient and capable of expanding access to credit and lower the cost. At the same time, it is also important to ensure the use of machine learning with alternative data does not result in disparate and discriminatory treatment.

Most of the time, machine learning algorithms are viewed as a "blackbox," but they should not be handled as invisible designs. Policymakers must understand the limitations of machine learning before endorsing these tools. Machine learning needs to be effective, but also transparent, about how they are making decisions about the reasons that certain applicants are denied credit or obtain different interest rates than others.

The CFPB is dedicated to expanding fair, equitable, and nondiscriminatory access to credit and to ensuring that consumers are protected from discrimination. The CFPB remains optimistic that the financial sector will find valuable ways to employ machine learning techniques in the near future.

Briefing Questions

Having examined the proposals, please review the materials included in the attached Appendix and prepare to brief the Director. In particular, the Director is eager to know your thoughts as to the following:

- To what extent is the practical and desirable to impose a requirement of transparency or explanation in the use of algorithms in decision-making in order promote fairness?
- What are the comparative strengths and weaknesses of the two options: ex post enforcement versus an ex ante review process of some sort for algorithm driven loan underwriting?
- How well does each option align with the decision and analysis in *Inclusive Communities*?
- How might each option impact the likelihood of success of disparate impact claims brought in the future?
- What are the pros and cons for choosing between "legally sufficient justification" and "business justification" requirements in the burden-shifting framework?
- What would be a likely reaction from industry players to these proposals? Consider the following examples:
    - Civil Rights Groups
    - Local government
    - Banks
    - AI companies
- How might each option increase or decrease costs and economic burden to relevant parties?
- For the ex ante option, consider the following questions
    - What is considered a "neutral" database for evaluating an algorithm?
    - What are the potential challenges in constructing a neutral database?
    - Will there be proprietary concern for the lenders to give their algorithms to regulator to test, or should the lenders take control of running the outcome-based analysis?
    - What principles should be used to determine who is similarly situated?

- How will each option impact credit-accessibility to underserved communities?
- Who should be liable if a practice is found discriminatory? The lender or the third-party companies who design the algorithm?

Appendices

1. Executive Order on Maintaining American Leadership in Artificial Intelligence https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

Credit Invisible

2. Building a bridge to credit visibility - a report on the CFPB's credit visibility symposium, https://www.consumerfinance.gov/about-us/blog/report-credit-visibility-symposium/.

Federal Fair Lending Laws

3. Equal Credit Opportunity Act (Regulation B) 12 CFR Part 202.
4. Discriminatory Conduct Under The Fair Housing Act. 24 CFR Part 100.

Disparate Impact

5. Federal Fair Lending Regulations and Statutes. Page 2-3, https://www.federalreserve.gov/boarddocs/supmanual/cch/fair_lend_over.pdf.

Supreme Court Caselaw

6. Texas Department of Housing and Community Affairs v. Inclusive Communities Project, Inc., 135 S. Ct. 2507 (2015).

The Role of CFPB

7. CFPB Fair Lending Report. Page 7-11, Page 15, https://files.consumerfinance.gov/f/documents/201906_cfpb_Fair_Lending_Report.pdf.

Machine Learning

8. Gentle Introduction to Machine Learning, https://youtu.be/Gv9_4yMHFhI

Machine Learning in Lending

9. An update on credit access and the Bureau's first No-Action Letter, https://www.consumerfinance.gov/about-us/blog/update-credit-access-and-no-action-letter/.
10. An Update from CFPB on Upstart's No-Action Letter, https://www.upstart.com/blog/an-update-from-cfpb-on-upstarts-no-action-letter.

Challenges of Machine Learning in Lending

11. Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process, https://www.govinfo.gov/content/pkg/FR-2017-02-21/pdf/2017-03361.pdf

Optional reading:

12. R Rahul Bhargava, "The Algorithms Aren't Biased, We Are," Medium (Jan. 3, 2018), https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2.
13. Finale Doshi-Velez et al., "Accountability of AI Under the Law: The Role of Explanation," https://arxiv.org/abs/1711.01134.

Ex post Regulation and Ex Ante Regulation

14. Steven Shavell, Foundations of Economic Analysis of Law 87-91, 428-30, 479-82, 492-520, 572-78 (2004).
15. HUD's Implementation of the Fair Housing Act's Disparate Impact Standard: A Proposed Rule by the Housing and Urban Development Department on 08-19-2019, https://www.federalregister.gov/documents/2019/08/19/2019-17542/huds-implementation-of-the-fair-housing-acts-disparate-impact-standard.
16. Talia Gillis, *False Dreams of Algorithmic Fairness: the Case of Credit Pricing*, talk at Harvard Empirical Legal Studies, Harvard Law School (Sept. 20, 2019), available at https://scholar.harvard.edu/files/gillis/files/gillis_jmp_191101.pdf.
17. Talia Gillis, *Discrimination Stress Testing*.