Testimony of Aaron Rieke
Managing Director, Upturn

Before the
Task Force on Financial Technology
United States House Committee on Financial Services

Hearing entitled
"Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit"

Thursday, July 25, 2019
Rayburn House Office Building, Room 2128

Chairman Lynch, Ranking Member Hill, and distinguished members of the Task Force, thank you for the opportunity to appear before you today to discuss the use of alternative data in credit underwriting.

My name is Aaron Rieke. I'm the Managing Director at Upturn, a nonprofit organization that advances equity and justice in the design, governance, and use of digital technology. Upturn is an integrated team of computer scientists, lawyers, and policy professionals. We conduct technically-grounded research and advocacy in partnership with some of the nation's leading civil rights and consumer groups. My work at Upturn focuses primarily on the ways that new technologies either help or hinder economic opportunity, particularly for low-income consumers and people of color.

Alternative data is an important topic because approximately 45 million Americans do not have access to fair and affordable credit. They cannot get traditional credit scores because national consumer reporting agencies do not currently collect and maintain sufficient data about them.[1] This is an economic justice and civil rights issue because consumers who are low-income, Black, and Hispanic are disproportionately likely to be unscorable.[2]

***Some* types of alternative data have the potential to benefit millions of underserved consumers — if properly regulated and used in appropriate circumstances.** However, the details matter immensely. Certain uses of alternative data could be harmful or discriminatory.

[1] Consumer Financial Protection Bureau, Data Point: Credit Invisibles, May 2015, 6, http://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf.
[2] *Id.*

My testimony highlights five key points that are critical for this Task Force to consider as it weighs the collection and use of alternative data for credit underwriting.

1. ***Conventional*** **alternative data holds significant promise, while** ***fringe*** **alternative data raises a range of concerns.**

The term "alternative data" typically refers to any information not traditionally used by the national consumer reporting agencies to calculate credit scores.[3] This broad definition masks a crucial distinction between *conventional alternative data* (data about how individual consumers handle their financial commitments) and *fringe alternative data* (data that is much less directly related to individual consumers' finances, such as social media data).[4] Conventional alternative data holds significant promise, while fringe alternative data raises a range of concerns.

The data traditionally maintained by national consumer reporting agencies have strong predictive power for credit underwriting. The reason is simple and intuitive: Data about how consumers handle their existing financial obligations — for example, whether they make timely payments on existing lines of credit — are closely related to their ability to take on new ones. Scores like FICO and VantageScore derive most of their predictive power from consumers' payment histories.[5]

This same logic applies to conventional alternative data. The best available evidence suggests that many kinds of bill repayment histories are predictive of credit repayment, and can help otherwise unscoreable consumers access credit. Some kinds of alternative data have already been used for many years. For example, nationwide consumer reporting agencies have incorporated rental and utility payment data, when available, in some of their scoring products.[6]

---

[3] United States Government Accountability Office, Agencies Should Provide Clarification on Lenders' Use of Alternative Data, December 2018, https://www.gao.gov/assets/700/696149.pdf.

[4] *See, e.g.,* Comments of Consumer Action, CFPB Request for Information: Use of Alternative Data and Modeling Techniques in Credit Process, May 2017, *available at* https://www.regulations.gov/document?D=CFPB-2017-0005-0058 (distinguishing "cellphone usage, social media habits, online purchases, payday loan usage, bank account balances, and marketing and collections data from text messages, education levels, apps downloaded and mobile wallet balances.")

[5] VantageScore estimated that 75% of all defaulting consumers are identified by just one variable (payment history) in one of the company's credit scoring models. Sarah Davies, Big Data and Credit Scoring, 2015 NCRC Annual Conference Presentation (on file with author).

[6] *See, e.g.,* Vantage Score, VantageScore Credit Scores and the Mortgage Market, 2018, https://www.vantagescore.com/images/resources/FAQs%20VantageScore%20Credit%20Scores%20and%20the%20Mortgage%20Market.pdf; TransUnion, Rent Payment Reporting with TransUnion and PayLease is a "No-Brainer" for Property Management Companies and Residents June 25, 2018, https://www.transunion.com/blog/rent-payment-reporting-with-transunion-paylease-is-a-no-brainer. A study using Experian's consumer credit profiles of New York City residents from each of the city's neighborhoods showed that the addition of rental data had the potential to raise scores for 76 percent of tenants. New York City Comptroller's Office, Making Rent Count: How NYC Tenants Can Lift Credit Scores and Save Money, October 23, 2017, https://comptroller.nyc.gov/reports/making-rent-count/rent-and-credit-report/.

Bank account transaction data are another type of conventional alternative data that shows promise in enhancing access to credit. This data can provide a timely portrait of a consumer's income and expenses, which can be strong evidence of their ability to repay a loan. It is important to ensure that consumers are considered on the merits of their overall financial picture, and not the details of their specific behaviors (*e.g.*, *where* a consumer chooses to shop). With the cooperation of banks, transaction data can be shared — with a consumer's express permission — in a way that is relatively easy for lenders to authenticate. Accordingly, we support the Consumer Financial Protection Bureau's efforts to ensure that consumers have robust choices to use and share their financial account data, subject to appropriate privacy and data security standards.[7]

On the other hand, evidence about the predictiveness and fairness of fringe alternative data is equivocal.[8] Some companies claim to be using "thousands of data points" in their credit scoring models, but these models are rarely subjected to independent scrutiny. Expansive datasets about people's social connections, where they live, how they behave, where they shop, and how they communicate are fraught with fair lending concerns. Even players in this new underwriting industry themselves emphasize that fringe credit models "can raise serious risks for institutions and consumers."[9]

There is significant hype and misinformation about fringe alternative data. For example, media reports have long suggested the existence of a "Facebook credit score."[10] But in fact, Facebook has a longstanding policy that prohibits *any* use of Facebook data for making eligibility determinations, regardless of whether a consumer consents to such use.[11] This doesn't mean some companies won't try to violate Facebook's rules, but it does mean that we shouldn't expect it to happen at scale anytime soon.

---

[7] *See generally* Request for Information Regarding Consumer Access to Financial Records, Docket No.: CFPB-2016-0048,
http://files.consumerfinance.gov/f/documents/112016_cfpb_Request_for_Information_Regarding_Consumer_Access_to_Financial_Records.pdf.

[8] Upturn, Knowing the Score: New Data, Underwriting, and Marketing in the Consumer Credit Marketplace, October 2014, 12, https://www.upturn.org/static/files/Knowing_the_Score_Oct_2014_v1_1.pdf.

[9] Testimony of Dr. Douglas Merrill, ZestFinance, U.S. House Committee on Financial Services, Perspectives on Artificial Intelligence: Where We Are and the Next Frontier in Financial Services, June 26, 2019, *available at* https://financialservices.house.gov/calendar/eventsingle.aspx?EventID=403824.

[10] *See generally* Aaron Rieke, Don't let the hype over "social media scores" distract you, August 11, 2016, https://medium.com/equal-future/dont-let-the-hype-over-social-media-scores-distract-you-33b5ccfbceb0; *see also* Robinson Meyer, "Could a Bank Deny Your Loan Based on Your Facebook Friends?" The Atlantic, September 25, 2015, *available at* https://www.theatlantic.com/technology/archive/2015/09/facebooks-new-patent-and-digital-redlining/407287/.

[11] Facebook Platform Policy 3.15, https://developers.facebook.com/policy/ (last visited July 23, 2019).

Fringe data can be predictive in particularly blunt ways. For example, consumers purchasing iPhones are usually more affluent than consumers purchasing other smartphones.[12] Accordingly, the difference in default rates between customers using Apple and Android phones is significant.[13] However, this correlation does not suggest it would be desirable — or even legal, from a fair lending standpoint — to start evaluating people for credit based on the type of mobile device they use. In short, seemingly innocuous data about people's behaviors might provide lenders with statistical stereotypes, but they are far less likely to capture individual circumstances to expand access to credit for those who need it most.

2. **Use of alternative data should take into account consumer protections at the state and local level.**

Use of alternative data, even high-quality conventional alternative data, should not be used in ways that diminish important consumer protections. For example, many states shield vulnerable populations from loss of electric and natural gas utility service during high cost months and times of illness or financial hardship.[14] As such, any use of utility bills in the credit process should not undermine these protections.[15] Similarly, use of rental history data should not penalize tenants who invoke their rights under state or local laws to withhold rent due to poor conditions. These are preventable issues, but will require close coordination with consumer groups.

3. **Alternative data should be deployed in ways that minimize non-credit uses.**

It is critical to ensure that the Fair Credit Reporting Act (FCRA) be applied to all companies that collect and share third-party data that is used or expected to be used as a factor in determining eligibility purposes, consistent with the plain text of the statute.[16] At the same time, it is important to remember that consumers' credit files are also often used for non-credit purposes.

---

[12] Tobias Berg, et. al, On the Rise of FinTechs – Credit Scoring Using Digital Footprints, Michael J. Brennan Irish Finance Working Paper Series Research Paper No. 18-12, July 15, 2019, *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3163781.

[13] *Id.*

[14] National Consumer Law Center, Full Utility Credit Reporting: Risks to Low-Income Consumers, July 2012, https://www.nclc.org/images/pdf/energy_utility_telecom/consumer_protection_and_regulatory_issues/ib_risks_of_full_utility_credit_reporting_july2012.pdf.

[15] For example, one research group offered some potential consumer friendly choices, including "not reporting small unpaid balances on accounts that are closed; not indicating that a customer is subsidized, on a payment plan or in forbearance; only reporting payments over 60 days overdue as late; not reporting retrospective data when the furnisher first begins reporting; and clearly communicating with customers that their payments be fully reported to CRAs." Policy and Economic Research Council, A New Pathway to Financial Inclusion: Alternative Data, Credit Building, and Responsible Lending in the Wake of the Great Recession, June 2012, 23, http://www.perc.net/wp-content/uploads/2013/09/WEB-file-ADI5-layout1.pdf.

[16] For recent developments on the application of the FCRA *see* Chi Chi Wu, Data Gatherers Evading the FCRA May Find Themselves Still in Hot Water, June 14, 2019, https://library.nclc.org/data-gatherers-evading-fcra-may-find-themselves-still-hot-water.

For example, some employers check job applicants' credit history, a practice that has been criticized as an illegitimate barrier to employment.[17] And although the FCRA restricts the use of credit reports for marketing, credit bureaus still sell aggregate credit records that include individual households' finances for marketing purposes.[18]

The specter of non-credit uses complicates the task of deciding whether alternative data is more likely to help or hurt consumers. We encourage Congress and the broader industry to explore ways to limit the use of alternative data to the sole purpose for which they were introduced — credit underwriting.

4. **Alternative data will require renewed efforts to protect consumers from predatory financial practices.**

As alternative data becomes more readily available, it will inevitably be used by predatory actors to identify and pursue financially vulnerable consumers. Federal and state regulators have uncovered instances where consumer reporting data, or data *derived* from consumer reporting data, have been used in ways that violate the FCRA.[19] Moreover, my organization's research has shown that internet payday lenders are incredibly savvy at snaring online users into loans with usurious interest rates.[20] To ensure that alternative data actually benefits consumers, Congress must ensure that the collection and use of this data is coupled with strong financial product protections and vigorous enforcement.

5. **Alternative data will require new regulatory guidance to avoid new sources of discrimination and ensure new credit models are transparent.**

We recommend that all new alternative data sources — even those that have predictive power — should be studied carefully for their association with protected class status.[21] Measuring for discrimination can be difficult as credit models become more complex, and because most non-

---

[17] *See, e.g.,* Amy Traub, Discredited: How Employment Credit Checks Keep Qualified Workers Out of a Job, 2012, http://www.demos.org/discredited-how-employment-creditchecks-keep-qualified-workers-out-job (finding that one in four survey participants who were unemployed said that a potential employer had requested to check their credit report during a job application).

[18] Upturn, Knowing the Score: New Data, Underwriting, and Marketing in the Consumer Credit Marketplace, October 2014, 17, https://www.teamupturn.com/static/files/Knowing_the_Score_Oct_2014_v1_1.pdf.

[19] *See, e.g.,* FTC Settlements Require Equifax to Forfeit Money Made by Allegedly Improperly Selling Information about Millions of Consumers Who Were Late on Their Mortgages, October, 2012, https://www.ftc.gov/news-events/press-releases/2012/10/ftc-settlements-require-equifax-forfeit-money-made-allegedly.

[20] Upturn, *Led Astray: Online Lead Generation and Payday Loans,* October 2015, https://www.upturn.org/reports/2015/led-astray/.

[21] For example, in 2007, the Federal Reserve Board used a nationally representative sample of more than 300,000 credit records, enriched with demographic data from other sources, to test for what it called "differential effect." The Federal Reserve Board, Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit, August 2007, https://www.federalreserve.gov/boarddocs/rptcongress/creditscore/. But we are not aware of similar studies for alternative data sources.

mortgage lenders do not collect protected demographic information from borrowers in order to test for fair lending compliance. As part of its work, this Task Force should encourage regulators to develop new guidance and set high standards for credit model development, validation, and testing. It is likely that many fringe alternative data sources will raise serious fair lending issues, particularly because there are likely to be "less discriminatory alternatives" based on conventional alternative data.[22]

The Task Force should also encourage regulators to set and enforce a high bar for model interpretability. It is critical — both as a matter of legal compliance and of policy — that industry data scientists are able to understand what their models are doing and why, and that consumers are able to understand why they do or do not qualify for credit.[23] Today, both the FCRA and Equal Credit Opportunity Act (ECOA) require a certain level of interpretability in credit scoring, but it is not clear how these provisions should be applied for more complex models, especially those relying on diverse kinds of alternative data.[24]

Conclusion

The Task Force should focus its efforts on encouraging the use of alternative data that is closely related to loan performance, has an understandable relationship to an individual applicant's creditworthiness, and has been evaluated for compliance with antidiscrimination law.[25] In addition, it should ensure that alternative data is used only for credit underwriting purposes, consistent with the FCRA and other key consumer protections. If these conditions are met, millions of historically underserved consumers stand to benefit.

---

[22] For an extended discussion *see* Robinson Meyer, "Could a Bank Deny Your Loan Based on Your Facebook Friends?" The Atlantic, September 25, 2015, https://www.theatlantic.com/technology/archive/2015/09/facebooks-new-patent-and-digital-redlining/407287/.

[23] *See, e.g.,* Rich Caruana, Intelligible Machine Learning for Critical Applications Such As Health Care, https://aaas.confex.com/aaas/2017/webprogram/Paper19142.html.

[24] 15 U.S.C. § 1681g(f)(1)(C) and 12 CFR § 1002.9(b)(2), respectively.

[25] In 1997, the Office of the Comptroller of the Currency (OCC) issued guidance stating that it would conclude that a variable used in a credit scoring model "is justified by business necessity and does not warrant further scrutiny if the variable is statistically related to loan performance, and has an understandable relationship to an individual applicant's creditworthiness." Office of the Comptroller of the Currency, Credit Scoring Models, OCC Bull. No. 97-24, app. 11, May 20, 1997, *available at* http://www.occ.gov/news-issuances/bulletins/1997/bulletin-1997-24.htm. See also National Consumer Law Center, Credit Discrimination 137 n.116 (6th ed. 2013) (stating that the OCC's guidance "may be indicative of how other federal regulators will view this issue."). Although this guidance was originally promulgated in the context of fair lending compliance, we believe it is a useful starting point when considering different types of alternative data.