Equitable Algorithms: How Human-Centered AI Can Address Systemic Racism and Racial Justice in Housing and Financial Services

Testimony of Stephen F. Hayes
Partner, Relman Colfax PLLC

Before the Task Force on Artificial Intelligence
United States House Committee on Financial Services

May 7, 2021

Chairman Foster, Ranking Member Gonzalez, and distinguished members of the Task Force on Artificial Intelligence, thank you for hosting this important hearing and for giving me the opportunity to submit this testimony.

My name is Stephen Hayes and I am a Partner at Relman Colfax PLLC. Relman Colfax is a national civil rights law firm. We have a litigation practice focused on combating discrimination in areas such as housing and lending, and we regularly represent individuals, non-profits, and municipalities bringing redlining, reverse redlining, and other civil rights claims. We also provide legal counsel to entities such as financial institutions, Internet-based companies, and nonprofits on civil rights compliance. My work focuses largely on providing fair lending and fair housing advice, including legal counsel on testing models for discrimination risks. I previously worked in the Legal Division of the Consumer Financial Protection Bureau ("CFPB"). I hope that my testimony today furthers the Committee's understanding of algorithms and discrimination, and the potential for using Artificial Intelligence and Machine Learning ("AI" and "ML") to increase opportunity, equity, and inclusiveness in financial markets.

A. Credit Discrimination, Alternative Data, and AI/ML

Our Nation's history of housing, employment, and credit discrimination has contributed to disparities in income, wealth, rates of home and small business ownership, and access to other important life opportunities. Existing credit markets reflect that history: studies have found evidence of racial disparities in credit scoring and factors on which traditional scores rely.[1] Disparities also exist with respect to populations without credit histories—those that are "credit invisible" or "thin file." Black and Hispanic Americans are more likely than white or Asian

---

[1] *See, e.g.*, Lisa Rice & Deidre Swesnik, *Discriminatory Effects of Credit Scoring on Communities of Color*, 46 Suffolk U. L. Rev. 935, 952-959 (2013); Aaron Klein, Brookings Institution, "Reducing bias in AI-based financial services" (July 10, 2020), https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services; Bd. of Governors of the Fed. Reserve Sys., Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit (2007); CFPB, "Analysis of Differences Between Consumer- and Creditor-Purchased Credit Scores" (Sept. 2012); National Consumer Law Center, Credit Discrimination § 6.4.1.1, "Studies Showing that Minorities Have Lower Credit Scores as a Group" (7th ed. 2018).

Americans to be credit invisible or to have unscored records.[2] These gaps are pernicious because it can be difficult to build a credit history without access to credit, limiting certain consumers' ability to improve their financial circumstances.

At the same time, many of these individuals may in fact be good credit risks. There is evidence that some "alternative data"—information not traditionally found in the credit files of the nationwide consumer reporting agencies and not commonly provided by consumers on credit applications—can help lenders make accurate underwriting and loan pricing decisions for these consumers.[3] Accordingly, some market participants supplement traditional lending strategies with alternative data, and regulatory agencies have worked to facilitate the responsible use of certain types of alternative data.[4]

Alternative data is distinct from, but often discussed in combination with, "alternative models" like ML models. The term "alternative models" can refer to methods for constructing predictive models from historical data without requiring human modelers to explicitly specify relationships among the variables that can be used in the model.[5] Like with alternative data, there is some evidence that the use of alternative models (and, in particular ML algorithms), has the potential to expand credit access by allowing lenders to evaluate the creditworthiness of consumers who are difficult to score accurately using traditional techniques.[6]

The use of alternative data and alternative models can also raise real concerns, including serious risks related to accuracy, completeness, explainability, validity, barriers to improving credit standing, and discrimination.[7] Moreover, fair lending concerns are not resolved solely because a practice increases access to credit; increases in access to even favorable credit or financial products can drive persistent inequality—and disparate impact—if distributed

---

[2] Kenneth P. Brevoort, et al., CFPB Office of Research, "Data Point: Credit Invisibles," at 6 (May 2015), https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf.

[3] *See, e.g.,* CFPB, Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process, 82 Fed. Reg. 11183, 11185 (Feb. 21, 2017).

[4] *See, e.g.,* FinRegLab, "The Use of Cash-Flow Data in Underwriting Credit" at 8 (July 2019), https://finreglab.org/wp-content/uploads/2019/07/FRL_Research-Report_Final.pdf.; Bd. of Governors of the Fed. Rsrv. Sys., CFPB, FDIC, NCUA, OCC, "Interagency Statement on the Use of Alternative Data in Credit Underwriting," at 1 n.1 (Jan. 2019), https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20191203b1.pdf.

[5] *See* Nicholas Schmidt & Bryce Stephens, "An Introduction to Artificial Intelligence and Solutions to the Problems of Algorithmic Discrimination," Consumer Finance Law Quarterly Report, Vol. 72, No. 2, 131, at 133 (2019), https://arxiv.org/ftp/arxiv/papers/1911/1911.05755.pdf.

[6] Patrice Alexander Ficklin, et al., CFPB Blog, "Innovation spotlight: Providing adverse action notices when using AI/ML models" (July 7, 2020), https://www.consumerfinance.gov/about-us/blog/innovation-spotlight-providing-adverse-action-notices-when-using-ai-ml-models/; Fed. Rsrv. Bd. Governor Lael Brainard, *Supporting Responsible Use of AI and Equitable Outcomes in Financial Services,* Speech at the AI Academic Symposium hosted by the Bd. of Governors of the Fed. Rsrv. Sys. (Jan. 21, 2021), https://www.federalreserve.gov/newsevents/speech/brainard20210112a.htm.

[7] Brian Kreiswirth, et al., CFPB Blog, "Using alternative data to evaluate creditworthiness," (Feb. 16, 2017), https://www.consumerfinance.gov/about-us/blog/using-alternative-data-evaluate-creditworthiness/; *see also, e.g.,* Solon Barocas and Andrew Selbst, "Big Data's Disparate Impact," 104 Cal. L. Rev. 671 (June 2016); FinRegLab, "AI in Financial Services: Key Concepts" (May 2020), https://finreglab.org/wp-content/uploads/2020/05/FinRegLab_AIFAQ_Key-Concepts_AI-in-Financial-Services.pdf; Caroline Wang et al., "In Pursuit of Interpretable, Fair and Accurate Machine Learning for Criminal Recidivism Prediction" (May 2020), https://arxiv.org/abs/2005.04176.

unequally. That said, whether alternative data and alternative models will fairly increase access to credit in any given situation depends on a number of criteria, such as the type of alternative data and models at issue, how those tools are deployed, and characteristics of applicant pools.

B. Legal framework

Two primary federal antidiscrimination laws—the Equal Credit Opportunity Act ("ECOA") and the Fair Housing Act ("FHA")—prohibit institutions from discriminating in lending on the basis of characteristics such as race, national origin, religion, and sex.[8] ECOA applies to nearly all lending, including lending to businesses. The FHA applies to housing discrimination, including lending related to residential real estate transactions.

These laws prohibit intentional and overt discrimination, sometimes called "disparate treatment," as well as an unintentional type of discrimination called "disparate impact." Disparate treatment occurs when an entity explicitly or intentionally treats people differently based on prohibited factors, such as race, national origin, or sex. Unlike disparate treatment, disparate impact does not require any showing of intent to discriminate or that the protected characteristic was considered at all. Instead, the focus of disparate impact is on outcomes. Generally, unlawful disparate impact occurs when a (1) facially neutral policy or practice disproportionately adversely impacts members of protected classes, and either (2) the policy or practice does not advance a legitimate interest, or (3) is not the least discriminatory way to serve that interest.[9] These frameworks translate well to assessments of lending models, including AI/ML models.

C. Testing Models for Discrimination

Questions regarding how to ensure that algorithms are not discriminatory have received a significant amount of attention in recent years, particularly with the increased use of alternative data and alternative models. At the same time, these questions are not being written on a blank slate, either legally or with respect to institutions' internal compliance practices. Certain companies, including many financial services companies, have been testing models for discrimination for years and have systems in place guiding those assessments. Of course, even the most robust existing systems can be improved, and disparities in credit markets remain; although essential, improving model fairness alone will not solve those disparities. At the same time, existing programs demonstrate that model discrimination testing is both possible and effective, and—even if it may not make sense to incorporate these frameworks wholesale into other markets—these systems can serve as guides for markets where testing is nonexistent or nascent.

The methodologies that institutions use for fair lending testing their models vary, but as a general matter the most effective systems are designed to align with regulatory expectations and traditional principles gleaned from antidiscrimination jurisprudence. These systems often include: (1) ensuring that models do not include protected classes or close proxies for protected

---

[8] 15 U.S.C. § 1691(a); 12 C.F.R. § 1002.2(z); 42 U.S.C. § 3605.

[9] *See, e.g.*, 12 C.F.R. part 1002, Supp. I, ¶ 6(a)-2 (ECOA articulation); 24 C.F.R. § 100.500(c)(1) (FHA articulation); 42 U.S.C. § 2000e-2(k) (Title VII articulation).

classes, for example as variables or segmentations; and (2) assessing whether facially-neutral models are likely to disproportionately lead to negative outcomes for a protected class, and if such negative impacts exist, ensuring the models serve legitimate business needs and evaluating whether changes to the models would result in less of a disparate impact while maintaining model performance.[10]

This final step—identifying whether less discriminatory alternatives exist—is key. In the case of traditional statistical models, identifying less discriminatory alternatives often involves a process of adding, dropping, or substituting variables in the model, with the goal of identifying variations that maintain reasonable performance but that have less disparate impact on protected classes.[11] Newer methods exist that can improve upon that process for ML models. Advancements in computing power, along with sophisticated algorithms, can help analyze the impact of many different sub-combinations of variables, which allows institutions to explore numerous iterations of variable combinations and adjustments to hyperparameter settings. Other techniques also exist, such as training models to optimize for performance and metrics of fairness. In short, not only can these searches work for ML models, they can be more effective than traditional methods because there are more options for model adjustments.

As noted, disparities in the current credit system are stark. That said, disparate-impact testing—including the adoption of less discriminatory alternatives—has proven critical in reducing credit inequalities. It has caused lenders to search for and implement model variations that predict accurately and reduce disparate outcomes. This process can benefit both businesses and consumers. Among other things, identifying less discriminatory practices can help businesses responsibly diversify their borrower pools, which can lead to more representative training samples and increases in access to credit over time. Less discriminatory models also help mitigate disparities, counteract the legacies of historic credit discrimination, and close unnecessary credit gaps.

In robust programs, these quantitative statistical tests are paired with more holistic compliance measures: fair lending training for relevant staff, including modelers; ensuring teams have diverse backgrounds and are empowered to identify and remedy issues; reviewing policies and procedures within which models operate; and assessing areas of discretion to ensure that the potential for judgmental bias is mitigated.

D. Why Do Some Institutions Test Models But Others Do Not?

Although some companies routinely test their models for discrimination risks, many do not. Several legal and structural characteristics contribute to this unevenness.

---

[10] *See* Initial Report of the Independent Monitor, Fair Lending Monitorship of Upstart Network's Lending Model at 7 (April 14, 2021) ("Initial Upstart Report"), https://www.relmanlaw.com/media/cases/1086_Upstart%20Initial%20Report%20-%20Final.pdf; Nicholas Schmidt & Bryce Stephens, "An Introduction to Artificial Intelligence and Solutions to the Problems of Algorithmic Discrimination," Consumer Finance Law Quarterly Report, Vol. 72, No. 2, 131, at 141–142 (2019), https://arxiv.org/ftp/arxiv/papers/1911/1911.05755.pdf; David Skanderson & Dubravka Ritter, Federal Reserve Bank of Philadelphia, "Fair Lending Analysis of Credit Cards" 38–40 (2014).

[11] *See* Initial Upstart Report at 11.

First, agency supervision spurs internal testing. Many lenders are supervised by the CFPB, and have a long history of supervision by banking regulators such as the Federal Deposit Insurance Corporation, Office of the Comptroller of the Currency, the Federal Reserve Board, and the National Credit Union Administration. Supervisory expectations with respect to models have not always been consistent, but in general supervisory authority means that lenders' models can be (and at times have been) subject to scrutiny and therefore are expected to be fair and transparent in ways that are not true for non-supervised companies.

Second, lenders have been on notice for years that, via ECOA and the FHA, disparate treatment and disparate impact apply to their credit models, and so many lenders have developed methodologies to address these risks.[12] However, discrimination and disparate impact is not clearly prohibited in all markets.[13] For example, financial regulatory agencies focus on credit discrimination but historically have not regulated discrimination related to other core consumer financial activities like acquiring checking accounts, credit reporting, or third-party debt collection—all areas in which models may be used. To the extent some antidiscrimination laws explicitly apply in these areas, they are generally not enforced by federal agencies and prohibit only disparate treatment. Disparate treatment, standing alone, is unlikely to ensure that models are non-discriminatory. Simply removing explicit protected class information from models will not eliminate bias and discrimination. Disparate impact, on the other hand, has historically proven effective for increasing equitable access to credit and housing.

Third, general regulatory model risk management ("MRM") expectations that are not directly related to discrimination can nonetheless complement and encourage appropriate fair lending model compliance.[14] MRM principles are articulated through guidance meant to help supervised banks avoid adverse consequences like financial loss and safety and soundness risks that can occur because of inaccurate or misused models. These MRM expectations foster responsible model development, accuracy, validation, use, and monitoring. They can facilitate fair lending testing because, among other reasons, they can help create an "audit trail" for models. Effective MRM programs ensure that modelers catalog models, assess the representativeness of data, validate that models work across populations, ensure that models are only used as intended, and establish a routine cadence for reviewing models.

These characteristics make it more likely that supervised entities providing services like credit will have systems for addressing discrimination risks arising from model usage. Importing similar principals and methodologies to model use in other markets could help advance equity and ensure models do not perpetuate historic disparities unnecessarily. At the same time, there is significant variation with respect to testing models for discrimination even within supervised financial institutions. This variation could be addressed through clear regulatory expectations, a

[12] HUD, DOJ, OCC, OTS, Fed. Rsrv. Bd., FDIC, FHFB, FTC, NCUA, Policy Statement on Discrimination in Lending, 59 Fed. Reg. 18266 (Apr. 15, 1994) ("Joint Policy Statement on Lending Discrimination"); HUD Final rule, Implementation of the Fair Housing Act's Discriminatory Effects Standard, 78 Fed. Reg. 114460, 11476 (Feb. 15, 2013).
[13] Stephen Hayes & Kali Schellenberg, "Discrimination is 'Unfair,'" Student Borrower Protection Center at 10-13 (April 2021), https://protectborrowers.org/wp-content/uploads/2021/04/Discrimination_is_Unfair.pdf.
[14] *See, e.g.,* Bd. of Governors of the Fed. Rsrv. Sys., OCC, SR Letter 11-7, "Supervisory Guidance on Model Risk Management" (Apr. 4, 2011).

need for which has become increasingly important given the rising interest in and use of alternative data and alternative models.

E. Moving Forward to Ensure Alternative Model Use is Fair and Equitable

Policymakers can take concrete steps to ensure models, including AI/ML models, are built in a transparent and accountable manner and result in fair and equitable outcomes. If appropriately tested, these models can serve as important tools to spur innovation, improve customer experiences, promote inclusiveness, and help overcome historic disparities experienced by protected classes.

1. <u>Regulatory agencies should routinely test models for discrimination, including assessing disparate impact and identifying less discriminatory alternatives.</u>

There is an uneven landscape with respect to how or whether institutions assess their models for discrimination. The CFPB and other agencies with supervisory authority can promote uniform internal testing by making clear that the agencies will review institutions' model testing results. Where internal testing is insufficient, the agencies should conduct independent testing, including assessing whether less discriminatory alternatives exist.

2. <u>Regulatory agencies should announce the methodologies they use to test models for discrimination.</u>

Even among financial institutions that conduct rigorous fair model testing, questions exist as to acceptable methodologies. This leaves institutions in a precarious situation, with some dedicating resources towards compliance without a clear picture of regulatory expectations. Institutions might also be overly cautious about using promising alternative data or modeling techniques that could benefit consumers for fear of regulatory risk. The CFPB should explain what methodologies it will use in supervision and enforcement so that entities can align their internal systems accordingly. The CFPB should also tailor MRM-like guidance specifically to fair lending assessments, and encourage more widespread adoption of these techniques by entities beyond those that are directly supervised by the federal banking agencies.

3. <u>Regulatory agencies should clarify that discrimination, including unnecessary disparate impact, is illegal across markets. To the extent ambiguities exist, Congress can explicitly codify coverage.</u>

Regulatory agencies should articulate that discrimination—including unnecessary disparate impact—is prohibited across markets outside of just credit and housing. This clarification would spur internal antidiscrimination measures, including testing, in areas that have historically been unregulated. Agencies like the CFPB and FTC currently have the tools to regulate many of these activities.[15] To the extent there are statutory authority concerns, Congress should consider explicitly codifying coverage.

[15] Hayes & Schellenberg, "Discrimination is 'Unfair,'" at 20-21.