

EXAMINING THE USE OF ALTERNATIVE
DATA IN UNDERWRITING AND CREDIT
SCORING TO EXPAND ACCESS TO CREDIT

HEARING
BEFORE THE
TASK FORCE ON FINANCIAL TECHNOLOGY
OF THE
COMMITTEE ON FINANCIAL SERVICES
U.S. HOUSE OF REPRESENTATIVES
ONE HUNDRED SIXTEENTH CONGRESS
FIRST SESSION

JULY 25, 2019

Printed for the use of the Committee on Financial Services

Serial No. 116-42



U.S. GOVERNMENT PUBLISHING OFFICE

40-160 PDF

WASHINGTON : 2020

HOUSE COMMITTEE ON FINANCIAL SERVICES

MAXINE WATERS, California, *Chairwoman*

CAROLYN B. MALONEY, New York	PATRICK McHENRY, North Carolina,
NYDIA M. VELAZQUEZ, New York	<i>Ranking Member</i>
BRAD SHERMAN, California	PETER T. KING, New York
GREGORY W. MEEKS, New York	FRANK D. LUCAS, Oklahoma
WM. LACY CLAY, Missouri	BILL POSEY, Florida
DAVID SCOTT, Georgia	BLAINE LUETKEMEYER, Missouri
AL GREEN, Texas	BILL HUIZENGA, Michigan
EMANUEL CLEAVER, Missouri	SEAN P. DUFFY, Wisconsin
ED PERLMUTTER, Colorado	STEVE STIVERS, Ohio
JIM A. HIMES, Connecticut	ANN WAGNER, Missouri
BILL FOSTER, Illinois	ANDY BARR, Kentucky
JOYCE BEATTY, Ohio	SCOTT TIPTON, Colorado
DENNY HECK, Washington	ROGER WILLIAMS, Texas
JUAN VARGAS, California	FRENCH HILL, Arkansas
JOSH GOTTHEIMER, New Jersey	TOM EMMER, Minnesota
VICENTE GONZALEZ, Texas	LEE M. ZELDIN, New York
AL LAWSON, Florida	BARRY LOUDERMILK, Georgia
MICHAEL SAN NICOLAS, Guam	ALEXANDER X. MOONEY, West Virginia
RASHIDA TLAIB, Michigan	WARREN DAVIDSON, Ohio
KATIE PORTER, California	TED BUDD, North Carolina
CINDY AXNE, Iowa	DAVID KUSTOFF, Tennessee
SEAN CASTEN, Illinois	TREY HOLLINGSWORTH, Indiana
AYANNA PRESSLEY, Massachusetts	ANTHONY GONZALEZ, Ohio
BEN McADAMS, Utah	JOHN ROSE, Tennessee
ALEXANDRIA OCASIO-CORTEZ, New York	BRYAN STEIL, Wisconsin
JENNIFER WEXTON, Virginia	LANCE GOODEN, Texas
STEPHEN F. LYNCH, Massachusetts	DENVER RIGGLEMAN, Virginia
TULSI GABBARD, Hawaii	
ALMA ADAMS, North Carolina	
MADELEINE DEAN, Pennsylvania	
JESUS "CHUY" GARCIA, Illinois	
SYLVIA GARCIA, Texas	
DEAN PHILLIPS, Minnesota	

CHARLA OUERTATANI, *Staff Director*

TASK FORCE ON FINANCIAL TECHNOLOGY

STEPHEN F. LYNCH, Massachusetts, *Chairman*

DAVID SCOTT, Georgia
JOSH GOTTHEIMER, New Jersey
AL LAWSON, Florida
CINDY AXNE, Iowa
BEN McADAMS, Utah
JENNIFER WEXTON, Virginia

FRENCH HILL, Arkansas, *Ranking Member*
BLAINE LUETKEMEYER, Missouri
TOM EMMER, Minnesota
WARREN DAVIDSON, Ohio
BRYAN STEIL, Wisconsin

CONTENTS

	Page
Hearing held on:	
July 25, 2019	1
Appendix:	
July 25, 2019	41

WITNESSES

THURSDAY, JULY 25, 2019

Evans, Lawrance L., Managing Director, Financial Markets and Community Investment, U.S. Government Accountability Office (GAO)	10
Girouard, Dave, CEO and Co-Founder, Upstart Network, Inc.	11
Johnson, Kristin N., McGlinchey Stafford Professor of Law, Tulane University Law School	8
Rieke, Aaron, Managing Director, Upturn	6
Wu, Chi Chi, Staff Attorney, National Consumer Law Center (NCLC)	5

APPENDIX

Prepared statements:	
Evans, Lawrance L.	42
Girouard, Dave,	54
Johnson, Kristin N.	57
Rieke, Aaron	74
Wu, Chi Chi	80

ADDITIONAL MATERIAL SUBMITTED FOR THE RECORD

Lynch, Hon. Stephen F.:	
Written statement of the Cato Institute's Center for Monetary and Financial Alternatives	96
Domino: A Blog About Student Debt	99
Written statement of the Financial Data and Technology Association of North America (FDATA North America)	102
Written statement of FICO	105
Written statement of VantageScore	107
McHenry, Hon. Patrick:	
Written statement of the Cato Institute's Center for Monetary and Financial Alternatives	111

EXAMINING THE USE OF ALTERNATIVE DATA IN UNDERWRITING AND CREDIT SCORING TO EXPAND ACCESS TO CREDIT

Thursday, July 25, 2019

U.S. HOUSE OF REPRESENTATIVES,
TASK FORCE ON FINANCIAL TECHNOLOGY,
COMMITTEE ON FINANCIAL SERVICES,
Washington, D.C.

The task force met, pursuant to notice, at 10:02 a.m., in room 2128, Rayburn House Office Building, Hon. Stephen F. Lynch [chairman of the task force] presiding.

Members present: Representatives Lynch, Scott, Gottheimer, Lawson, Axne, McAdams, Wexton; Hill, Luetkemeyer, Emmer, Davidson, and Steil.

Ex officio present: Representative McHenry.

Also present: Representatives Green, Himes, Porter; Gonzalez of Ohio, and Hollingsworth.

Chairman LYNCH. Good morning. The Task Force on Financial Technology will come to order.

Without objection, the Chair is authorized to declare a recess of the task force at any time.

Also, without objection, members of the full Financial Services Committee who are not members of this task force are authorized to participate in today's hearing consistent with the committee's practice.

Today's hearing is entitled, "Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit." I now recognize myself for 4 minutes to give an opening statement.

I want to thank everyone for being here at our second Financial Technology Task Force hearing. Today's hearing will focus on the use of alternative data, the financial and nonfinancial data that is not traditionally used by national consumer reporting agencies in credit underwriting. With an estimated 26 million consumers lacking in any credit history, and another 19 million with an outdated or short credit history, lenders have looked to other means of assessing the creditworthiness of applicants.

As a result, alternative data has become a hot topic. It has the potential to expand credit access but also raises concerns over the nature and sources of its data points. There is also significant regulatory uncertainty surrounding its use. Today, we will hear testimony and discuss questions on all of these issues. The promise of fintech lending has been to lower costs and bring new consumers

into the market. This promise has been fueled by data points outside of the traditional factors used by underwriters like payment history and credit utilization. Today, lenders use an array of financial and nonfinancial data in their decision-making. Some factors, such as utility bill or rent payments, resemble traditional factors. Others, such as living in public housing, who your friends are, and what their credit scores are, your ZIP Code, your reading choices, educational attainment, educational institutions, and driving habits or online shopping habits, are a significant departure from traditional factors.

We know that Facebook has up to 52,000 data points on each of its 2.7 billion daily users, and they sell access to that data to its advertisers. Use of this and other data can potentially help 45 million Americans who might have trouble accessing credit with traditional factors alone. Take for instance, a 28-year-old woman in a modest-paying job, or maybe with 2 very modest paying jobs, who has never had a credit card or taken out a personal loan or mortgage loan. She might find herself denied access to credit based on traditional factors, even though she is working 12 hours a day. But a lender using alternative data might take into account that she went to a reputable school, had a job with a reputable employer, and always pays her rent and utility bills on time. In that case, they might approve her application for credit. It is very likely we have one or two staffers working here on Capitol Hill who fit that exact description.

However, it is not hard to imagine a similar scenario with much different results. Say, a young man with a decent but short credit history might be right on the cusp of being deemed creditworthy by traditional factors. However, a lender using alternative data sees in his rental history that he moves frequently, moves around a lot. In the last few years, he has had several domiciles. They may also see he doesn't have a college degree and that his Facebook friends have below average credit scores. So, they deny him access to credit. Unfortunately, this probably describes a number of our military personnel as they repeatedly move domiciles as a result of multiple redeployments during their careers.

Without question, there are instances when using alternative data in credit underwriting has potential positive impacts. However, right now, oversight of its use is either highly fragmented or completely nonexistent, leading to uncertainty for lenders and potential harm for consumers. That is why we are here today, to better understand how to harness the benefits and mitigate the harms of using alternative data.

I look forward to the testimony of our witnesses and the discussion of our Members. With that, I now recognize the ranking member of the task force, the gentleman from Arkansas, Mr. Hill, for 5 minutes for an opening statement.

Mr. HILL. I thank the chairman. I appreciate you convening this hearing and I appreciate our witnesses appearing today. We are grateful for your advice and counsel today. Analyzing the use of alternative data in the marketplace lending industry is an important sector within our broader study of the fintech ecosystem. I am pleased that we were able to bring everybody together and do a deeper dive on this topic.

Marketplace or fintech lenders are categorized through their digital or online focus and have recently emerged and grown quite a bit over the last decade. According to S&P Global, marketplace lending grew by 30 percent in 2017. They provide unsecured credit to individuals and working capital to small businesses. They have unique funding models with financing provided by investors, credit facilities, securitization, and, of course, balance sheet cash.

Typically, these lenders currently lend through two primary models: a bank partnership model, in which a bank originates the loan, which is generally sourced and served by the marketplace lender; or a direct lender model, in which a marketplace lender acquires the applicable regulatory licensing in all of the States of our country.

To help determine a borrower's creditworthiness, marketplace lenders often use some form of alternative data, hence the topic today. Traditional lenders typically have used FICO scores, 3 years of tax returns, payment history for credit cards, mortgages, or student loans in order to establish a risk profile for their borrowers.

However, marketplace lenders robustly combine FICO scores with alternative data points to better gauge a borrower's character and economic situation. Examples of these data points include education level, employment status, utility and rent payments, et cetera. Analyzing these data points has the potential to widen the universe of borrowers and provide greater access to affordable credit.

Importantly, a report by TransUnion outlined that lenders that utilized alternative data were able to lend to an additional 66 percent of borrowers in current markets and 56 percent in new markets.

Today, we will explore concerns about how alternative data can best comply with critical fair lending requirements, which will be discussed in more depth. However, I do want to remind my colleagues that we don't want overregulation to stifle innovation and prevent the American consumer from now being able to access affordable credit through this new methodology.

As to compliance obligations, obviously, I want to highlight some of the ongoing issues that have been evolving within the marketplace lending industry. The Treasury report—which I regularly reference in these hearings—released a year ago now, provides a comprehensive review of the fintech sector. It has a robust analysis of this industry. It generally favors innovation, but identified certain important policies that need to be highlighted and discussed today, such as codify the valid-when-made doctrine, codify the role of the bank as the true lender of a loan that is made, allowing the testing of new credit models and data sources for financial institutions, and think through this issue of special charters or harmonization of this process across our States.

The report also mentioned the third-party lender guidance. I know the FDIC and the OCC have been focused on this due to the rise of marketplace lenders and strong bank partnerships. As a former community banker, I well understand the compliance responsibilities around vendor partner, due diligence, onboarding of new partnerships, and board of director reviews.

Also, as a result, as a banker, I understand the importance of banks maintaining a robust level of safety and soundness and constantly facing changing technology but assuring a vigor in compliance on both data security and privacy.

I look forward to hearing the thoughts of the panel today, and over the years, I think this is going to be a fascinating way for Mr. Lynch and I to make recommendations to our full Financial Services Committee on how we can broaden marketplace lending. So, with that, I would like to yield the balance of my time to my friend, the ranking member of the Full Committee, Mr. McHenry of North Carolina.

Mr. MCHENRY. Thank you. And, look, technology is creating new pathways for more consumers to access the financial system. That is a good thing. We are talking about people who are otherwise credit invisible or lack enough credit history to finance things like a mortgage, credit cards, or other loans. Alternative data draws on nontraditional sources of financial history, including bill payment history in areas like mobile phones, that are now essential ingredients, with 81 percent of Americans owning a smartphone at this point or using a smartphone, and rent. And by harvesting this type of data about the consumer, lenders have a more holistic picture about the consumer to whom they are lending.

Yet, this new era is not without its challenges. We need to ensure that alternative data remains nondiscriminatory and that consumer data and privacy are protected. So, it is our job to ensure responsible innovation continues to be a driving force of the American economy, but in particular, in financial services. I yield back.

Chairman LYNCH. The gentleman yields back.

The Chair now recognizes the gentleman from Georgia, Mr. Scott, for 1 minute for an opening statement.

Mr. SCOTT. Thank you very much, Chairman Lynch, and let me commend you and Mr. Hill for providing this excellent bipartisan leadership on what I refer to as the thrilling new frontier. We are in a situation now where technology is moving at such a rapid pace, and where we need to look at where we need to adjust the sails and make sure everybody has an opportunity to be able to participate in this. And, of course, as we look at this scoring technology, we look at how it is impacting our financial system. There is no group that needs our help more than the 26 million Americans who have no credit history. There are also 19 million Americans who have a very limited credit history. And when you put the totality of the unbanked and the underbanked in there, we can see that we must not leave these parts of our population behind.

So, I am looking forward to this, to making sure that we determine effectively how data is used in lending decisions and credit scoring, so all the American people can participate in this glorious new frontier.

Thank you, Mr. Chairman.

Chairman LYNCH. I thank the gentleman. On behalf of this bipartisan task force, I want to welcome our distinguished panel. I would like to welcome the testimony of: Chi Chi Wu, a staff attorney with the National Consumer Law Center, based in Boston, my hometown; Aaron Rieke, managing director at Upturn, which is a nonprofit focused on promoting equity and digital technology

through research and advocacy; Kristin Johnson, McGlinchey Stafford Professor of Law at Tulane University Law School; Lawrance Evans, Managing Director of Financial Markets and Community Investment at the Government Accountability Office; and Dave Girouard, founder and CEO at Upstart, which is a fintech lender focused on direct-to-consumer loans.

Witnesses are reminded that your oral testimony will be limited to 5 minutes. And without objection, your written statements will be made a part of the record.

Ms. Wu, you are now recognized for 5 minutes to give an oral presentation of your testimony.

**STATEMENT OF CHI CHI WU, STAFF ATTORNEY, NATIONAL
CONSUMER LAW CENTER (NCLC)**

Ms. WU. Mr. Chairman, Ranking Member Hill, and members of the task force, thank you for inviting me to testify today. I am testifying on behalf of the low-income clients of the National Consumer Law Center. We have heard several times today that there are tens of millions of consumers who are credit invisible. The topic of this hearing, alternative data, is often promoted as the solution. The thing is, alternative data includes lots of different types of data used in lots of different ways. Some types of data and uses can be helpful; others can hurt. As we say, the devil is in the details.

The number one consideration for alternative data should be consumer choice. That should be the touchstone for all data collection. Now, we have heard with respect to the Equifax data breach a repeated complaint: Hey, none of us gave Equifax permission to collect our data.

Let's get this right with respect to alternative data. Let's make sure it is the consumer's choice, that consumers make knowing and affirmative decisions to allow the use of this data, and the data is only used in the ways that consumers give permission for and expect. Another consideration for alternative data is whether it is used to create second-chance scores for just credit-invisible consumers or whether it is dumped wholesale into traditional credit reports where it might damage the records of consumers who already have a score. We want to give credit-invisible consumers a chance to be seen without hurting any of the nearly 200 million consumers who are already visible.

As for types of data, bank account transaction data has shown a lot of promise, but it is also a juicy target. Debt collectors would love to get ahold of it. And bank account data can include sensitive information, such as where a consumer shops. There should be appropriate guardrails for sharing bank account data.

Rent payment information is another type of data looks promising, specifically when no additional late payments are reported. But we don't want to penalize tenants who invoke their rights to withhold rent over poor conditions.

Payday loan information, in contrast, is probably harmful. It is designed to lead to a cycle of debt, and just reporting it can hurt a consumer. And it is probably not necessary because most payday borrowers actually have credit records.

Gas and electric utility data can be potentially harmful if added in the wrong way. If reported monthly without giving consumers a choice, it has the potential to hurt tens of millions of low-income consumers by adding new reports of 30- or 60-day late payments. In contrast, efforts to include utility data on a voluntary basis could be useful, and new voluntary products show there is no need for utility credit reporting where the consumer has no choice.

And then, of course, there is Big Data—things like social media profiles, web browsing history, and behavioral data. There are a lot of unanswered questions about the predictiveness and the accuracy of Big Data. Some of it is also troubling because it strongly reinforces inequality. For example, education, that is, what kind of degree a consumer has, is highly correlated with the income and education of one's parents. And using social media profiles, particularly friend networks, raises concerns about racial disparities, given who most people's friends and families are likely to be.

Speaking of racial disparities, we know there are tremendous racial disparities with respect to traditional credit scores. It is the result of centuries of slavery and discrimination which led to the huge racial wealth gap. Alternative financial data is also likely to have racial disparities for the same reasons. The critical question is whether the alternative data or algorithms lessen or increase racial disparities and whether it is more predictive or less than traditional models.

These two questions are closely tied to the test for disparate impact under the Equal Credit Opportunity Act. If the alternative data is less predictive, there is less of a business justification for it, under the disparate impact test. On the other hand, if it creates less of a racial disparity, it could be a less discriminatory alternative than traditional scoring.

In terms of regulation, all third-party alternative data used for credit should be considered a consumer report under the Fair Credit Reporting Act (FCRA). Unfortunately, several courts of appeals haven't respected the plain language of the FCRA and its broad coverage. We urge Congress to reaffirm this broad coverage, because the FCRA has critical protections. One of the key issues with alternative data is accuracy, the FCRA addresses accuracy, and it gives consumers the right to dispute errors.

The FCRA, as well as the ECOA, also requires notices for the purpose of transparency, requiring lenders to disclose the source and type of information so consumers aren't left in the dark as to the reasons for credit decisions. Having black boxes to evaluate creditworthiness should be a thing of the past. I thank you for the opportunity to testify and I look forward to your questions.

[The prepared statement of Ms. Wu can be found on page 80 of the appendix.]

Chairman LYNCH. Very good, thank you.

Mr. Rieke, you are now recognized for 5 minutes.

**STATEMENT OF AARON RIEKE, MANAGING DIRECTOR,
UPTURN**

Mr. RIEKE. Chairman Lynch, Ranking Member Hill, and distinguished members of the task force, thank you for the opportunity to testify today. We are here because approximately 45 million

Americans do not have access to credit because there is a lack of quality data with which to underwrite them. Alternative data can certainly help. I want to echo Ms. Wu and say that the devil is in the details, and to suggest that we are really talking about two categories of data here: conventional data; and fringe data.

Conventional data consists of things like various payment histories, bank account balances, information about an individual person's financial capacity. Fringe data consists of things like social media data, information that may be correlated with this financial capacity but is much further removed.

Conventional data is promising; fringe data raises concerns. To understand why, think about traditional FICO credit scores. These credit scores are not conceptually complex. Most of their predictive value comes from people's payment histories. That is really the number one factor in the recipe of FICO scores. The logic is simple. If a consumer is keeping up with their current financial obligations, it is reasonable to predict that they can take on new financial obligations. As it turns out, the same basic logic applies to many kinds of conventional data. The best available evidence suggests that bill payment histories are similarly predictive and can help otherwise unscorable consumers access credit.

Another example, cash-flow data obtained from a consumer's bank account with their express permission, can provide an immediate high-quality picture of that person's ability to repay a loan, even without a credit bureau being involved at all. That is conventional data.

The story gets murkier when we talk about fringe data. Expansive data sets about people's social connections, the kinds of websites they visit, where they shop, and how they talk do not have the same simple, intuitive connection to each individual's ability to repay a loan. These can yield blunt stereotypes that might be predictive, but for the wrong reasons.

Let me offer you an analog analogy. Imagine I offered to build you a credit-scoring model that relied on a person's ZIP Code. That should feel intuitively wrong. I want to unpack why. First, we know that geography reflects deep-seated social inequities. The result would almost certainly be textbook disparate impact.

Second, judging from ZIP Codes would paint with too broad a brush. It would do little to help many of the unscorable consumers we seek to help most who already live in low-income neighborhoods. Latching on to traditional markers of wealth and privilege aren't going to get us to where we want to be. My point is that thousands of behavioral data points thrown into a complicated, machine-learning, artificial intelligence model can actually act and behave just like a ZIP Code. In the absence of rigorous public scrutiny, we should be skeptical of fringe data.

I want to note, because Facebook was brought up in opening remarks, that Facebook has for a number of years had a policy that prohibits third parties from using Facebook users' data for any kind of eligibility purpose. So, if you see a start-up company touting their use of Facebook data, ask them why they are violating Facebook's policies. That may not be the case forever, but I think today that indicates that we are not ready to embrace this new data set. In short, this task force should focus its efforts on encour-

aging the use of alternative data that is closely related to loan performance, has an understandable relationship with an individual applicant's creditworthiness, and has been evaluated for compliance with anti-discrimination laws.

Fortunately, this is all doable. More collection and use of alternative data makes the Fair Credit Reporting Act and the Equal Credit Opportunity Act more important than ever before. I would urge Congress to ensure that new kinds of alternative data are only used for credit underwriting, where we have researched and understood their role, and not for things like employment and insurance.

Finally, as you are all aware, thanks to the advocacy of Ms. Wu and her colleagues, any new policies around alternative data must respect important State and local consumer protections. Thank you again for the opportunity to testify, and I welcome your questions.

[The prepared statement of Mr. Rieke can be found on page 74 of the appendix.]

Chairman LYNCH. I thank the gentleman.

Ms. Johnson, you are now recognized for 5 minutes for a summary of your testimony.

STATEMENT OF KRISTIN N. JOHNSON, MCGLINCHEY STAFFORD PROFESSOR OF LAW, TULANE UNIVERSITY LAW SCHOOL

Ms. JOHNSON. Good morning, Chairman Lynch, Ranking Member Hill, Ranking Member McHenry, members of the committee, and members of the task force. Thank you for inviting me to participate in this hearing to discuss the use of alternative data in credit underwriting and credit scoring. I am a professor of law and associate dean of faculty research at Tulane University Law School, but I have previously worn other hats. I was an analyst at Goldman Sachs, a vice president and associate general counsel at JPMorgan, and an associate at a New York law firm with a globally recognized transactional practice. During my tenure in financial services and as an academic, I have learned a few things about financial markets, including the lesson that credit is a critical resource.

Individuals and families increasingly rely on credit to finance household purchases and overcome significant unanticipated expenses. Without access to credit on fair and reasonable terms, it can be extraordinarily expensive to be poor. For families with fragile financial circumstances, credit may serve as a lifeline, enabling consumers to meet short-term debt obligations and to pay for education, transportation, housing, medicine, childcare, and even food.

Two critical developments create promise for the 26 million Americans referenced earlier as credit invisible, those without credit histories, and the 19 million Americans who have thin, impaired, or stale credit histories described as unscorable. First, the birth of Big Data. The collection, storage, and analysis of vast volumes of consumer data fuels artificial intelligence or automated decision-making platforms. Similar to the proliferation of AI in health care, employment, criminal law, surveillance, and communications, the rise of AI in finance monetizes consumer data. Consumers' web browsing, click-stream data, and social media networking, which we could describe as consumers' digital interface, is matched with

or paired with consumers' financial transactions, checking and saving account cash flows, and credit and debit card transactions, fueling data mining and engendering a new set of behavioral criteria we can describe as alternative data.

While fintech firms integrating alternative data offer great promise, it is very much worth noting that this new species of financial market intermediaries also presents great concerns. In my limited time this morning, I note three challenges that arise when we integrate and endeavor to regulate alternative data.

First, alternative data may, as mentioned earlier, disadvantage vulnerable, marginalized consumers, particularly those who are members of legally protected classes. Under the behavioral scoring model, your friends on Facebook, the people in the pictures you post on Instagram, and those you chat with on WhatsApp—I am happy to deconstruct that later for those unfamiliar—may signal more than whether or not you have street cred. These connections may determine the interest rate on your next mortgage.

It is not yet clear how these new sources of data will impact credit invisibles and unscorable groups often disproportionately comprised of women and people of color. Unsavory lending practices, detestable marketing tactics, and usurious interest rates have too often plagued these marginalized consumers.

Second, learning algorithms evaluate facially neutral, alternative data, yet may result in variables that function as proxies for protected traits or result in decisions that may have a disparate impact on members of legally protected classes.

Consider, for example, Amazon's recent experiment with an algorithm tasked with reviewing resumes for a software programmer position. Armed with the resumes of previous hires and general instructions regarding qualifications, the algorithm went rogue. Because previous hires were predominantly men, the algorithm began to discount references to women, including references to women's chess club captain or all-women's colleges. Unknowingly, the algorithm replicated historic discriminatory hiring biases. In credit decisions, these results may be actional, as noted earlier, under the Equal Credit Opportunity Act and fair lending and fair housing regulations.

Finally, alternative data raises concerns regarding consumer privacy and cybersecurity concerns. Beyond Equifax's settlement this week, there is more breaking cybersecurity news. A 20-year-old computer programmer successfully launched a cyber attack against another nation's national revenue agency, signaling that it is imperative to ensure that any entities that collect, store, and transfer consumer data have developed sufficient security mechanisms.

CRAs may also struggle with respect to the obligation to describe and explain adverse credit decisions. Because of the inscrutable nature of learning algorithms, they are non-intuitive, opaque, and their operations are not often easily explained.

Finally, in my written testimony I note as well that there is an even newer class of emerging financial intermediaries within the fintech ecosphere, or ecosystem—blockchain-based CRAs. I reference in my written testimony Bloom, one example of a blockchain-based credit reporting agency or an entity that will op-

erate in a manner similar to a credit reporting agency, that is also presumably to rely on alternative data.

For these reasons, I encourage and urge Congress to think carefully about comprehensive legislation outlining the appropriate uses for alternative data and data governance, storage, transfer, and cybersecurity protections, as well as enforcement of anti-discrimination norms.

[The prepared statement of Ms. Johnson can be found on page 57 of the appendix.]

Chairman LYNCH. Mr. Evans, you are now recognized to give us a 5-minute summary of your testimony.

**STATEMENT OF LAWRENCE L. EVANS, MANAGING DIRECTOR,
FINANCIAL MARKETS AND COMMUNITY INVESTMENT, U.S.
GOVERNMENT ACCOUNTABILITY OFFICE (GAO)**

Mr. EVANS. Thank you very much, Chairman Lynch. I am pleased to appear before you, Ranking Member Hill, Ranking Member McHenry, and the members of the task force to discuss the use of alternative data in underwriting. My testimony is largely based on our December 2018 report, which covered several fintech lending issues. The problem with the current credit-granting ecosystem has been well-articulated, namely its limits in its ability to reach certain borrowers. We know that alternative data provides an opportunity to improve the status quo by expanding access to credit, improving prices, speeding up decision-making, and preventing fraud, but it is also important to know that some of what we refer to as alternative data is not new.

However, the types of alternative data available have expanded significantly due to the ability to secure large volumes of consumer and behavioral information, including data on consumer spending and shopping habits, internet browsing history, online social media networks, educational affiliations, and other factors that may not have a clear nexus with creditworthiness.

In combination with analytic techniques like machine-learning, these factors provide predictive power for fintech companies looking to enhance their ability to determine who is eligible for credit. But alternative data is not a panacea. Depending on the specifics of these data and the analytical techniques used to extract information from them, these innovative approaches can bring significant risk. One of the major concerns is that usage of that data may produce lending outcomes that result in disparate impacts or violations of fair lending laws, unintentionally in some cases.

For example, according to a Federal Reserve newsletter, it has been reported that some lenders consider whether a person's online social network includes people with poor credit histories, which can raise concerns about discrimination against those living in disadvantaged areas.

Another concern is that there may be a lack of transparency about what alternative data are being used and how they ultimately factor into credit decisions. This potential opacity could raise issues, not only for consumers, but for fintech firms themselves looking to comply with fair lending requirements. It may also be unclear whether a borrower has the ability to dispute the accuracy of the information used.

The great challenge ahead is to manage the risk-reward balance of innovation and ensure our experience with alternative data does not mimic our experience with alternative mortgage products leading up to the financial crisis. To better ensure the risks are managed without stifling innovation, which is extremely important, policymakers and regulators will need to sort through a number of different tradeoffs and considerations.

In the meantime, implementing key recommendations that GAO has offered to regulators would assist them in addressing some important deficiencies as we see them. Fintech lenders and their banking partners we spoke to indicated they face challenges due to regulatory uncertainty about the appropriate use of alternative data. Representatives of one bank said that a fintech partner's use of alternative data may be attractive from an innovation and business perspective, but the bank would likely hesitate to use this data due to regulatory uncertainty. While Federal agencies monitor the use of alternative data, they have not provided firms with the types of communication that they need to really think through the appropriate use of this data in the underwriting process.

We believe coordinated guidance from the regulators may better position fintech lenders and their bank partners to responsibly use alternative data. In our prior work, we have also recommended that agencies formally evaluate the feasibility and benefits of adopting knowledge-building initiatives. We believe these initiatives will help firms understand the applicable regulations, improve regulators' knowledge of fintech products, and facilitate interactions between all parties.

Chairman Lynch, Ranking Member Hill, Ranking Member McHenry, and members of the task force, this concludes my opening statement. I look forward to any questions you may have.

[The prepared statement of Mr. Evans can be found on page 42 of the appendix.]

Chairman LYNCH. Thank you, Mr. Evans.

Mr. Girouard, you are now recognized for 5 minutes. Welcome.

**STATEMENT OF DAVE GIROUARD, CEO AND CO-FOUNDER,
UPSTART NETWORK, INC.**

Mr. GIROUARD. Chairman Lynch, Ranking Member Hill, Ranking Member McHenry, and members of the Task Force on Financial Technology, thank you for the opportunity to participate in today's conversation. My name is Dave Girouard, and I am co-founder and CEO of Upstart, which is a leading artificial intelligence lending platform. I founded Upstart more than 7 years ago, in order to improve access to affordable credit. In the last 5 years, almost \$4 billion in bank quality consumer loans have been originated on our platform, using a model that combines alternative data with AI and machine-learning algorithms to determine a borrower's credit-worthiness.

Concerns about fairness in algorithmic lending, particularly in the use of alternative data, are well-founded. As a company focused entirely on reducing the price of credit for the American consumer, fairness is an issue we care about deeply. In the early days at Upstart, we conducted a retroactive study with a large credit bureau, and we uncovered a jarring pair of statistics: Just 45 percent of

Americans have access to bank quality credit, yet 83 percent of Americans have never actually defaulted on a loan. This is not what we would call fair lending.

The FICO score was introduced in 1989 and has since become the default way that banks judge a loan applicant, but in reality, FICO is extremely limited in its ability to predict credit performance because it is narrow in scope and inherently backward-looking.

At Upstart, we decided to use modern technology and data science to find more ways to prove that consumers are indeed creditworthy, to bridge that 45 percent versus 83 percent gap. We believe that consumers are more than their credit scores, and by going beyond the FICO score and including a wide variety of other information such as a consumer's employment history and educational background, we have built a significantly more accurate credit model.

While most people believe a better credit model means saying no to more applicants, the truth is just the opposite. Because Upstart's model is more accurate, we have significantly higher approval rates and lower interest rates than a traditional model.

But we also understood that consumer protection laws weren't to be taken lightly. Thus, we proactively met with the appropriate regulator, the Consumer Financial Protection Bureau (CFPB), before launching our lending program. After several years of good-faith efforts between Upstart and the CFPB to determine the proper way to measure bias, we demonstrated that our AI-driven model doesn't result in unlawful disparate impact against protected classes of consumers.

Because AI models change and improve over time, we developed automated tests with the regulators' input, in order to report on the impact of our credit decisions across underserved groups on a quarterly basis. We have been providing this information to the CFPB for the last 18 months. Moreover, we were able to report to the CFPB that our AI-based system improved access to affordable credit; specifically, our model approves 27 percent more consumers and lowers interest rates by 3.57 percentage points compared to a traditional lending model. For near-prime consumers in the 620 to 660 FICO range, our model approves 95 percent more consumers and reduces interest rates by 5.42 percentage points compared to a traditional model. And, most importantly, Upstart's model provides higher approval rates and lower interest rates for every traditionally underserved demographic. That is the type of consumer benefit we should all get excited about.

In September 2017, Upstart received the first ever no-action letter from the CFPB, recognizing that Upstart's platform improves access to affordable credit without introducing unlawful bias. The concern that use of alternative data and algorithmic decisioning can replicate or even amplify human bias in lending is well-founded. However, in Upstart's experience, the fair-lending laws enacted in the 1970s and the substance of fair-lending enforcement, that is, monitoring and testing the impact on actual consumers who apply for loans, translates very well to the AI-driven world of today.

But in reality, the path we walked at Upstart is insufficient to create a robust and competitive market that will maximize finan-

cial inclusion and credit access. In our early days at Upstart, we couldn't know for certain whether our model would be biased. It wasn't until loans were originated that we were able to demonstrate that our platform was fair. As an early-stage startup, this was a risk worth taking, but it is not a risk a large bank would have considered.

If broader and deeper financial inclusion among American consumers is important to this committee, it is worth considering rule-making or legislation that will provide some type of limited sandbox for model development and testing. By combining regulatory support with model innovation, with rigorous and standardized testing, we can ensure that we don't forego the clear and obvious benefits that AI and alternative data can offer to the American consumer. Thank you.

[The prepared statement of Mr. Girouard can be found on page 54 of the appendix.]

Chairman LYNCH. Thank you. I now yield myself 5 minutes for questioning. Thank you all. This is a great group. One of the nicer things about this task force is that it is bipartisan, and we are here for the same purpose: We are looking for guidance. We have an assortment of issues that we are confronting. Obviously, the banking industry is transformed, I think, because of technology so that we have an old banking culture that is very much rule-based, and it seems to be merging or morphing into this sort of tech hybrid where you move fast and break things. And so, there is that clash of cultures.

But I can generally group our concerns into four areas. One is the whole issue of companies vacuuming up this personal data, this behavioral surplus, as Shoshana Zuboff describes in her book, "The Age of Surveillance Capitalism." And under what conditions do consumers have a choice in terms of what gets vacuumed up and what gets used in terms of the algorithms that are employed to judge their creditworthiness or on other matters. There is that whole permission aspect that Ms. Wu and Ms. Johnson both raised. Actually, all of you, I think, addressed that in some regard.

Then the data use, how that gets used, what data is permissible to use and what is not. Then, one of the concerns that this committee has is regarding the security of that data. We had Facebook in, and Mr. Marcus, who is heading up their Libra cryptocurrency project. And it is obvious from our history with Facebook, that Facebook does not do privacy well, and so we worry about that. If you look at the terms of service agreement, the one that is on your phone with Facebook, it is about 20 pages long. And if you look at it closely, it basically is the opposite of a privacy agreement. It basically gives Facebook the ability to gather all your data and then sell it to their advertisers. And if you don't agree, if you don't click, "I agree," you don't get Facebook. So, I am worried about fintechs using that same sort of adhesion contract to get people to surrender their data, in order to get the value of what Mr. Girouard has described, which is perhaps lower rates, better access to credit, all the benefits that might flow from one of the fintech lenders.

And then, lastly, we are struggling with how to hold people accountable with financial data. Should there be—I asked Mr. Marcus, but he wasn't forthcoming with an answer—I asked him,

I said, would you accept fiduciary liability for the mishandling of consumers' personal financial data because of the consequences that can occur because of that mishandling?

So, Ms. Wu, let's talk about, how do we get into this? How do we introduce this permission regime where people can—and, Ms. Johnson, I will go to you on this as well—how do we introduce this? Right now, it is a permissionless vacuuming up of data. How do we change the paradigm and the model from what we have now to a more rule-based, if you will, structure with some of the fintech that is emerging?

Ms. WU. Thank you for the question, Chairman Lynch, it is an excellent one. There is the sort of limited issue of alternative data for credit purposes, where we would urge that any legislation always be on an opt-in basis, that consumers have a choice, and that that choice be real and meaningful, that it not be in mice type of 20 pages of fine print that you mentioned. From a broader perspective of privacy in general, yes, we all should have more control over our own data, the right to opt in, opt out, or even have our data deleted.

Chairman LYNCH. Very good.

Ms. Johnson?

Ms. JOHNSON. I think that Ms. Wu's point is absolutely consistent with what our expectations ought to be. I think the challenges are two-part, one part technical, so I might defer to Mr. Girouard to respond as to how their model might address this very specific and technical point. But for AI to be effective, as I describe in my written testimony, there has to be a certain quantity or volume of observations available. They have to be uniform to a certain extent, and that facilitates the learning algorithm's ability to work through the data in a manner that is exceptionally efficient and reduces operating costs, thereby enabling fintech lending platforms to reduce the cost of borrowing for consumers.

One of the challenges I am very curious about how we will navigate is the extent to which we are rightly asking that consumer's consent prior to their data being used, and how we reconcile that with how machine-learning algorithms operate. So, I think there is a gap there that we have to have enough of a conversation about, to be successful in crafting regulation.

The other thing I just mentioned really quickly about consent is that the extent that the data is being gathered really may be the point of departure for some of our concerns. In many instances, consumers are completely unaware that the data is being gathered. And in some instances, they are voluntarily giving the data over for the better credit opportunities or reduced price credit opportunities, which is disconcerting, to be quite honest, because it suggests those who are most vulnerable might be exposed to—or exploited, in fact, by arrangements whereby they share the most intimate details of their financial lives or their personal lives for the purpose of getting better access to credit.

Chairman LYNCH. Thank you very much. I now yield to my friend, the gentleman from Arkansas, Mr. Hill, for 5 minutes.

Mr. HILL. Thank you, Mr. Chairman.

And, again, thanks to the panel. This is another really excellent panel that has been assembled for the task force work, and I think all of you bring a great perspective.

Certainly, this issue of customer choice is an important one, and we all are frustrated, I think, with moving away from passwords into a more robust authentication, which is critical to a digital world, critical to fintech being successful, whether you are working at the biggest bank in the country or a great startup. We need to get beyond “password1” and our name as authentication, and we have been talking about that a lot.

Secondly, this issue of, I own the data, I am the consumer, and I am allocating you some data for a project we are working on together, and so broadening that transparency in access to my data for the purpose of taking a decision that I want to have with an online partner. These are really important areas and thank you for bringing those up.

Mr. Girouard, I want to talk a little bit about your model and the alternative use or, as was described by Mr. Rieke, your expansion, I would say, of conventional data. I will ask if you use “fringe data” or not, as he defined it, but we will find out. But I am very impressed that you have been working 18 months with the CFPB, which is a beloved institution in Washington and certainly in this committee and, therefore, has imminent authority over that relationship. And congratulations for having a no-action letter. We think that is a great improvement for CFPB operations as an absolutely serious comment and a great way for them to demonstrate the ability for fintech innovation in a mini sandbox if you want to call that a derivative of that.

My first question is, the conventional data you expand beyond FICO, what is the nature of that in your business?

Mr. GIROUARD. Sure. I want to say first that the data we use in our models comes entirely from two sources: one, is a credit reporting agency; and two, is directly from the consumer themselves. So we aren’t “hoovering data in many places.” We don’t take data from Facebook, et cetera. What we do is include information—and I had mentioned a couple of them—somebody’s work history, where do they work, are they a nurse, are they a policeman, et cetera, their educational history, the degree of education obtained, their area of study. These are things that are unique to our model. We also look at some behavioral things when they interact with us, what sort of—how much—what size loan do they ask for, how did they find us, things of this nature. These all end up being helpful and predictive toward our model.

Mr. HILL. Do you consider that—of course, it is provided by the customer. They are seeking the loan, so they have granted you permission to do that. Are you also seeing their cash-flow data for a period of months by access to their bank account in making your determination?

Mr. GIROUARD. Today, that is not something we do. We do request and, with consent, get access to a bank account really for verification purposes and to avoid fraud and such. But it is, as of today, not part of our credit decisioning.

Mr. HILL. You talked about how you are doing that, and, of course, the CFPB is learning, too, and you keep some of your cred-

it, and your partner bank has some of your credit originated by you on their books, and then you securitize credit. So, for the loans that you keep and for the loans that are on the bank's books, of course, those are being reviewed by compliance officials for compliance with all fair lending laws and the like? Isn't that right?

Mr. GIROUARD. Sure. There are many layers of oversight and governance over what we do. The vast majority, almost all Upstart loans are originated through bank partners—some of which are FDIC-regulated, and some of which are OCC-regulated. So, we are beholden to all of them and go through very regular audits and such.

Mr. HILL. What is your view of what statutorily ought to change about the creation of a sandbox at our bank regulatory agencies? What does that mean to you? I see it in your testimony. You don't really explain what you mean by that. How do you define it?

Mr. GIROUARD. Our belief, as I said, is that the right way to handle regulation for alternative data, and the use of alternative data is actually to measure the outcome, to look at its impact on consumers and whether there is bias in the outcome. The challenge with that, and the way the world works today, is, you don't know until you originate the loans. So, you are taking on some risk that, during that period of evaluation of building and testing that model, you could be in violation of the law, of fair-lending laws. The sandbox concept is, how do you actually make progress there? How do you actually build a better model that is both more effective and more accurate, but also fair and unbiased without testing and moving? And the notion of a sandbox is to provide some freedom, not just for a startup like we were 5 years ago, but to a large financial institution, a bank, to do the same thing.

Mr. HILL. This is like a phase one or a phase two clinical trial in the drug research industry. How long do you think that would take and how much of a Big Data set would that be, in just your world of personal lending, do you think would be necessary to prove out a concept like that, analytically? 18 months?

Mr. GIROUARD. Yes. That is—

Mr. HILL. Do you look at it in time, or do you look at it in total data set, or both?

Mr. GIROUARD. It is a little of both.

Mr. HILL. Because you have to go through the economic cycle of these borrowers, to some degree, some seasoning of these borrowers.

Mr. GIROUARD. That is really about the efficacy question, meaning, does this model work well? But the fairness question actually is answered quite quickly because you know right away who you are approving and who you are not approving.

Mr. HILL. Thank you. I yield back, Mr. Chairman.

Chairman LYNCH. The Chair now recognizes the gentleman from Georgia, Mr. Scott, for 5 minutes.

Mr. SCOTT. Mr. Evans, let me start with you, because in your testimony you provide a very good survey of the literature of the potential benefits of alternative data, but you also mention the risks. First of all, I think it would be helpful if you gave us some examples. What are we talking about when we say alternative data? What would that be?

Mr. EVANS. This could range from data that we have had significant experience with, like on-time rental payments, mobile payments, and the like. But it could also be data that we glean from your digital footprint online or your browser history.

Mr. SCOTT. But these data points also must uphold the fair lending laws and standards that we have in place. I think the critical question is, how do we strike the balance? How do we strike the necessary balance, particularly given the innovative nature, the rapidity of our technology moving?

Mr. EVANS. Excellent question. And there are two things that I would point out from our body of work. One, we looked across the globe, and we looked at some of the innovative things other countries were doing, and they were things like the regulatory sandboxes, and innovation offices. We have to understand the technology, and the way to understand the technology is to engage. We have recommendations that are open to regulators to make sure they are carefully thinking through whether these innovation offices and other types of knowledge-sharing initiatives would be appropriate here in the United States.

Also, guidance is extremely important because it sets the rules of the road. It sets parameters. And if the fintech firms aren't getting that kind of guidance, they are not—

Mr. SCOTT. And do you think the regulators are living up to that? Do you think they are giving this guidance properly now?

Mr. EVANS. I would say no. There are certainly places where you can find good information from the Federal Reserve and others, but they haven't communicated this guidance in a written, formal way, so that people understand that this is relevant guidance for firms to follow. When you get many touches across the fragmented regulatory system, it is helpful to know that the guidance is coordinated; it is not coming from just one regulator.

Mr. SCOTT. Ms. Johnson, you said something in your statement that I agree wholeheartedly with: you said that credit is a critical choice. It is almost a life-and-death choice. Can you imagine not having a checking account? Not having a savings account? Not having a credit card? Not having any history in this time? And yet, we have almost 60 million Americans in that shape. How critical, in your words, is this, at this point, with our unbanked, and if we fail in this ability to make the alternatives work, what would that look like? How serious is this situation facing these 60 million unbanked, or what you refer to as invisibles, and making them visible?

Ms. JOHNSON. This is a great question. Thank you. I think I might dissect or sort of bifurcate the question into two parts, one part just being thoughtful at the outset about the idea that credit, as we are describing it, originates from—or the decision-making process, or determinations about credit, originates from an evaluation of eligibility, right? The notion that credit reports are used merely for credit is mistaken. We know that credit reports might be used in other processes to determine employment and access to other resources. So, in some instances, we are talking about credit and the data that is evaluated to determine whether or not someone has access to credit, as a gateway. This is a sort of a gateway to a variety of critical access, to a variety of critical and important resources in our society. Credit is a critical resource and credit re-

ports are a critical factor in the lives of individuals because it may impact their ability to access other resources beyond credit, right?

Mr. SCOTT. Yes.

Ms. JOHNSON. That is the first point, just to segregate out the ideas that what we are evaluating here, the data that is being gathered, there are many important impacts with respect to that data, that are beyond just simply whether or not one qualifies for a credit card or a home mortgage loan. Although, access to those resources is important as well.

I would also underscore—Congresswoman Porter was one of my colleagues in the academy before joining you all here on the committee and in Congress, and her work has historically, along with others, underscored the significance of the financial status of individuals as impacting a variety of elements of their lives, and your point underscores that as well. I just suggest that credit and the data that is being gathered for the purposes of evaluating credit will impact access to financing, but it impacts access to a number of other things, including education.

Mr. SCOTT. And, Mr. Chairman, may I just ask this—one of the values of—

Chairman LYNCH. The gentleman has gone a minute-and-a-half over. Go ahead, though.

Mr. SCOTT. Thank you. One of the values of the fintechs is that they are now providing help and services to the unbanked that our traditional banks are not doing, will not do. And I am not going to ask you to answer that, but I am sure you will agree that that is an area we can develop more of, to use our emerging fintechs to be a valuable asset, because many of the existing actors in the financial services industry are not going to touch these unbanked and underbanked. But, anyway, thank you.

And thank you, Mr. Chairman. I'm sorry.

Chairman LYNCH. Quite all right.

The Chair now recognizes the ranking member of the full Financial Services Committee, the gentleman from North Carolina, Mr. McHenry, for however much time he may consume.

Mr. McHENRY. I will respect the Chair. Thank you, Mr. Lynch. And thank you, Mr. Hill, for your leadership.

It is my hope that this task force can—we can build some consensus around financial technology. This is a nonideological space in an otherwise highly polarized Washington. And I think it shows that we can use technology to get better societal outcome—well, the same or better societal outcomes that we seek in current law.

We have very important provisions of law that have been put in place through a massive amount of work to ensure that we don't discriminate against people based off of what I would describe as superficial reasons. And that work, where you are located, what you look like, who your parents were, any of that stuff, right?

And what we see now in China is that you have this—you have a social score as well. And it is political connections and all of this stuff. And I hear this underlying the whole panel, we don't want that. Just because you tweet and you are a jerk on Twitter doesn't mean you are uncreditworthy. Or if you follow nuts on the left or the right on Twitter, that should not make you more or less creditworthy.

Getting into the fundamentals of this, how you use alternative data, Mr. Girouard, you brought this up. Let's talk about the sandbox approach that Mr. Hill brought up in his question.

So, the question of innovation and financial inclusion, I think, should go hand-in-hand. What are the benefits of a sandbox approach, Mr. Girouard, in your view?

Mr. GIROUARD. As someone who has gone through the process, as we did over 4 years, frankly, with the CFPB, the sandbox isn't to our advantage. We already walked the walk and walked over the coals.

But honestly, in the interest of the American consumer, you want a robust environment where not just small companies but the largest banks have an opportunity to innovate in modeling and in credit decisioning, because it can only benefit the consumer.

A sandbox is necessary because—let me just give an example. In the very early days of our lending, I met with the CEO of one of the top banks in the country, one of the largest card issuers in the country, and his words to me were, "I love what you are doing. I am really glad you are doing it, because we will never be able to do that."

And I think honestly, it may be to my business advantage that that is the case, but it is not to the American consumer's advantage. We need innovation across the industry, not just in unsecured personal loans, but in mortgages, in auto lending, in HELOCs, in all flavors of credit.

Mr. MCHENRY. What will the benefit be if you use alternative data and somebody has, under a traditional score, less than A-plus credit, but you see through alternative data that they actually pay their rent, they pay their cellphone bill, and they have never missed those payments, it enhances that credit score, right?

Others, it would actually say that that credit score is not as good, because they are not paying or they continue to have issues.

There is this picking and choosing, when you say, we only want to use good stuff, if it is alternative data. Well, that is not representative that everyone is a good credit risk, right? How do you prove that out in terms of ensuring it is not discriminatory based off of our traditional metrics under Federal law?

Mr. GIROUARD. Let me just say, the important background is that FICO and income, which are the two anchors of almost any lender, are terribly biased. And they are so biased that the additional of alternative data, whether that is education, whether that is the name of the company you work for—there are a variety of other things—actually reduces the bias and the credit decisioning, because the baseline is so biased itself. That is why it represents such an opportunity.

The other really important—

Mr. MCHENRY. Okay. Across the panel, does anybody disagree with that statement?

Ms. JOHNSON. I would add something.

Mr. MCHENRY. But any disagreement with the contents in the last 5 sentences of what Mr. Girouard said?

Anyone on the panel?

Ms. JOHNSON. There is bias certainly in the existing data, because it is the result of systemic—we just talk about data collection

for algorithms generally. We have to acknowledge that at the outset, the data that is being collected is biased.

One of the best and easiest, most accessible examples, would be in criminal law enforcement. To the extent that an area is overpoliced by police in a particular city or area, there will be more arrests in that area—

Mr. MCHENRY. No, but I am talking about consumer credit, and I am talking about the specifics of this. That is a larger societal issue. We are the Financial Services Committee and not the Judiciary Committee. That is a major issue; I certainly understand that. And I appreciate that.

But let's talk about what we are going to fix here in the Financial Services Committee. When you say that alternative data can be an enhancement—and I understand all of the caveats that all of you in a very loyal sort of way, if I would say, say, yes, it has great opportunities but there are risks. Of course there are, right?

But when we are talking about getting unbanked or credit invisible people and making them visible, I think that is a proper societal tradeoff in order to get more people into the world of being banked, rather than underbanked or unbanked.

And so, I appreciate the hearing. And with that, Mr. Chairman, I yield back.

Mr. SCOTT [presiding]. Thank you, Mr. McHenry.

The gentleman from New Jersey, Mr. Gottheimer, is recognized now for 5 minutes.

Mr. GOTTHEIMER. Thank you, Mr. Chairman, and thank you to all of the witnesses for being here today. I appreciate it.

Traditional information used to make lending decisions and establish credit scores often does not account for the 26 million customers and consumers without a credit history or the 19 million consumers with a short or outdated credit history to form a credit score, groups that are often labeled as thin file or credit invisible.

Thankfully, lenders and CRAs have started using alternative data to make lending decisions, determine credit scores, and expand consumers' access to data.

I personally believe that this is the future in the era of renting and Venmo and Uber, that we need to give the next generation of consumers the ability to build a stronger credit file through non-traditional data sources. That is why I am working on the Credit Access and Inclusion Act, legislation that would allow the reporting of certain alternative data like rent and telecom payments to consumer reporting agencies to help thin-file consumers build their credit scores and hopefully access credit.

We also must ensure traditional credit bureaus and those using alternative financial service data still comply with the Fair Credit Reporting Act, also known as FCRA.

Ms. Wu, if I can start with you, how can we ensure that alternative data sources comply with FCRA data furnishing requirements?

Ms. WU. Thank you for the question, Congressman Gottheimer.

One of the things we need to clarify is that any time third-party data is used for credit decisioning, it should be covered by the Fair Credit Reporting Act.

The example of Facebook, for example. Facebook may have a disclaimer in its website saying you are not supposed to use it for credit. But if they are doing it wink, wink, nudge, nudge, and lenders are using it for credit, it should be covered by the Fair Credit Reporting Act.

And so Congress should clarify that, but I also want to say in the area of sandboxes, the devil is also in the details. Sandboxes shouldn't be a license to ignore things like the Fair Credit Reporting Act and the requirements for accuracy, predictiveness, and notices.

Mr. GOTTHEIMER. Thanks for your answer.

Just a follow-up to that, what kinds of alternative information would you seek to use that is not already shared by applicants or regularly requested as part of loan applications, rental payments, bank statements, and, of course, under the Fair Credit Reporting Act?

Ms. WU. First of all, the most important aspect is consumer choice. The consumer should be allowed the option of sharing it or not. So if they want to share their bank account data, if they want to share their utility payment or rent payment data, they should be permitted to. But if they don't want to, if they want to say, hands off my data, that also should be respected. And then the lender should consider that in the same way they consider credit data.

The other side of this equation of alternative data is, are the lenders actually going to use it? We have seen lenders who won't even upgrade to the latest FICO model, let alone use an alternative score. So, one of the tough parts is actually getting the lenders to look at it.

And I think one of the things that this committee has done that is useful is passing Chairman Lynch's bill giving the CFPB authority to regulate the scoring models. We have heard from Mr. Evans that there needs to be guidance from the regulators. The best thing to do is have the experts at the CFPB review these models and ask, is this predictive, is this accurate, does this create disparate impact? And the bill that this committee passed does that.

Mr. GOTTHEIMER. Do you see that changing with some of the financial institutions? I know that many aren't considering other datasets. Do you see that changing? Is there a desire to—how is the trend line on that? What do you think would really spur that along?

Ms. WU. I think the things that will spur it along are things like, Fannie Mae and Freddie Mac are going to be needing to update their scoring models, and we have actually encouraged the use of pilots, limited pilots with alternative scores.

Mr. GOTTHEIMER. And are we seeing good news out of that? Are we getting more access to credit for people? I really am grateful for your leadership in this space, because I think it is very, very important that more people have access who should get it, who qualify for it, but just because of traditional, the way we have done things forever, they are not getting access to it, or because it is so black box that you don't know what is in it. And I think that lack of transparency also has a big impact.

Ms. WU. Fannie and Freddie have not adopted the new scoring models yet, but some of the other testing that has gone on has shown some promise.

Again, the devil is in the details. We need to be careful. There is going to be some disparate impact. But the thing about the disparate impact test is, it doesn't say, okay, there are some racial disparities you have to stop. Are there more racial disparities or less? Is it predictive? Predictiveness is so key here. And if it is not predictive, you shouldn't be using it.

Mr. GOTTHEIMER. Thank you so much. I yield back.

Mr. SCOTT. Thank you. The gentleman from Ohio, Mr. Davidson, is recognized for 5 minutes.

Mr. DAVIDSON. Thank you, Mr. Chairman. I thank our witnesses, and I thank all of my colleagues for thoughtful questions and good dialogue. And hopefully, this will yield some progress in this really important space.

Mr. Girouard, I want to follow up where Mr. Hill left off when he was talking with you about how much time would this take and how would a sandbox work in a regulatory framework where we have maybe provided certainty for this path with legislation.

And in your response to him, you said, well, we don't really need 18 months; you can know pretty quickly whether it is discriminatory or not; i.e., is it working? And I just want to pick up from there, because it seems incomplete.

Because if you give credit to everyone at low rates or, say, free, it is not discriminatory; it is all free to everyone, whomever shows up, or it is a fixed rate for everyone, no matter what, it is not discriminatory.

But if there is a massive default rate, it really doesn't work, right? You do care about defaults, correct? How far into that process could we know is it both nondiscriminatory and actually effective in the sense that it provides a useful tool?

Mr. GIROUARD. That is a good question. It certainly varies based on the nature of the product. A mortgage, for example, plays out over many, many more years. But you do need enough data, you do need to understand both fairness and efficacy. Fairness can be sorted out fairly quickly. Efficacy takes time. You need to see how a loan performs.

Mr. DAVIDSON. Is it really fair to give money to somebody who has no hope of repaying it?

Mr. GIROUARD. No, it is actually against everybody's best interest to do that. Ability to repay—

Mr. DAVIDSON. Efficacy is inherently linked to fairness is, I guess, the point. And so, I am just curious. If you look at probabilistic models and you look at the statistics and say, hey, if you have this pattern, is there a dataset that shows what the—95 percent certainty, 99 percent certainty, what range of probability of payment history in the early years, could you say the sandbox has produced an effective tool so that it is both nondiscriminatory and it is efficacious?

Mr. GIROUARD. Congressman, you are asking exactly the right questions. The sandbox has to be defined in a way that allows the lender to decide if this new model works. And it won't be the same sandbox for every type of credit product for a variety of reasons—

Mr. DAVIDSON. Okay. That gives me concern, because there is no real hope to pass a law that could provide certainty. It is essentially like, go negotiate your own deal with a regulator.

Mr. GIROUARD. With all due respect, I think rulemaking could absolutely define a sandbox in terms of number of loans, how long the sandbox can operate for, the total dollars in it. There is no question in my mind that a reasonable process could define rules that put a sandbox in place for the major areas of credit for consumers. That would make a significant improvement in the ability to see innovation in this area.

Mr. DAVIDSON. Yes. Thanks for your expertise, and I appreciate your experience in the matter.

Mr. Rieke, your background in privacy at the FTC is interesting, because so much of this links on privacy. And in the United States, particularly in banking, with Gramm-Leach-Bliley, financial institutions have a carveout where they treat data differently.

In a way, financial institutions, and frankly all sorts of institutions, if they were looking at their balance sheet, they might treat their dataset as a valuable asset. Consumers, however, don't necessarily realize that some places they are considered to have a property right in their data. Is it an asset for both?

And as people give up this data, one of the concerns is, how do we reconcile the de facto impact of GDPR and the looming patchwork of privacy laws coming in the United States and Congress' failure to act on privacy with that framework so that consumers can control their data some and not find themselves, well, wait, I was denied credit. Well, yes, you blocked all access to your background, if you go to the far end. And on the other hand, the idea that, gee, if you click these terms and conditions, anything that is in it is fair game.

How do we regulate privacy in this space with respect to credit?

Mr. RIEKE. That is a great question. I think the FCRA is a strong start. If you squint at the text of the FCRA, what comes out of that is if your data is used for important eligibility purposes, certain rights and protections attach.

Now, the FCRA is pretty old now. And as Ms. Wu said, if I am giving permission to Facebook to hand my data over to a lender, it is questionable whether that framework would attach. But I think looking at the spirit of the FCRA, which was created especially for these concerns and were some expansion so that statute might make sense for the digital age, would be where I would start.

Mr. DAVIDSON. All right. Thank you. My time has expired and I yield back.

Mr. SCOTT. Thank you.

And now the gentlewoman from Virginia, Ms. Wexton, is recognized for 5 minutes.

Ms. WEXTON. Thank you very much.

And thank you to the panelists for coming today. This is really fascinating, and you are giving us all a lot to wrap our heads around.

Mr. Girouard, I am really interested in your model and especially the fact point—the datapoint that it reduces interest rates by 5.42 percentage points and approves 95 percent more consumers in that near-prime area.

What kind of response are you getting from lenders about your model? Are they enthusiastic about it?

Mr. GIROUARD. By lenders, do you mean banks we partner or mean to partner with?

Ms. WEXTON. Yes.

Mr. GIROUARD. Thank you, Congresswoman. I would generally say there is a lot of excitement about the potential for a model like this to be able to serve more customers, to be able to build on their side, lower the risk of lending. A more accurate model is intuitively compelling to a bank officer.

Having said that, there certainly remains a lot of concern about regulatory uncertainty. And there is not in any sense a clear-eyed statement or a sense from the regulators how to think about this area of technology to a bank. A no-action letter that we received from the CFPB is a great start. It is not by any sense a panacea, because there are many other regulators. There are many limits to a no-action letter, so there is plenty of room for either regulatory action or rulemaking to provide more clarity.

Ms. WEXTON. I understand that there is some question about regulatory certainty. But are the lenders willing to accept that your model is a more accurate credit reporting model?

Mr. GIROUARD. I think I can comfortably say yes. I am almost universally seen acknowledgment that our model is more accurate and more inclusive.

Ms. WEXTON. Okay. And Ms. Wu had indicated that one of the things that we should consider is making any of these alternative datapoints that are being used for credit to be considered as a report under the FCRA.

Would that impact your ability to create this algorithm, or is that something that would not be an issue for you?

Mr. GIROUARD. FCRA is to cover third-party data, data reported to—and then can be shared with a lender. And that is one part of our data. The other part, which is important to us, is the data that a consumer, with our consent, with their consent, submits to us.

And again, that can be—if they are stating their income to you. That is not something generally a credit reporting agency has information on.

There will always, at least in my mind, be two paths for data to come to a bank and a lender, one through FCRA-related data, through credit reporting agencies, and the other provided by the consumer themselves. And they are both important.

Ms. WEXTON. Okay. Thank you.

And, Ms. Wu, you had also indicated that there should be an opportunity for consumers to opt out of these alternative datasets being used for credit purposes, is that correct?

Ms. WU. Thank you for the question, Congresswoman.

I actually would urge that it would be an opt-in process, that any time you are creating these large new datasets, consumers give their written authorization to have their utility or their bank account information included, to be considered.

Ms. WEXTON. So, they would have to affirmatively opt in—

Ms. WU. Yes.

Ms. WEXTON. —and then get it used.

Okay. And I guess a part of that would be a declination or a refusal to opt in could not be used against them, right? It wouldn't factor into the algorithm, but it wouldn't be down counted for not—

Ms. WU. If they already have a traditional credit file and score and they decline to opt in to alternative data, we would say the lender should go ahead and use the traditional credit score. If they don't opt in, then the data can't be used, obviously.

Ms. WEXTON. All right.

Ms. JOHNSON, as a law professor, I know that you are familiar with the difference between de jour discrimination and de facto.

Is there a way to be proactive in this space and make sure that we don't end up with de facto discrimination in these algorithms, or is it always going to be retrospective, looking back and seeing what the analysis provides us?

Ms. JOHNSON. Thank you for the question, Congresswoman.

I think that there is a way for us to be thoughtful in advance of the release of these types of products in financial markets.

I think earlier, Chairman Lynch referenced the “move fast, break things” mantra that was adopted by a number of technology firms, and now as fintech firms are entering into spaces and operating, as Mr. Girouard mentioned, without clear regulatory guidance, there will be a temptation to use information or data, alternative data, to facilitate what may be faster, more efficient, lower-cost credit evaluation processes.

We do have some knowledge in advance of the types of data that tends to lead to bias or discrimination, based on a long history of legislation and court decisions and agency actions in this space.

I think one of the things we can do is really identify red flags and target areas. Some of the data Ms. Wu mentioned earlier and has been talked about over the course of this hearing, that it is useful and be thoughtful about would be rental payment history, but there are any number of reasons why—and Ms. Wu's organization and others have thought about—that information may disadvantage or utility bill payment may disadvantage certain—

Mr. SCOTT. Ms. Johnson, the time is running out.

Ms. JOHNSON. Thank you very much.

Ms. WEXTON. I yield back.

Mr. SCOTT. Thank you very much.

The gentleman from Missouri, Mr. Luetkemeyer, is now recognized for 5 minutes.

Mr. LUETKEMEYER. Thank you, Mr. Chairman.

Mr. Girouard, since all banks are required to follow the ECOA and you partner with a lot of banks, what due diligence and ongoing monitoring does your company provide your bank partners to ensure that 100 percent certainty for those banks of no fair-lending violations?

Mr. GIROUARD. Sure. That is a very good question. For sure, providing this technology to banks is not for the faint of heart. There is what I would say is a process of probably more than a year of them getting to understand and do diligence on our processes, fair lending being just one of many, to make sure that loans originated using this type of system are within the law. And also, of course, that the creditworthiness is real, the efficacy of the model is real.

So, there is real, significant work before anything happens, before any relationship is signed.

After the fact, there is a constant reporting and auditing like function. The same report that we provide for CFPB for all loans, we can do for an individual bank. And that gives the bank comfort that we are actually monitoring on a regular basis to make sure the loans originated in their name, under their charter, are within the bounds of fair lending regulation.

Mr. LUETKEMEYER. I would like to follow up on the previous colleague's questions here with regards to the no-action letter.

I am assuming that because you have a no-action letter, it is very helpful when you go approach other banks to become partners with them. Because it would sort of seem like you are—it is a get-out-of-jail free card from the standpoint that you have already been sort of preapproved by CFPB, that the modeling you are doing is something that falls within the guidelines of everything.

How important is that no-action letter whenever you start negotiating with the other entities?

Mr. GIROUARD. It is certainly very important. And the reason we were willing to invest information and be as transparent as we were for several years with CFPB, I think, it is important because it demonstrates to banks that we are not a “move fast and break things” company. That may be the name—or sort of a label you want to paint Silicon Valley startups with. But we are not in that class. We are a company that takes regulation and working transparently with regulators seriously.

However, as I said earlier, it is absolutely not a panacea. They care about the FDIC, they care about the OCC, they care about State regulators, all of whom could decide to accept the CFPB's no-action letter and its conclusions or could choose not to. And that is why I think ultimately it is important to clarify regulation.

Mr. LUETKEMEYER. Why do you think more entities like you have not gone the no-action letter route? There are not very many, if any, that have done this, is that correct?

Mr. GIROUARD. There is none other to date, as far as I am aware.

Mr. LUETKEMEYER. Why do you think that you are the only one that has done this? It would seem to me to give you a marketing advantage from the standpoint—if I am a bank and you come to me and you say, look, I have already had my modeling fall within the guidelines of the CFPB and all of the other entities out here that are regulating this, and I will continue to put these processes in place to protect the integrity of our data, it looks to me like you have to a win/win there. Why is there nobody else doing that?

Mr. GIROUARD. My only conclusion I can draw from that is one of a few things. Number one, they are not actually using alternative data in a meaningful way.

Number two, they are using it, but they have found another way, another path to creating comfort that they are within the bounds of fair lending laws.

Or, three, they are using it, but they are not using it responsibly. And I don't necessarily know which of those is the answer.

Mr. LUETKEMEYER. Very good.

Mr. Evans, your testimony points out that CFPB has developed fair lending examinations related to credit models and the Federal

banking regulators have issued guidance to the depositor institutions on third-party or vendor management, including fintechs. However, despite this regulatory framework, there seems to be a disconnect between lenders and fintechs and the regulators that provide uncertainty in the fintechs' place.

Can you explain this?

Mr. EVANS. Well, yes. And I think it goes back to ultimately the fragmented nature of the regulatory system. Fintechs experience uncertainty in that regard, because there are a number of actors in that particular space.

CFPB's position on one thing may differ from the Federal Reserve or the OCC's position. And so oversight of fintech lending requires significant coordination. And the knowledge-building initiatives that I talked about in my opening statement would allow regulators to really understand the fintech products and ensure that the regulatory framework is adaptable and flexible.

Mr. LUETKEMEYER. Okay. Very good. I see that my time has expired. Thank you.

Thank you, Mr. Chairman.

Chairman LYNCH. The Chair now recognizes the gentleman from Texas, Mr. Green, for 5 minutes.

Mr. GREEN. Thank you, Mr. Chairman. And I thank you and the ranking member for hosting this hearing.

I would also like to thank Mr. McAdams for allowing me to proceed at this time. In fact and in truth, it would be his turn, and he has allowed me to have the opportunity to proceed.

I would like to move first, if I may, and rather expeditiously to Mr. Girouard.

Sir, in the model that you currently utilize, do you maintain the traditional credit score and then do you add these other, what you are calling, alternative datapoints to the traditional score?

Mr. GIROUARD. We do. We vary—we use FICO score. We use—

Mr. GREEN. That is going to be enough, because I have a lot to cover. I appreciate it.

Mr. GIROUARD. Okay.

Mr. GREEN. Thank you. I don't mean to be rude, crude, and unrefined.

Mr. GIROUARD. Not at all.

Mr. GREEN. Okay. Thank you.

Friends, I started with Mr. Girouard for a reason. What we are calling alternative data, in most circumstances—there may be some that I am not covering—is really additional data. It is additional data. My bill that I have is not about alternative data, alternative meaning one or another. It is about additional data. It is about what Mr. Girouard does when he takes the traditional data and then he adds what we are calling alternative, but it really is more data that we are adding. We are not leaving out the traditional scores.

My bill does not require consumers to opt in. Consumers do this of their own volition. They can allow their additional data to be scored, and it can help a good many consumers, as evidenced by what Mr. Girouard has called to our attention.

The bill is a bill that has metamorphosed. I confess that initially we used the term, "alternative," but we soon realized that when

people heard the term, “alternative,” they assumed that we were somehow going to negate what was already there as a traditional score.

Now, understanding that we are talking about additional—we are talking about the utilities, we are talking about the rent, but we simply added to what is already there, and in doing this, I think we will give many consumers the opportunity to own a home, and to make purchases that they would not ordinarily be able to make.

Those that don’t opt in will not be—they won’t have that traditional score in any way encroached upon, infringed upon. It won’t have an impact on that. Only those who opt in.

With that said, I want to give you an opportunity to ask me a question. Let’s turn the tables, if you don’t mind, so that we can become as clear as possible, perhaps perspicuously, so that there is a better understanding of what this bill is about.

I am not going to debate persons who want to have an alternative credit scoring model. That is perfectly acceptable to me. I would only suggest that if we focus on this bill, that we use the term, “additional credit scoring.”

Questions from any member of the panel, please?

Ms. JOHNSON. I have a question actually.

Mr. GREEN. Thank you.

Ms. JOHNSON. And Mr. Girouard may answer it, but it grows directly out of your question. Thank you, Congressman, for inviting us to ask.

In the first instance, we have described credit invisibles as those who do not have a traditional credit score under the FICO criteria.

To the extent that inclusion is our goal, which I think is bipartisan motivation for the committee and our thoughtfulness today—if inclusion is the goal and the idea that you propose is that alternative data is additional data supplementing an already robust methodology for analyzing consumer—the likelihood of consumer default or predicting creditworthiness, I am not sure I follow how credit invisibles are actually captured if the data that is being used is not the primary source of evaluation.

Mr. GREEN. If I may answer, because there are only 32 seconds left.

You could be a great Member of Congress, by the way, with your question.

Here is how they are captured. Because they can opt in. And if they have nothing more, that will be there, plus the nothing, plus the something. I hate to be so elemental with the explanation. But what I want to do is make it as clear as possible that what we are doing is leaving the traditional, whatever it happens to be, and then we bring these additional points of data to the scoring process.

Now, given that my time is almost up, and by some standards up, I see—

Mr. SCOTT. Will the gentleman yield for a moment?

Mr. GREEN. I will yield and beg that the Chair would not look at the clock, if you will, please.

Mr. SCOTT. Okay. Very quickly, I think another part of this—

Chairman LYNCH. The gentleman will suspend. We can't be doing this. You are over. If the gentleman wants to conclude his thought, he can, but—

Mr. GREEN. I can't yield?

Chairman LYNCH. The gentleman's time has expired. I'm sorry.

The Chair now recognizes the gentleman from Ohio, Mr. Gonzalez, for 5 minutes.

Mr. GONZALEZ OF OHIO. Thank you, Mr. Chairman, and Ranking Member Hill for holding this hearing today, and thank you to our witnesses. I believe this area is an incredible opportunity to explore how new technologies can be deployed to allow more Americans to gain access to credit. That is sort of the promise or the hope, anyway, of the machine-learning technology.

And I share the sentiment that Ms. Johnson just raised, which is the goal is to expand credit to as many Americans as possible.

Mr. Girouard, I want to focus on your company specifically in the context of the sandbox. And so, we will go there.

Bear with me for a second. You were founded in 2012, according to Crunchbase anyway, and have raised, I think it was \$144 million in total funding.

At what point did you start working with the CFPB directly in the funding stream?

Mr. GIROUARD. I believe our first meeting with the CFPB was either in 2012 or 2013, about that time.

Mr. GONZALEZ OF OHIO. Okay. So really, from the beginning, this was a concerted effort and a decision on your part?

Mr. GIROUARD. That is correct.

Mr. GONZALEZ OF OHIO. Okay. How big was the A, if you are—I don't know if you are allowed to share that, but—

Mr. GIROUARD. I'm sorry?

Mr. GONZALEZ OF OHIO. How big was the series A run, roughly? I will tell you where I am going so you can maybe answer this.

I want you to talk about the benefits of the sandbox in terms of allowing for more startups to enter this space. Because you talked about the big banks potentially being able to get into it. But I want to see more innovation. You guys have an incredible team. I was on your site, a bunch of ex-Googlers and very smart folks. I know there are plenty of folks in Silicon Valley who would love to get into this space.

How would the sandbox enable that?

Mr. GIROUARD. The sandbox brings some clarity, which tends to make the money flow in terms of these companies, first of all, more entrepreneurs wanting to enter this space. When you have a very highly regulated area with a lot of confusion, most entrepreneurs will opt for something else.

If you want more entrepreneurial effort in this area, bringing clarity will bring both the interest of the entrepreneurs and the money from the investors, and that will create companies that are going to make a difference over time.

Mr. GONZALEZ OF OHIO. So, one of the benefits of the sandbox is not just that it gives Wells Fargo a chance, but that it gives the next group of startups a chance as well?

Mr. GIROUARD. Without question.

Mr. GONZALEZ OF OHIO. Great. And then I want to shift to some of the data privacy laws that you have kind of alluded to as well.

California's privacy law is going to be coming into effect. And we hear a lot throughout the industry about the problems that is going to create.

Can you comment on how you see it affecting your business specifically and AI in general?

Mr. GIROUARD. Sure. I believe there are real issues related to privacy and large technology companies that need to be addressed, and I know are being addressed. And I am very appreciative of our home State, California, taking the lead on this.

We are, of course, already preparing, reviewing, and planning to adapt our practices, our product, to the California law. What I would just generally add, of course, is a business like ours operates at a national level, so it would certainly be a step forward for us to have something of that sort, sort of managed at a Federal level more than at a State level. But having said that, we appreciate that is not the way the world works, and we will adapt to California's law.

Mr. GONZALEZ OF OHIO. Yes. I think one of my concerns—and, again, that I keep hearing is when you have this patchwork of 50 different State laws and you want to operate all over the country, as does everybody, you are creating—not you—but California has created a bit of chaos. And I know one thing this committee is committed to is to solving that, which I am excited about.

And then I guess kind of with my last question, as we are thinking through what that national standard should be, what is it about what the California law that you like, and what is it that you think should be changed or different?

If you are not comfortable answering, that is fine.

Mr. GIROUARD. I am not sure I am comfortable enough to try to state that here. Thank you.

Mr. GONZALEZ OF OHIO. Okay. Thanks. With that, I yield back. Thanks.

Chairman LYNCH. The gentleman yields back.

The Chair now recognizes the gentleman from Utah, Mr. McAdams, for 5 minutes.

Mr. MCADAMS. Thank you, Mr. Chairman.

I want to thank the panelists for being here today. And I care deeply about expanding financial inclusion. But I want to make sure we are supporting an environment where all Americans can access credit.

I do know that credit decisions can mean the difference between a family qualifying for a home or a loan to buy a car and the incredible life consequences that those decisions have for each and every potential borrower.

We need to have appropriate consumer protections, and consumer protections shouldn't be ignored while we get the dial right to maximize the benefits while minimizing any potential negative impacts.

But I want to zero in on that balance, the potential benefits and the potential harm or questions that are raised from the use of alternative data.

First, speaking towards the benefits, we have heard testimony today that these alternative data factors are giving lawmakers more confidence in who they can responsibly lend to, meaning more consumers have access to credit, ideally at competitive rates.

My first question is to you, Mr. Girouard. What percentage of your loan portfolio would you estimate that your company can make loans to because of the inclusion of alternative data sources?

Or stated another way, if you were only allowed to use traditional data sources, what percentage of your customers would you not be able to lend to because you couldn't assess their credit-worthiness?

Mr. GIROUARD. Thank you, Congressman. That is a great question.

That is exactly the data that I presented in my up-front statements. What the CFPB asked us to do is to look at our model, if we removed all what you might call alternative data and used only traditional data.

The difference is, among the general population, we described as—and this is among people who have applied for loans at Upstart—is about 27 percent. More people are approved because of the alternative data. But importantly, when you look at the near-prime segment, which is people with somewhat lower FICO scores, it is a 95 percent increase in approvals.

So, it is a very significant difference in improvement on who we can approve due to the alternative data that we include.

Mr. MCADAMS. Thank you. And further on that point, some alternative data factors are now being used to include—or maybe not furthering the point—but in a different direction. Some of those alternative data factors are now being used to include online behavioral data such as online shopping habits, and social network connections.

My next question is for Mr. Rieke. I believe you made this distinction in your testimony between the types of datapoints and the conventional alternative data.

Out of curiosity, how much of these alternative data sources are moving the needle on a credit score? And I am not referring to alternative data such as bill payments, or online utility payments, of which I think most Americans would intuitively understand why that could be included in this credit score. But the online shopping habits, social network usage, how much does that affect an individual's data score?

Are we talking 5 points of credit? One point? Fifty points? And for someone who has a thin credit file, how much are these factors weighted compared to traditional data factors?

Mr. RIEKE. The short answer is, we don't know. Most of the fringe alternative data that has to do with social media and shopping habits and web behavior, there is a lot more hype than reality, in terms of what I have been able to ascertain in our research. There are a lot of start-up companies making some pretty strong claims to the media, and then maybe once they hire a lawyer kind of backing off of those or starting to practice overseas.

And so there is some academic research studies that showed web signals like what website you come from, whether you are using an iPhone or Android phone, can really help kind of narrow in on

what kind of person you are, mostly because those are proxies for wealth.

But in terms of the real science and research around the predictiveness of fringe alternative data, it is really hard to say, because companies hold that data close, and I think there is a lot less of that really happening in the United States today because of issues with the ECOA.

Mr. MCADAMS. Do consumers understand what information on them is being collected and used in their credit decisions, and are there industry standard practices on disclosure?

Mr. RIEKE. I am not aware of any kind of formal industry standard best practices. There are some private businesses, like Credit Karma, that do, in my view, a pretty good job of showing the basic FICO score factors and helping people make sense of that. I have seen nothing resembling that for more complex or fringe datasets.

Mr. MCADAMS. Ms. Wu?

Ms. WU. If I may address that, Congressman, the Fair Credit Reporting Act and the Equal Credit Opportunity Act do require that if someone is turned down or priced higher for credit, a notice goes out explaining what the reasons were. That is really, really important because of the impact these decisions have on people's lives.

One of my concerns is with machine-learning and AI, where the machine itself is determining what factors to use. How do you make sure consumers have adequate information about what is going on inside the black box?

The other thing that I wanted to quickly mention is another type of nonfinancial data that is being used, and Upstart is using, which is education. And I worry about the impact of using education as a form of alternative data. Because we know of the great inequality and racial disparities in terms of what kind of degrees people get.

Mr. MCADAMS. Thank you, Ms. Wu.

My hope would be that we can use this data to not only expand access to credit for more individuals but it can also be used as a form of improving financial literacy, if individuals know what data is being used and what things they might do as individuals to move the needle as well. And hopefully, that doesn't include unfriending their friends on Facebook.

Thank you. And I yield back.

Chairman LYNCH. I thank the gentleman.

The gentleman from Florida, Mr. Lawson, is now recognized for 5 minutes.

Mr. LAWSON. Thank you, Mr. Chairman.

And I welcome the witnesses to the committee.

You might have already responded on this particular issue, but it is important to me. As most of you know, credit reports do not tell the full story of one's economic status. As a matter of fact, they could have a false narrative, depending on the circumstances.

It is estimated that the use of alternative data such as utility payments, rent payments, cellphone payments, and other forms could expand access to credit to over 40 million consumers here in the United States.

Can everyday payments such as rent payments or cellphone payments paint a more accurate picture of someone's ability to pay?

Are we headed in the right direction by saying that this would be a true picture of the individual's ability to get credit?

And everyone—all of you, if you care to respond to that, it would be great. I would just like to know—and you might have already talked about it. But this is talked about all over the place, especially in Florida, where we have a high concentration of students in my district, about 80,000 or 90,000 of them.

So, I am anxious to know what your statement is going to be. And will I tell you, the reason being is when I was coming out of college, I was given all of these credit cards, Exxon, all of them, and so I started using them. And because the invoices, I guess, were going to the dormitory where I used to live, nobody forwarded them to me.

When I got ready to try to get a loan or do some other thing, it came up. And it had been over 1 year or 2 years or so. And I just thought maybe, because I had graduated from school, they just gave me free credit. I didn't know. And that is one of the things that affects a lot of students, because they move around to different places.

That is the reason I wanted to bring that question up and have all of you respond to it.

Ms. WU. Congressman, that is a great point. And you are absolutely right. Traditional credit scores and credit reports often don't reflect the true financial behavior of a consumer, precisely because of things like your experience or the fact that there are a lot of negative marks for things like medical debt, where people got sick and debts were sent to debt collectors.

And we know that even among people with a subprime score, most of them, if you give them credit, will pay it back. Something like 80 percent of consumers who score a 600 will pay it back.

So, alternative data could be useful, especially things like bank account data or rent, and if people choose, if consumers want to supply their utility and cellphone payments. Again, the devil is in the details; how you do it is important. Second-chance scores are better than putting this information in the traditional credit reports.

We are concerned about factors that lead to more inequality or reflect inequality. As Mr. Rieke said, using geographic neighborhood or using what kind of degree you have, because we know that over 36 percent of non-Hispanic whites have a college degree, but less than 16 percent of Hispanics and 23 percent of African Americans do. So, if you use whether or not a consumer has a college degree, it is going to have some stark racial disparities.

Mr. RIEKE. Congressman, I want to just say I think the question of ability to repay is a really good target for this. We are talking about expanding access to credit, but we are not doing anyone any favors by giving them predatory products or too many credit products. That can destroy lives.

So, I think ability to repay is a really important nexus between this question of alternative data and what are we trying to find out, but also a pretty strong consumer protection standard.

Ms. JOHNSON. And I would just echo the earlier reflections. Thank you, Congressman, for this very important question about a really important demographic: students.

We know from the New York Federal Reserve that households face \$13 trillion in debt as of the end of the year, fourth quarter 2018, and \$1.5 trillion in student debt.

Student debt for a particular population, and most recently graduated generations of students, is staggering and crippling. And unlike past generations, these students are moving out of their parents' houses later, and they are having extended job searches.

So, the predatory credit card tactics, the idea of drawing them into spaces where their credit histories will be marred, or they won't have credit histories at all because of how long it is taking them to dig themselves out of educational debt, really does prompt a need, a very significant need for alternative mechanisms, pathways for them to gain access to credit.

I think we are all just thoughtful about how to do that in a way that is effective for consumers, protects their privacy, and is thoughtful about discrimination.

Mr. LAWSON. Mr. Chairman, I know I am out of time, so I yield back.

Chairman LYNCH. I thank the gentleman.

We have agreed to just do one more brief round of questioning, so I yield myself 5 minutes for questioning.

In our discussions with Facebook, in an effort to try to get some accountability on the protection of personal financial data, the issue of assigning fiduciary responsibility for the handling of information was suggested. And I would say that the response from Facebook was evasive, to be generous.

What about that concept that there would be liability for mishandling the financial data that we surrender to fintech companies? Is that something that is workable, do you think, Ms. Wu?

Ms. WU. I thank you for the question, Chairman Lynch. I think that whether you call it a fiduciary duty, or you have legal duties or legal accountability for losing someone's data, there should be a regulatory scheme in place that holds Big Data companies, whether they be credit bureaus or Facebook, to accountability for losing sensitive personal information and data.

Chairman LYNCH. Yes. I guess I should just put a finer point on that.

When I say "fiduciary", I mean in the classic financial services sense where a fiduciary is required to handle that information in the best interest of the customer, and not sell it or deploy it for other purposes. That is what I am getting at.

I am trying to make sure what happens with personal financial data is not what happened with general data that is being vacuumed up and used and deployed without the knowledge or consent, meaningful consent, of individual consumers. That is what I am trying to get at.

Ms. Johnson?

Ms. JOHNSON. Yes. I'd just say thank you, Congressman.

We have examples and models of how to protect financial transaction data that exist in current regulation. The Gramm-Leach-Bliley Act, for example, specifically requires that financial institutions disseminate initial and annual privacy notices to customers regarding financial transactions. The provision of the Gramm-Leach-Bliley

ley Act that I am describing enables consumers to specifically opt out in certain instances of other uses of financial data.

It also requires financial institutions to anonymize data, essentially to the extent that they use data for other purposes, to aggregate the data and ensure that the data is anonymous and not directly reflective or you couldn't easily discern that it refers to a particular consumer based on the profile.

Now, I will say that data scientists at Princeton and Stanford recently published a study illustrating that they could successfully decode, if you will, anonymized data and establish users' identities based on social networking profiles. The idea that this could happen is obviously concerning and gives us pause.

But I do think that we have some examples in existing legislation and regulation that could offer a point of departure for having a conversation about how to create accountability, responsibility, and transparency for anyone who—or entities who are gathering, storing, and distributing personal consumer financial information.

Chairman LYNCH. Great. Let me just jump over to Mr. Girouard. I pulled up Upstart's terms of service agreement. And it is a lot shorter than Facebook's. Thank you very much. It is about 8 pages.

But there is one section in here on limitation of liability. And it says the customer—"you agree that all access and use of the site and its contents and your use of the products and services is at your own risk."

In no event shall we or any lender be held liable for any damages, including direct or indirect, special, incidental or consequential damages, losses or expenses arising in connection with the site or any linked site or use thereof or inability to use by any party or in connection, or for failure of performance, error, omission, interruption, defect, delay in operation, transmission, computer viruses, et cetera." It is very, very broad.

And this is one of those things where you have to click, "I agree." And either you agree to all of this or you don't use the site, you don't use Upstart.

Is that fair to the consumer, do you think?

Mr. GIROUARD. Chairman Lynch, I certainly wish we had a better option. But it is a complicated world. And certainly a business needs to protect its interests.

Somebody could say the internet crashes and I was in the middle of getting a loan, and that just cost me my ability to buy a home or do something else.

Chairman LYNCH. This basically shuts off the consumer from any recovery at all under any circumstances. I understand cases like that where the technology breaks down, you could say in that case, we don't accept any liability.

But in the terms of this, it is airtight where, you basically block off any type of accountability; you are beyond reach by this agreement. This is the type of thing I worry about.

And I just—

Mr. GIROUARD. It is a fair concern. I genuinely believe we have the highest consumer ratings we have ever found in our industry in terms of our respect and the way we treat customers or prospective customers.

Chairman LYNCH. I appreciate that. I am just concerned that no one has any recourse based on the terms of this agreement.

With that, I yield to my friend, the gentleman from Arkansas, for 5 minutes.

Mr. HILL. Thank you, Mr. Chairman.

Just following up on that, Mr. Girouard, that particular thing he read, which obviously we haven't read, but I admired him in real-time going to your website—and it is a thing of beauty. And that is the difference between the House and the Senate, Ms. Johnson. You seem rather concerned about our technological capabilities here.

That is really talking about your—the connectivity between the customers and you, isn't it, protecting you from liability, from the internet or from the website or the connection? Isn't that what that is mostly addressing?

Mr. GIROUARD. Certainly. Any commercial agreement between a consumer and a business has to have reasonable protections in it. I am not an attorney, let me just admit that. So, for me to say what is an appropriate limitation of liability is not something I am probably equipped to speak about today.

Mr. HILL. But we appreciate that. And that is something that we all deal with in any kind of commercial transaction. And I think it is made worse sometimes over the internet, because you don't have any kind of face-to-face explanation and it is a little bit more passive. But I think making sure consumers know what they are getting into is important.

Mr. EVANS. I wanted to ask you. I read your testimony—thanks for it—about this harmonization between the regulatory agencies. You have urged them to adopt a harmonized approach to guidance under use of alternative data and also on the sandbox issue.

Did they give you a timeframe when they would have a harmonized view on that?

Mr. EVANS. They did not. They all agreed with the recommendation and appreciated the spirit of it.

Mr. HILL. Right. I think that is something we have all talked about here. We will be certainly pressing them for this more unified approach on vendor due diligence and an IT exam, guidance on what is an appropriate bank risk profile in this arena, how to do the board review of vendor due diligence. All of this is important.

Mr. EVANS, did you, in your work on this, see any reason to make statutory changes to the Fair Credit Reporting Act or the Equal Credit Opportunity Act?

Mr. EVANS. There are certainly some issues. The scope of the work didn't allow us to rigorously collect all the evidence for us to provide a conclusion on that, but for sure, the complexity of some of the algorithms could limit the type of information that a company is able to provide if they were to deny credit to an individual.

Mr. HILL. Yes. I read the reliability of data point in your testimony, and we have talked about that with other witnesses at previous panels, with just asking the simple question, using an AI-based model, a machine-learning model, that uses additional data, just as Mr. Girouard has described, we have asked the question, is it auditable? And in the instances that we have had, the answer

has been consistently yes. And the evidence of that is not hypothetical because the loans originated are subject to a fair lending exam by a commercial bank or portfolio buyer. And then, of course, in Mr. Girouard's case, they are also auditable by the CFPB's analysis of this data. So, as long as a commercial bank is a partner in it, from my point of view, that seems like the disparate impact test, the HMDA test if it would be a mortgage, or fair lending, or equal opportunity type assessments would be made. Is that generally your view from the work that you did? While you may have found six industry stakeholders who you had concerns about, it is doable to validate a model and have an audit trail as to how the determination was made, isn't it?

Mr. EVANS. In the models that we considered, I would say yes, but I would say our work was limited to the fintech companies we actually talked to. There could be classes of models that—

Mr. HILL. Yes, but the obligation is on that user, that innovator, whether it is a bank partner or a fintech nonbank partner, to demonstrate that they comply with all the compliance obligations of the Federal Government.

Mr. Girouard, you have said already, it is auditable, and the CFPB audits it, and then your bank partner audits it. That is correct, right, you can validate your model and backtrack it?

Mr. GIROUARD. That is correct.

Mr. HILL. And one question that didn't come up today—I didn't hear it—is back-testing. We have had the most ideal circumstances of the past 11 years, thanks to the unbelievable policies of our Federal Reserve, so that we have a very benign interest rate environment, we have rising real wages, we have a rising economy. What about back-testing your \$4 billion you have originated? Looking back under more adverse credit circumstances, what have you learned?

Mr. GIROUARD. Sure. That is a valid concern, and certainly any lender, to earn its stripes, really needs to perform through an economic cycle. First, we test the best we can by simulating higher unemployment. So, there are ways we can simulate higher unemployment and look at the impact we expect it to have on our loan portfolio.

Second, because there actually are recessions, what I might call microrecessions in parts of the country, small parts of the country, we can actually look at loan performance in those particular areas. It is not a perfect proxy, but it is a way to understand how our loans would perform in a weaker economy.

Mr. HILL. I thank the panel, and I yield back, Mr. Chairman.

Chairman LYNCH. The gentleman yields back.

The Chair now recognizes the gentleman from Georgia, Mr. Scott, for 5 minutes.

Mr. SCOTT. Yes, Mr. Lynch, this has indeed been a very, very informative hearing, and our panelists are well-prepared and very informative. Thank you for this. As I said in my opening statement, we are at a new frontier here, and it is an exciting frontier. But let's go back to the alternative data because I think that is really the fundamental foundation of this hearing.

Now, there are different kinds of alternative data that I am hearing. So, you may say a utility bill, or you may say your online hab-

its, or you may say your educational attainment. Tell me, how does a lender weigh these? Would they give more preference to your educational attainment? Does that have the same weight as your utility bill? How is it used by lenders when they are making these underwriting decisions? Let me start with you, Ms. Wu—or is there anybody who has an immediate answer to this? Or if we don't have an answer to it, don't you think we should?

Mr. GIROUARD. I am glad to answer, Congressman, as one who does exactly what you are asking about. Traditional lending systems are what you might describe as rules-based. Okay? A series of, if the person's FICO is between this and that, if their income is between this and that, here is what we can offer them. That is what you might call a rules-based system. These newer models, what are sometimes termed machine-learning or AI, are far more sophisticated than that. What they really do is they look at the history of loans and the data that has come in about those loans, and it learns about how each of those factors actually impacted the performance of the loans. So, it is not a human sitting there trying to evaluate whether education or FICO or something else is more predictive. The software will learn over time what is the best combination of that information, that will be the most accurate model.

The goal of a company like Upstart is to build a more accurate model, which tends to lead to higher approval rates, and we do that by relying on the software to do things that humans can't realistically do, which is to consider not two or three variables, but hundreds or maybe even a thousand variables, and that results in a more accurate credit model. And, fortunately, it also results in one that approves more people at lower rates.

Mr. SCOTT. So, you are saying that the machine has a more accurate ability to give a certain alternative data more weight over the other? I guess what I am asking is, is there more benefit for one type of alternative data to be helpful to the unbanked or underbanked? There is a variety of things. Maybe also added in there, did he serve in the service? What was his rank? Was she a schoolteacher? What was the caliber of her employment structure? Do you see what I am saying? There seems to me that if we just put all this up to the machine, I am not sure it is giving it—there ought to be some weight here.

Ms. WU. That is a great point, Congressman Scott, and something I am concerned about. Because data that kind of looks like credit, as Mr. Rieke said, rent or bank account or utility bills, everybody, if they have a good history, can benefit from that. But if you are talking about things like education, how many college grads really are credit invisible? Are we really expanding access to credit if we say, we will give you a higher score if you graduate from college, especially if you graduate from an elite institution?

Mr. SCOTT. Right.

Ms. WU. And then, in terms of machine-learning, one thing I want to add is, yes, it might be up to the lender to give that weight, but the lender has to be able to explain it. And if all this data is going into a big black box and the machine is deciding what is more important or not, you have to be able to put it on a piece of paper and explain to the consumer what was more important. The law requires it, because we need transparency in lending.

Mr. SCOTT. Yes. Thank you, Mr. Chairman. Great panel.

Mr. EVANS. And so there is—I'm sorry.

Chairman LYNCH. If you can be quick, Mr. Evans.

Mr. EVANS. There is an important tradeoff to think about: transparency versus predictability. And that is something we have to grapple with and it is something about which the regulators can offer guidance.

Chairman LYNCH. Thank you very much. I appreciate that.

The Chair now recognizes the gentleman from Florida, Mr. Lawson, for 5 minutes.

Mr. LAWSON. Thank you.

Mr. HILL. Mr. Chairman, I need to just, I think, politely object. I thought we were going to do just ourselves for a final round of questioning, and I have no more Members here. And so, with all due respect to my friends, that is not really what we agreed to, so—

Chairman LYNCH. Okay. I understand the gentleman is short on time, and I totally respect him, but when I asked for a second round, I meant a second round for the Members.

Mr. HILL. But you said a second round for the two of us, sir.

Chairman LYNCH. Sir, I was not aware that that is the way you understood that.

Mr. HILL. That is the way you said it, and that is the way I understood it.

Chairman LYNCH. Perhaps I meant the two of us, meaning the two sides. I know the gentleman had no other—

Mr. HILL. We have no other Members, so I think just in fairness under the rules, with the deference of Mr. Scott being the last questioner, that would be appreciated.

Chairman LYNCH. The gentleman, Mr. Green, has yielded and—

Mr. LAWSON. I yield back.

Chairman LYNCH. —the gentleman, Mr. Lawson, agrees as well?

Mr. LAWSON. Yes.

Chairman LYNCH. Okay.

Without objection, the Chair moves to include in the record of this hearing a letter from the Cato Institute, Center for Monetary and Financial Alternatives, dated July 24, 2019; a letter from the Financial Data and Technology Association, dated July 23, 2019; and also an article from the Student Borrower Protection Center, entitled, "Educational Redlining: The Use of Educational Data in Underwriting." Without objection, it is so ordered.

I would like to thank our witnesses for their testimony today.

The Chair notes that some Members may have additional questions for this panel, which they may wish to submit in writing. Without objection, the hearing record will remain open for 5 legislative days for Members to submit written questions to these witnesses and to place their responses in the record. Also, without objection, Members will have 5 legislative days to submit extraneous materials to the Chair for inclusion in the record.

This hearing is now adjourned.

[Whereupon, at 12:12 p.m., the hearing was adjourned.]

A P P E N D I X

July 25, 2019



United States Government Accountability Office

Testimony

Before the Task Force on Financial
Technology, Committee on Financial
Services, House of Representatives

For Release on Delivery
Expected at 10:00 a.m. ET
Thursday, July 25, 2019

FINANCIAL TECHNOLOGY

Agencies Should Provide Clarification on Lenders' Use of Alternative Data

Statement of Lawrance L. Evans, Jr., Managing Director,
Financial Markets and Community Investment

GAO Highlights

Highlights of GAO-19-694T, a testimony before the Task Force on Financial Technology, Committee on Financial Services, House of Representatives

Why GAO Did This Study

Fintech refers to the use of technology and innovation to provide financial products and services. Fintech lenders are nonbank firms that operate online and may use alternative data to make loan decisions. In December 2018, GAO issued a report that examined fintech lenders' use of alternative data. The three fintech lending segments that GAO reviewed are personal, small business, and student loans.

This statement—based on GAO's December 2018 report (GAO-19-111)—discusses (1) fintech lenders' use of alternative data, (2) potential benefits and risks of alternative data, and (3) the extent to which federal agencies monitor lenders' use of these data. For that report, GAO reviewed literature and agency documents; analyzed relevant federal guidance; conducted interviews with agency officials and industry stakeholders; and interviewed a sample of 11 fintech lenders (selected based on size, products offered, and other factors).

What GAO Recommends

In its December 2018 report, GAO recommended that CFPB and the federal banking regulators communicate in writing to fintech lenders and banks that partner with fintech lenders on the appropriate use of alternative data in the underwriting process. The agencies concurred and stated that they plan to take action to address GAO's recommendations.

View GAO-19-694T. For more information, contact Lawrence L. Evans, Jr., at (202) 512-8678 or evansl@gao.gov

July 25, 2019

FINANCIAL TECHNOLOGY

Agencies Should Provide Clarification on Lenders' Use of Alternative Data

What GAO Found

Some of the financial technology (fintech) lenders GAO interviewed said they use nontraditional data—also referred to as alternative data—to supplement the traditional data used to make credit decisions or to detect potential fraud. Federal agencies and stakeholders generally define alternative data as information not traditionally used by the national consumer reporting agencies to calculate a credit score. Examples of alternative data include utility payments, cash flow statements, education information, social media activity, and internet browser history.

Using alternative data in credit decisions presents both potential benefits and risks. According to industry stakeholders and literature GAO reviewed, potential benefits include expansion of credit availability and faster credit decisions. For example, alternative data could be used to allow fintech lenders to offer loans to borrowers whose traditional credit history may have been insufficient for banks to extend them credit. Potential risks include disparate impact and other fair lending issues and cybersecurity concerns.

The Consumer Financial Protection Bureau (CFPB) and federal banking regulators have monitored fintech lenders' use of alternative data by collecting information and developing reports on alternative data, but they have not provided lenders and banks with specific guidance on using the data in underwriting. For example, CFPB's fair lending examination procedures and the banking regulators' third-party guidance on risk do not clearly communicate the agencies' views on the appropriate use of alternative data. Nine of the 11 fintech lenders GAO interviewed said additional guidance would be helpful to clarify regulatory uncertainty, which some lenders identified as a barrier to further financial innovation in expanding access to credit. Federally regulated banks that partnered with fintech lenders also told GAO that clarification on appropriate use of alternative data would help them manage their relationships with those lenders. Clear communication from CFPB and the federal banking regulators on appropriate use of alternative data in the underwriting process would bring fintech lenders greater certainty about their compliance with fair lending and other consumer protection laws, and help federally regulated banks better manage the risks associated with partnering with fintech lenders that use these data. Additionally, this communication might allow fintech lenders and their bank partners to innovate and expand access to credit through the responsible use of alternative data.

Chairman Lynch, Ranking Member Hill, and Members of the Task Force:

Thank you for the opportunity to testify today about the use of alternative data in underwriting by financial technology (fintech) lenders. Fintech refers to the use of technology and innovation to provide financial products and services, and fintech lenders are online, nonbank lenders that leverage this technology to provide consumers and small businesses with loans. Fintech lenders typically rely on the internet to offer their loan products and use different sources of funds than traditional banks. These lenders may use traditional means to assess borrowers' creditworthiness, such as credit scores. But their credit models also may analyze large amounts of data from nontraditional sources such as bank accounts—also referred to as alternative data—to determine creditworthiness. Using alternative data in credit decisions could expand responsible access to credit, particularly for some borrowers who lack traditional credit scores because they lack or have insufficient credit history. However, potential risks remain that using alternative data could lead to disparate impacts in credit decisions or pricing and that consumers' data will not be adequately secured.

My testimony today addresses (1) fintech lenders' use of alternative data, (2) potential benefits and risks of alternative data, and (3) the extent to which federal agencies monitor lenders' use of these data. This statement is based on our December 2018 report on fintech lending.¹ For that report, we reviewed reports by industry stakeholders and federal agencies (identified through a literature search) and responses to the Consumer Financial Protection Bureau's (CFPB) request for information on alternative data.² We also interviewed a sample of 11 fintech lenders, which we primarily selected based on size (\$1 billion or more in loans originated or facilitated since the formation of the firm) and products offered, and five fintech lending trade associations. To assess the extent to which federal regulators have monitored the use of alternative data by fintech lenders, we reviewed federal regulators' examination policies, guidance on third-party risk management, and other documents. We

¹GAO, *Financial Technology: Agencies Should Provide Clarification on Lenders' Use of Alternative Data*, GAO-19-111 (Washington, D.C.: Dec. 19, 2018).

²On February 21, 2017, CFPB requested information from all interested members of the public on the use or potential use of alternative data and modeling techniques in the credit process. Through this request, CFPB sought to learn more about current and future market developments, including existing and emerging consumer benefits and risks, and how these developments could alter the marketplace and the consumer experience.

interviewed officials from CFPB, the Department of the Treasury (Treasury), Federal Deposit Insurance Corporation, Board of Governors of the Federal Reserve System (Federal Reserve), Federal Trade Commission, National Credit Union Administration, Office of the Comptroller of the Currency, Securities and Exchange Commission, and the Small Business Administration. Further, we interviewed representatives from federally regulated banks and other industry stakeholders. A complete description of our scope and methodology can be found in our December 2018 report.

We conducted the work upon which this statement is based in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Fintech Lenders Use Alternative Data to Supplement Traditional Underwriting and for Other Purposes

Five of the 11 fintech lenders we interviewed for our December 2018 report said they used alternative data to supplement traditional data when making a credit decision. In addition, one fintech lender used alternative data exclusively in its underwriting decisions.³ As defined by federal agencies and industry stakeholders, alternative data includes any information not traditionally used by the three national consumer reporting agencies when calculating a credit score.⁴ Some of the information defined as alternative data is financial in nature and has characteristics similar to traditional data used by consumer reporting agencies. For example, on-time mortgage payments factor into credit scores, but on-time rental payments do not and are therefore considered alternative data. Other alternative data are nonfinancial. For example, two fintech lenders we interviewed consider the applicant's educational institution

³At least one lender in each of the three fintech lending segments (personal, small business, and student loans) we reviewed in GAO-19-111 stated that they use alternative data when making credit decisions.

⁴Consumer reporting agencies, also known as credit reporting companies and credit bureaus, collect information on consumers that is commonly used to determine eligibility for credit, employment, and insurance, including credit scores. Credit scores are typically calculated using information such as on-time mortgage payments, unpaid debt, number and type of loans, debt collection history, and bankruptcy. The three national consumer reporting agencies are Equifax, Experian, and TransUnion.

and degree when underwriting or pricing a loan. Other examples of alternative data include utility payments, cash flow statements, social media activity, and internet browser history.

According to the fintech lenders we interviewed that use alternative data in making underwriting decisions, the data may be obtained from the borrower, data aggregators, national databases, or other sources. For example, one lender stated that it makes credit decisions using information from multiple credit bureaus in conjunction with alternative data provided by the borrower on educational background, work history, occupation, and employer. Four lenders said that the addition of alternative data helps them better determine a potential borrower's credit risk.

Other lenders we interviewed used alternative data to detect potential fraud. For example, two of these lenders told us they verify a borrower's identity using information collected from the borrower's internet browser. One lender also discussed checking the email address provided by a borrower against a list of email addresses that a third party has identified as fraudulent.

All 11 of the fintech lenders we interviewed stated that they take steps to test their underwriting model for compliance with fair lending laws.⁵ For example, two lenders said they use CFPB's report on fair lending analysis to test for potential discriminatory impacts of their models, and one used model risk management guidance from the federal banking regulators to review its model. In addition, four lenders said they use third parties, including consulting and law firms specializing in fair lending issues, to test their model for compliance with fair lending laws.

⁵Lending practices that result in unequal treatment based on race and sex, among other borrower characteristics, would be a violation of the Equal Credit Opportunity Act. See 15 U.S.C. § 1691.

Stakeholders and Literature Identified Both Benefits and Risks of Alternative Data

Industry stakeholders and literature we reviewed identified several potential benefits associated with fintech lenders' use of alternative data, including expansion of credit availability and faster credit decisions.

- **Expansion of credit.** Several industry stakeholders and literature we reviewed noted that alternative data could be used to allow fintech lenders to offer loans to borrowers whose traditional credit history may have been insufficient for banks to extend them credit. As we reported in March 2018, CFPB officials stated that using alternative data could expand responsible access to credit, particularly for some borrowers who are among the estimated 45 million people who lack traditional credit scores due to the lack of a credit history or an insufficient credit history, including borrowers with a "thin" credit file.⁶
- **Improved pricing of products.** Two industry stakeholders and literature we reviewed discussed how using alternative data may enhance the assessment of a borrower's creditworthiness. As a result, the borrower may be placed in a better credit classification and receive lower-priced credit than would be available using traditional data alone.
- **Faster credit decision.** Two industry stakeholders and literature discussed how using alternative data may allow fintech lenders to reach credit decisions more quickly than traditional banks and improve convenience for borrowers. For example, lenders can utilize alternative data sources to verify application information, including borrower identity, almost immediately.
- **Fraud prevention.** As identified by five industry stakeholders and as discussed earlier, fintech lenders can use alternative data to verify borrowers' identities, which helps prevent fraud.

Using alternative data in credit decisions also presents potential risks, including for disparate impact and relating to cybersecurity.

⁶GAO, *Financial Technology: Additional Steps by Regulators Could Better Protect Consumers and Aid Regulatory Oversight*, GAO-18-264 (Washington, D.C.: Mar. 22, 2018); and Consumer Financial Protection Bureau, *Data Point: Credit Invisibles* (Washington, D.C.: May 2015). According to this publication, there are an estimated 26 million consumers who do not have a credit history with one of the national credit reporting companies. An additional 19 million consumers were estimated to have "unscorable" credit files, which means either that their file is thin and they have insufficient credit history (9.9 million) or that they have stale files and lack any recent credit history (9.6 million).

-
- **Disparate impact and other fair lending issues.** Five industry stakeholders and literature we reviewed discussed the potential for certain types of alternative data to be correlated with characteristics protected by fair lending laws. Accordingly, the use of alternative data in credit decisions raises concerns that borrowers who are part of protected classes may be adversely affected by the data's use.⁷ However, as previously discussed, all 11 of the fintech lenders we interviewed stated that they take steps to test their underwriting model for compliance with fair lending laws.
 - **Transparency of use.** Seven industry stakeholders and literature expressed concerns that there may be a lack of transparency about what alternative data are being used and how they are used in the credit decision. Furthermore, it may be unclear whether the borrower has the ability to dispute the information used.
 - **Reliability of data.** Six industry stakeholders stated that it is difficult to ensure that many forms of alternative data are accurate without validation of the reliability of the data sources.⁸
 - **Performance during credit cycle.** Fintech lending, including the use of alternative data in underwriting decisions, has not been tested in an economic downturn.
 - **Cybersecurity.** As identified by Treasury, recent cybersecurity breaches illustrate the potential for security risks, which may become

⁷For example, according to a Federal Reserve newsletter, it has been reported that some lenders consider whether a consumer's online social network includes people with poor credit histories, which can raise concerns about discrimination against those living in disadvantaged areas. The newsletter noted that instead of expanding access to responsible credit, the use of data correlated with race or national origin could serve to entrench or even worsen existing inequities in financial access. Federal Reserve System, *Consumer Compliance Outlook*, 2nd issue, 2017.

⁸Whereas the Fair Credit Reporting Act requires that borrowers have an opportunity to check and correct inaccuracies in credit reports, borrowers could face more challenges in checking and correcting alternative data that some fintech lenders use to make underwriting decisions because alternative data are not typically reflected in credit reports. See GAO-18-254.

a growing concern as lenders expand beyond traditional borrower data.⁹

Federal Agencies Could Provide Better Communication on the Appropriate Use of Alternative Data

In December 2018, we reported that CFPB and federal banking regulators have monitored fintech lenders' use of alternative data by collecting information and developing reports on alternative data, but they have not provided lenders and banks with specific guidance on using the data in underwriting.¹⁰ Specifically, CFPB has developed fair lending examination procedures that discuss identifying and detecting potential fair lending violations related to credit models. However, CFPB has not issued any procedures or guidance specifically discussing what the agency considers to be appropriate use of alternative data.

Furthermore, the federal banking regulators each have provided third-party or vendor management guidance to depository institutions, which describes the risk assessment, due diligence and risk monitoring, and oversight in which banks should engage when working with third parties, including fintech lenders. However, the guidance does not specifically mention fintech lending activities or provide specific information on how bank management should monitor a third party's use of alternative data and any associated risks. The federal banking regulators also have taken some steps to supplement the existing third-party guidance with more information that may be applied to banks' relationships with fintech lenders. However, these efforts do not include clarification on the regulators' views on alternative data. For example, in July 2016, the Federal Deposit Insurance Corporation issued proposed third-party lending guidance that outlines the risks that may be associated with third-party lending, and the expectations for a risk-management program, supervisory considerations, and relevant examination procedures.¹¹ However, the proposed guidance does not specifically address alternative data.

⁹We discuss cybersecurity and privacy concerns generally and among fintech firms in several other reports. For example, see GAO, *Consumer Privacy: Changes to Legal Framework Needed to Address Gaps*, GAO-19-621T (Washington, D.C.: June 11, 2019); *Consumer Data Protection: Actions Needed to Strengthen Oversight of Consumer Reporting Agencies*, GAO-19-196 (Washington, D.C.: Feb. 21, 2019); *Internet Privacy: Additional Federal Authority Could Enhance Consumer Protection and Provide Flexibility*, GAO-19-52 (Washington, D.C.: Jan. 15, 2019); and GAO-18-254.

¹⁰GAO-19-111.

¹¹Federal Deposit Insurance Corporation, *Proposed Guidance on Third-Party Lending*, Financial Institution Letters 50-2016 (Washington, D.C.: July 29, 2016).

Nine of the 11 fintech lenders we interviewed said additional guidance would be helpful to clarify regulatory uncertainty. For example, one lender stated that guidance that clarified the appropriate use of alternative data and outlined steps lenders should take to ensure compliance with fair lending laws would be helpful. Another lender said it would be helpful for federal regulators to develop additional guidance clarifying how to incorporate different types of data in lending decisions. Additionally, three fintech lenders stated that the uncertainty surrounding the use of alternative data arising from the lack of clear guidance acts as a barrier to further financial innovation and potential expansion of access to credit.

Similarly, federally regulated banks that partnered with fintech lenders told us that clarification on appropriate use of alternative data by federal regulators would help them manage their relationships with those lenders. For example, representatives of one bank said that a partner fintech lender's use of alternative data may be attractive from an innovation and business perspective, but the bank likely would hesitate to agree to using these data due to regulatory uncertainty. Bank representatives stated that more guidance therefore would be helpful to clarify what is and is not permissible in terms of alternative data.

Thus, in our December 2018 report, we recommended that CFPB and the federal banking regulators communicate in writing to fintech lenders and banks that partner with fintech lenders on the appropriate use of alternative data in the underwriting process. The agencies concurred and stated that they plan to take action to address our recommendations. Clear, consistent communication on alternative data can help provide fintech lenders with greater assurance that they are complying with fair lending laws when using these data in their credit decisions. It also can help the banks effectively manage the risks associated with partnering with lenders that use these data. Additionally, this communication may allow fintech lenders and their bank partners to innovate and expand access to credit through the responsible use of alternative data. We will continue to monitor the agencies' progress in addressing our recommendations.

Chairman Lynch, Ranking Member Hill, and Members of the Task Force, this concludes my prepared statement. I would be pleased to respond to any questions you may have at this time.

**GAO Contact and
Staff
Acknowledgments**

If you or your staff have any questions about this testimony, please contact Lawrance L. Evans, Jr., Managing Director, Financial Markets and Community Investment at (202) 512-8678 or evansl@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this statement.

GAO staff who made key contributions to this testimony are Winnie Tsen (Assistant Director), Erika Navarro (Analyst in Charge), Namita Bhatia-Sabharwal, Abigail Brown, Adrienne Cline, Farrah Stone, Robert Lowthian, Jessica Sandler, and Jennifer Schwartz.

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

**Obtaining Copies of
GAO Reports and
Testimony**

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (<https://www.gao.gov>). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to <https://www.gao.gov> and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <https://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube. Subscribe to our RSS Feeds or E-mail Updates. Listen to our Podcasts. Visit GAO on the web at <https://www.gao.gov>.

**To Report Fraud,
Waste, and Abuse in
Federal Programs**

Contact FraudNet:

Website: <https://www.gao.gov/fraudnet/fraudnet.htm>

Automated answering system: (800) 424-5454 or (202) 512-7700

**Congressional
Relations**

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800, U.S. Government Accountability Office, 441 G Street NW, Room 7149, Washington, DC 20548

**Strategic Planning and
External Liaison**

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707, U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548



Please Print on Recycled Paper.



Testimony of
Dave Girouard
CEO and Co-Founder, Upstart Network, Inc.

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

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

July 25, 2019

Chairman Lynch, Ranking Member Hill, and Members of the Task Force on Financial Technology, thank you for the opportunity to participate in today's conversation.

My name is Dave Girouard, co-founder and CEO of Upstart, which is a leading artificial intelligence ("AI") lending platform. I founded Upstart more than 7 years ago in order to improve access to affordable credit. In the last five years, almost \$4 billion in bank-quality consumer loans have been originated on our platform using a model that combines alternative data with AI and machine learning algorithms to determine a borrower's creditworthiness.

Concerns about fairness in algorithmic lending, particularly in the use of alternative data, are well founded. As a company focused entirely on reducing the price of credit for the American consumer, fairness is an issue we care about deeply.

In our early days at Upstart, we conducted a retroactive study with a large credit bureau and uncovered a jarring pair of statistics: just 45% of Americans have access to bank-quality credit, yet 83% of Americans have never actually defaulted on a loan. That's not what we would call fair lending.

The FICO score was introduced in 1989 and has since become the default way banks judge a loan applicant. But in reality, FICO is extremely limited in its ability to predict credit performance because it's narrow in scope and inherently backward looking.

At Upstart, we decided to use modern technology and data science to find more ways to prove that consumers are indeed creditworthy - to bridge that "45% versus 83%" gap. We believe that consumers are more than their credit scores. And by going beyond the FICO score, and

including a wide variety of other information such as a consumer's employment history and educational background, we've built a significantly more accurate credit model.

While most people believe a better credit model means saying no to more applicants, the truth is just the opposite: Because Upstart's model is more accurate, we have significantly higher approval rates and lower interest rates than a traditional model.

But we also understood that consumer protection laws weren't to be taken lightly. Thus we proactively met with the appropriate regulator - the Consumer Financial Protection Bureau (CFPB) before launching our lending platform.

After several years of good faith efforts between Upstart and the CFPB to determine the proper way to measure bias, we demonstrated that our AI-driven model doesn't result in unlawful "disparate impact" against protected classes of consumers. Because AI models change and improve over time, we developed automated tests with the regulator's input in order to provide reports on the impact of our credit decisions across underserved groups on a quarterly basis. We have been providing this information to the CFPB for the last 18 months.

Moreover, we were also able to report to CFPB that our AI-based system improved access to affordable credit. Specifically:

- **Our model approves 27% more consumers and lowers interest rates by 3.57 percentage points, compared to a traditional lending model**
- **For near-prime consumers (620-660 FICO) our model approves 95% more consumers and reduces interest rates by 5.42 percentage points compared to a traditional model**
- **Upstart's model provides higher approval rates and lower interest rates for every traditionally underserved demographic**

That's the type of consumer benefit we should all get excited about. In September 2017, Upstart received with the first-ever "No Action" letter from the CFPB, recognizing that Upstart's platform improves access to affordable credit without introducing unlawful bias.

The concern that the use of alternative data and algorithmic decisioning can replicate or even amplify human bias in lending is well-founded. However, in Upstart's experience, the fair lending laws enacted in the 1970s and the substance of fair lending regulation enforcement—that is, monitoring and testing the impact on actual consumers who apply for loans—translates very well to the AI-driven world of today.

But in reality, the path we walked at Upstart is insufficient to create a robust and competitive market that will maximize financial inclusion and credit access. In our early days at Upstart, we couldn't know for certain whether our model would be biased. It wasn't until loans were originated that we were able to demonstrate that our platform was fair. As an early-stage startup, this was a risk worth taking, but it's not a risk a large bank would have considered.

If broader and deeper financial inclusion among American consumers is important to this committee, it's worth considering rule making or legislation that will provide some form of limited sandbox for model development and testing.

By combining regulatory support for model innovation with rigorous and standardized testing, we can ensure that we don't forego the clear and obvious consumer benefits that AI and alternative data can offer to the American consumer.

Thank you.



Written Testimony of

**Kristin N. Johnson, McGlinchey Stafford Professor of Law and
Associate Dean of Faculty Research, Tulane University Law School**

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

**Examining the Use of Alternative Data in Underwriting
and Credit Scoring to Expand Access to Credit**

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

**The Future of Finance:
Alternative Data in Credit Underwriting**

Chairwoman Waters, Ranking Member McHenry, Stephen Lynch, Chair of the Financial Technology Task Force, and Bill Foster, Chair of the Task Force on Artificial Intelligence, Members of the Committee and Members of the Task Force:

Thank you for inviting me to participate in this Hearing to discuss the use of alternative data in credit underwriting and credit scoring to expand access to credit.

I am a professor of law and associate dean of faculty research at Tulane University Law School where I teach courses on corporate and securities law, the integration of emerging technologies into financial markets, systemic risk across financial markets and financial markets regulation.¹ I am an affiliate of the Murphy Institute at Tulane and director of the Institute's Financial Market Stability Program.² I am here today solely in my academic capacity and am not testifying on behalf of any entity.

Over the last several years, a body of sophisticated algorithms commonly described as artificial intelligence ("AI") and distributed ledger technologies have altered the financial market ecosystem, creating a new class of financial institutions – fintech firms.³

¹ I have previously served as an analyst at Goldman Sachs, a Vice President and Associate General Counsel at JP Morgan and an associate at a New York law firm supporting many of the largest financial institutions in the country. For the last decade, as an academic, I have published research examining the risk management implications of nascent financial products and services in credit and capital markets.

² The Murphy Institute is a privately-funded, interdisciplinary center at Tulane University that aims to support and advance applied research in public policy, public affairs and civic engagement and to inspire and educate students and the interested general public in the understanding and analysis of challenging economic, moral, and political questions.

³ The discussion presented here appears in forthcoming academic journal manuscripts.



Supplementing traditional credit underwriting data inputs and processes, fintech firms employ newer modeling techniques and consider a broader range of source data referred to descriptively (rather than normatively) as alternative data. These new inputs include information regarding consumers' financial transactions, recurring payments history and a behavioral score based on social networking and digital-interface. Fintech firms include both the non-depository digital platforms that operate independently and platforms that partner with legacy banks to originate loans.⁴ Fintech firms servicing credit scoring and underwriting markets offer great promise but also present unique concerns.⁵

The introduction of alternative data may improve access to credit for many consumers with nonexistent or insufficient credit histories. According to estimates, twenty-six million Americans do not have traditional credit histories and are considered "credit invisible." Another nineteen million Americans have thin (limited), impaired or stale (outdated) credit histories and, as a result, cannot obtain credit scores using traditional scoring methodologies ("credit unscorable").

Unsavory lending practices, detestable marketing tactics and usurious interest rates have too often plagued marginalized consumers who face persistently fragile financial circumstances.⁶ Unlike legacy credit scoring businesses such as Equifax, Experian and Transunion that rely on commercially available credit scoring models like the Fair Isaac Corporation Lenders ("FICO") methodology fintech firms increasingly rely on alternative credit scoring models and nontraditional source data. According to proponents, the development of nascent methodologies and alternative data enables fintech firms to expand access to credit to consumers historically deemed invisible or unscorable.

Legislative and regulatory authorities must, however, balance fintech firms' laudable promises of greater inclusion with the significant risks posed by integrating alternative data and new methodologies. Careful examination of the rise of alternative data and the evolution in consumer credit underwriting methods casts a spotlight on fintech firms' promises of inclusion and reveals the perils of relying on source data that may not be demonstrably predictive of creditworthiness as well as the potential for predatory or discriminatory practices to undermine the anticipated benefits of alternative source data and credit evaluation processes.

Fintech firms integrating alternative data and modeling techniques must satisfy long-standing fairness and accountability standards, engage in responsible innovation and commit to provide sufficient transparency, meaningful disclosure, auditing and necessary

⁴ Christopher K. Odinet, *Consumer Bitcredit and Fintech Lending*, 69 ALA. L. REV. 781 (2018).

⁵ FED. TRADE COMM'N, *BIG DATA: A TOOL FOR INCLUSION OR EXCLUSION?*, at i, (2016).

⁶ Kathleen C. Engel & Patricia A. McCoy, *A Tale of Three Markets: The Law and Economics of Predatory Lending*, 80 TEX. L. REV. 1255, 1261 (2002).



internal controls to meet statutory obligations regarding their methodologies and minimize the potential for discriminatory effects on legally protected classes.

Fintech Firms and Learning Algorithms

“Fintech” is a catch-all term used to refer to the digital platform or internet-based financial services firms that engage in digital transfers, storage, payments systems, digital asset origination (such as cryptocurrency) and secondary market trading, investment advising and digital credit scoring and origination. To capitalize on economic efficiencies, reduce transaction costs and mitigate commonly-identified enterprise risks, fintech firms integrate artificial intelligence technologies such as supervised or unsupervised machine learning, deep learning or neural networks (“AI”) or distributed ledger technologies into their business models. While there is no universally adopted definition of AI, the term refers to a diverse, but related, set of technologies that train through a reinforcement learning process, simulate human decision-making and cognitive behavior and engage in predictive analysis.⁷

Advancements in the collection, storage and analysis of vast volumes of data (“big data”) fuel AI platforms designed to automate decision-making in several key sectors including healthcare, education, employment, criminal law, security, surveillance, communications and finance. While the inclusion of data crunching algorithms is nothing new – investment banking firms, for example, have long relied on sophisticated algorithms to predict timing, pricing, risk and other factors that influence investment and trading decisions – the rapid adoption of learning algorithms that interpret alternative data in consumer credit markets presents significant risks.

Automating Credit Decisions

Learning algorithms at the center of fintech platforms’ credit evaluation processes analyze vast quantities of data in fractions of a second. Fintech platforms replace face-to-face meetings with loan officers and cumbersome and time-consuming paper-based credit application processes with applications accessible on internet-enabled smartphones, tablets and other mobile or personal devices. Removing human underwriting agents and their biases arguably reduces the likelihood of intentional discrimination. AI-based credit scoring methodologies may enhance consumer default predictions and lead to better credit classification and possibly lower-priced credit than traditional credit scoring methodologies. Together these process-oriented improvements enhance efficiency and accuracy, improve pricing, reduce operating and loan origination costs and enable fintech

⁷ Examples of AI modeling techniques include but are not limited to decision trees, random forests, artificial neural networks, k-nearest neighbors, genetic programming, and “boosting” algorithms. Given the limited time available and suggested scope, only an abbreviated description of artificial intelligence and other referenced technologies appears in the submitted written testimony.



firms to offer credit to a greater diversity of consumers, in particular those who have struggled to obtain credit.

Traditional credit evaluation processes like FICO consider tradeline information, including but not limited to existing and previous loan obligations, repayment history, credit limits, account status for revolving accounts, credit inquiries, public records such as civil judgments, tax liens and bankruptcies. Incumbent credit scoring methodologies predominantly use multivariate regression analysis to correlate past credit history to consumer credit outcomes and evaluate the likelihood of default or delinquency. Increasingly, incumbent credit scoring firms and traditional methodologies are shifting their evaluation criteria. As fintech firms tout the benefits of AI driven decision-making, both incumbent credit scoring firms and insurgent fintech platforms rely on alternative sources of data and scoring methodologies.

Alternative Data

According to industry and federal and state agency reports, alternative data refers to information not traditionally used by the national consumer reporting agencies ("CRA") in calculating a consumer's credit score. In some instances, alternative data simply expands the categories of payment history beyond those considered by CRAs. For example, some fintech platforms integrate telecommunications (mobile phone and cable bills), utilities or residential rental payment history. In other instances, fintech firms expand the types of information considered in credit scoring processes and include financial transaction data (checking account cashflows).

Alternative data may assist historically marginalized (credit invisible and unscorable consumers) to gain access to conventional credit markets. There is good reason to believe that capturing nontraditional data may enable consumers with thin, impaired or nonexistent credit files to demonstrate a history of timely bill payment. The frequency of telecommunications, utility and rental payments may enable consumers to generate a different but valuable track record or consistent, timely bill payment history.

Limitations and conflicts arising from the use of alternative data to expand access to credit. Consumer advocates have, however, expressed some concerns regarding the impact of integrating certain data points, such as utility bill payments. Relying on utility or cable bill payment histories may disadvantage low-income consumers for various reasons. First, dispute resolution processes for public utilities and cable services may differ from other types of recurring obligations. Second, utility bill balances may fluctuate seasonally, prompting some consumers to delay payments or fall behind on pay utilities bills. Low-income or fixed income families are particularly susceptible to these circumstances.



Consider, for example, the families living in areas of the country that face severe seasonal weather patterns. For families living in the northeastern part of the country, for example, home heating bills may present significant monthly demands during the winter and families may not be able to pay utility bills on-time or in full at the close of each billing cycle. Similar challenges may arise for families living in southern states during the summer months. Finally, the significance assigned to recurring residential bills may disadvantage families that migrate seasonally based on employment opportunities or periodically relocate based on service in the armed forces.

Consumer advocates also expressed concerns that H.R. 435 proposed during the 115th Congress would preempt consumer privacy protections by amending Section 623 of the Fair Credit Reporting Act ("FCRA") to permit utilities and landlord to furnish payment information to a CRA "notwithstanding any other provision of law." Without modification, the earlier version of proposed H.R. 435 would override federal requirements that a subsidized housing provider obtain a consumer's consent before sharing rental payment information. The recently proposed Credit Access and Inclusion Act addresses the preemption concerns and solicits a two-year study and report from the Government Accountability Office on the impact of furnishing additional information.

Financial transaction and social networking data. Expanding credit evaluation criteria beyond additional types of recurring payments, alternative data may also include personal consumer financial transaction data – bank account and credit/debit card transactions, including deposits, transfers or withdrawals. Methodologies integrating alternative data may also incorporate educational (major and university attended) or professional accomplishments.

Proponents of alternative data also advocate for the inclusion of nonfinancial, behavioral data. These data points may include digital interface information such as clickstream data, audio and text data, internet browsing and search habits, geo-spatial data and survey or questionnaire data.⁸ Beyond simply browsing preferences, fintech firms are also integrating highly-personalized reputational data. For example, fintech firms are assessing consumers' social network status, web-scraping data from consumers' financial transactions and social media activities and ranking consumers based on relational social connections (consumers' status as "social influencers") through analysis of exchanged messages and friends tagged in social media posted photos.

It is not yet clear how these new sources of data will impact those without credit reports or with thin or stale credit files. It is also unclear how credit invisibles and unscorables who do not have conventional checking and savings accounts or credit cards will generate financial transaction data. Similarly, ranking consumers based on higher educational or

⁸ Mikella Hurley and Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J. L. & TECH. 148, 159 (2016).



professional accomplishments seems likely to replicate the current credit scoring patterns. Finally, credit invisibles and unscorables that lack a presence on social media are unlikely to engender the relational benefits or rewards associated with social networking. In fact, familial and neighborhood associations may make it more difficult for consumers who have not traditionally qualified for credit on fair and reasonable terms to gain access to better, higher quality credit products.

Indisputably, however, the rising significance of alternative data has ignited interest across various markets for greater access to consumer financial data. Consistent with its dominance in the general technology market, Facebook has directly approached banks requesting access to consumers' financial transaction data⁹ and registered for a patent for a technology that assesses users based on social network connections. Technology firms often seek to gather sensitive data from consumers but resist transparency regarding the uses of consumer data.¹⁰

Regulating Alternative Data

The harvesting, distribution and integration of financial transaction and behavioral scoring data raises significant questions regarding consumer protections, privacy and discriminatory practices.

Alternative data such as financial transaction data - credit and debit card and checking account transaction history- may offer valuable insights. Information regarding financial transaction activities and behavior may better inform evaluations of factors that are correlated to consumer credit risk assessment. A consumer's financial history is, however, sensitive information. Unmonitored use and distribution of this information challenges consumer protections and privacy norms.

Privacy Concerns - Existing and Proposed Federal Oversight

A host of state and federal regulators and this Committee are actively seeking to clarify the types of alternative data and the method for including these new class of information in emerging and evolving credit scoring processes. This Committee has held multiple hearings to explore these questions.

More specifically, in February of 2017, the Consumer Financial Protection Bureau ("CFPB") announced a comprehensive Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process. The Government Accountability Office ("GAO") issued a report in March of 2018 - *Additional Steps by*

⁹ Emily Glazer, Deepa Seetharaman and AnnaMaria Andriotis, *Facebook to Banks: Give Us Your Data, We'll Give You Our Users*, WALL ST. J., Aug. 6, 2018.

¹⁰ FRANK PASQUALE, *THE BLACK BOX SOCIETY* (Harvard Univ. Press, 2015).



Regulators Could Better Protect Consumers and Aid Regulatory Oversight - and a second report in December of 2018 - *Agencies Should Provide Clarification on Lenders' Use of Alternative Data* - recommending a series of policies including proposals to coordinate agencies' regulatory efforts, clarify standards governing alternative data and minimize uncertainty regarding the use of alternative data in the underwriting process. In the absence of effective state or federal regulatory intervention, many warn that fintech firms will take advantage of gaps in oversight and engage in regulatory arbitrage.

Advocates argue that existing regulations sufficiently address consumer protection, privacy and antidiscrimination concerns. Under the Gramm Leach-Bliley Act, financial institutions may not distribute "raw" consumer data to third parties; instead, prior to distributing consumers' personal financial data, financial institutions must aggregate, anonymize and de-identify personalized transaction details. Financial institutions must also send consumers initial and annual privacy notices and allow them to opt-out of sharing their personal transaction information with unaffiliated third parties.

These protections are, however, weak and evidence suggests that they do not effectively protect consumers' confidential personal financial information. Using statistical methods, data scientists can decode or de-anonymize aggregated consumer social media and financial transaction data. In other words, data scientists can reverse the steps taken by financial institutions to de-identify consumer data and match consumer data with individual consumers' profiles. A recent study by Stanford and Princeton researchers details a theoretical methodology for de-identified web browsing histories and linking individual search histories to social media profiles using only publicly available data to facilitate the matching process.¹¹

Behavioral scoring presents even more pernicious concerns. According to proponents of behavioral scoring, the likelihood that a consumer will default on payment obligations may be determined by evaluating the consumer's network of friends, neighbors, folks with similar interests, income levels, and backgrounds. Unlike consumer financial transaction data and payment history evaluations, however, behavioral scoring may not be demonstrably predictive of financial responsibility.

Credit is, indisputably, a critical resource. Individuals and families increasingly rely on credit to finance household purchases or overcome significant, unanticipated expenses.¹² Without access to credit on fair and reasonable terms, it can be extraordinarily expensive to be poor. For families with fragile financial circumstances, credit may serve as a lifeline,

¹¹ Jessica Su, Ansh Shukla, Sharad Goel, Arvind Narayanan, *De-anonymizing Web Browsing Data with Social Networks*, Apr. 3, 2017, <https://sharad.com/papers/twivacy.pdf>.

¹² REPORT ON THE ECONOMIC WELL-BEING OF U.S. HOUSEHOLDS IN 2018, BRD, GOVERNORS FED. RESERVE SYS. (2019), <https://www.federalreserve.gov/publications/files/2018-report-economic-well-being-us-households-201905.pdf>.



enabling consumers to meet short term debt obligations, and to pay for education, housing, and even food.¹³

Consumers navigate an ever-widening web of debt. According to the Federal Reserve Bank of New York's Center for Microeconomics – at the close of the first quarter 2019, families and individuals face over \$13 trillion in debt obligations.¹⁴ Rising college and university tuition rates have fueled an increase in educational debt obligations. Students and their families currently owe approximately \$1.5 trillion in student loan debt.¹⁵ A parallel narrative in the home mortgage loan market has led American households to borrow over \$9 trillion in mortgage debt.¹⁶

Credit reporting agencies have a special role in financial markets and fintech firms operating at the intersection of startup innovation and consumer credit origination raise a number of the normative questions.¹⁷ As AI increasingly influences the terms and availability of credit, this nascent technology will also inevitably perform a gatekeeping function, determining who receives access to credit, and for those with access, learning algorithms will likely decide the most fundamental terms of any credit arrangement.

Privacy Concerns - Adopted and Proposed State Laws and Regulation

In the absence of definitive federal regulation addressing the use of alternative data, several state laws require disclosure regarding the use of alternative data by credit scoring platforms or limit the use of alternative data.

California Consumer Privacy Act. Signed by Governor Jerry Brown on June 28, 2018, the California Consumer Privacy Act ("CCPA") grants a consumer the right to request that a business "disclose the categories and specific pieces of personal information that it collects about ...consumer[s], the categories of sources from which that information is collected, the business purposes for collecting or selling the information, and the categories of [third] parties with which the information is shared."¹⁸ The CCPA also enables consumers to request the deletion of personal information, opt out of the sale of personal information, and access the personal information in a "readily useable format."¹⁹

¹³ See Abbye Atkinson, *Rethinking Credit as Social Provision*, 71 STAN. L. REV. 1093 (2019) (describing the dangers of making credit a key determinant of whether and how basic needs are met).

¹⁴ QUARTERLY REPORT ON HOUSEHOLD DEBT AND CREDIT (Q1 2019), CTR. MICROECON. DATA: FED. RESERVE BANK N.Y. (2019), https://www.newyorkfed.org/medialibrary/interactives/householdcredit/data/pdf/HHDC_2019Q1.pdf.

¹⁵ *Id.*

¹⁶ *Id.*

¹⁷ E. Gerald Corrigan, *Are Banks Special?: ANNUAL REPORT 1982*, FEDERAL RESERVE BANK OF MINNEAPOLIS (raising fundamental questions regarding the role of banks and prudential regulation).

¹⁸ Assembly Bill No. 375, https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

¹⁹ Dipayan Gosh, *What You Need to Know About California's New Data Privacy Law*, Harvard Business Review (July 11, 2018), <https://hbr.org/2018/07/what-you-need-to-know-about-californias-new-data-privacy-law>.



The CCPA construes “personal information” broadly. Under the CCPA, “personal information” means “information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household.”²⁰ Similarly, the law also offers a broad definition of the term “sell;” consequently, any of the following activities constitutes a sale of consumer data: “disclosing, disseminating, making available, transferring or otherwise communicating orally or in writing or by electronic or other means” a consumer’s personal data. Examples of personal information include consumer’s personal identifiers, education information, geolocation, biometric data, internet browsing history, psychometric data, and “inferences” drawn from information used to create a profile about a consumer, reflecting the consumer’s preferences, predispositions or behavior, among other attributes.

The CCPA requires companies to obtain consent from customers before selling their personal data to third parties, but it does not apply to consumer information that is de-identified. “De-identified” information is personal information that cannot reasonably identify, relate to, describe, or be linked to a particular consumer.²¹ In addition, the CCPA does not apply to “aggregate consumer information,” which is information that relates to a group or category of consumers, from which individual consumer identities have been removed, that is not linked or reasonably linkable to any consumer or household or device.²²

Critics have challenged the breadth of the CCPA and the likely impact that the law would have on established business models in the technology sector including several of the largest technology companies such as Facebook, Twitter, and Google.²³ This restriction may extend to internet service providers such as AT&T and Verizon, which collect broadband activity data (web browsing data) and may generate behavioral profiles to enable digital advertising.²⁴

New York State Senate 2302 and Department of Financial Services Regulatory Guidance

As of July 2019, the New York State Assembly is considering the adoption of Senate bill 2302 - a bill that would prohibit consumer reporting agencies from using information about the members of a consumer’s social network to evaluate the consumer’s creditworthiness.²⁵ The bill defines the term “members of a consumer’s social network” as “a group of individuals authorized by a consumer to be part of his or her social media communications and network.”²⁶ The bill prohibits consumer reporting agencies from

²⁰ Assembly Bill No. 375, https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

²¹ *Id.*

²² *Id.*

²³ *Id.*

²⁴ *Id.*

²⁵ <https://legislation.nysenate.gov/pdf/bills/2019/S2302>.

²⁶ *Id.*



“collect[ing], evaluat[ing], report[ing], or maintain[ing] in the file on a consumer the credit worthiness, credit standing or credit capacity of members of the consumer’s social network for purposes of determining the credit worthiness of the consumer; the average credit worthiness, credit standing or credit capacity of the consumer’s social network; or any group score that is not the consumer’s own credit worthiness, credit standing or credit capacity.”²⁷ In addition to pending legislation limiting the use of social networking behavioral information in consumer credit evaluation processes, New York state financial services regulators have expressly limited the use of alternative data in the context of life insurance underwriting methodologies.

New York State Department of Financial Services Insurance Circular: Use of External Consumer Data and Information Sources in Underwriting for Life Insurance

On January 18, 2019, the New York State Department of Financial Services (“NYSDFS”) issued an insurance circular with two guiding principles on the use of alternative data in life insurance underwriting. First, insurers must independently determine that external data sources do not collect or use prohibited criteria. Insurers may not rely on a vendor’s claim of that alternative data does not reflect bias or result in discrimination against protected classes. Insurers may not evade their obligations to comply with antidiscrimination laws by pointing to the proprietary nature of a third-party process.²⁸ Notwithstanding the fact that alternative data may be provided by third-party vendors, the NYSDFS emphasized that “the burden” to ensure compliance with antidiscrimination laws “remains with the insurer at all times.”²⁹

Second, insurers should not use external data unless they can establish that it is not “unfairly discriminatory.”³⁰ Insurers must be confident that the use of alternative data is demonstrably predictive of mortality risk. The Circular also notes that “transparency is an important consideration in the use of external data sources to underwrite life insurance.” Insurers using external data should be confident that the use of the data is demonstrably predictive of mortality risk and that they can explain how and why this is the case.³¹

Fair Credit Reporting – Alternative Data as a “Consumer Report”

The Fair Credit Reporting Act (“FCRA”) imposes obligations on CRAs - entities that provide consumer reports - as well as anyone who uses or furnishes information included in consumer reports. The FCRA defines consumer reports as “communication[s] of any information by a consumer reporting agency bearing on a consumer’s creditworthiness, credit standing, credit capacity, character, general reputation, personal characteristics, or mode of living which is used or expected to be used or collected in whole or in part for

²⁷ *Id.*

²⁸ *Id.*

²⁹ https://www.dfs.ny.gov/industry_guidance/circular_letters/cl2019_01

³⁰ *Id.*

³¹ *Id.*



determining a consumer's eligibility for credit, employment purposes, or any other purposes enumerated in the statute."

A number of questions arise as fintech firms begin to gather alternative data and generate credit assessments. If fintech firms' consumer credit assessments based on alternative data constitute "consumer reports," consumers and consumer advocates may assert that fintech firms are subject to the obligations imposed on CRAs under the FCRA. In addition, CRAs may only distribute consumer reports for limited purposes identified in the statute. Consumer reports may be furnished (i) in connection with a credit transaction involving the consumer, (ii) for employment purposes, (iii) in connection with insurance underwriting, or (iv) in accordance with the consumer's written instructions. Consequently, entities gathering data and fintech firms and other firms that obtain and resell data may violate the FCRA by impermissibly using and transferring assessments based on alternative data if such assessments constitute consumer reports. As described in the CFPB request for information and the GAO reports, federal regulators should clarify the contexts in which nontraditional data or alternative data will be deemed "consumer reports" and the instances in which fintech firms may be deemed CRAs.

Adverse Action Notices – Explainability

The FCRA and Equal Credit Opportunity Act ("ECOA") also impose an adverse action notice requirement for entities that take action with respect to any consumer that is based, in whole or in part, on any information contained in a consumer report. State law parallels federal obligations for adverse action notices.

Under relevant provisions of New York Insurance Law referenced above, for example, insurers must notify consumers of their right to receive the "specific reason or reasons for a declination, rate differential, or other adverse underwriting decision." According to the NYSDFS Circular issued earlier this year, if an insurer uses alternative data to underwrite insurance, the reason(s) provided to the consumer for any adverse action "must include details about all information" underlying the decision, including the specific source of the information.

Satisfying adverse action notice requirements may present a significant challenge for platforms using learning algorithms to review large volumes of alternative data. The inscrutable and non-intuitive nature of learning algorithms suggests that even developers may be unable to explain the specific rationale underlying an algorithm's credit or insurance underwriting decision.³² As a result, it may be difficult for CRAs to explain adverse actions as contemplated under the existing regulatory framework.

³² Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085 (2019)



Bias, Fairness and Inclusion

Under ECOA and federal fair lending regulations, intentional discrimination based on a protected trait is prohibited under antidiscrimination statutes. Facially-neutral algorithms mitigate the risk that consumers will face intentional discriminatory treatment based on legally protected traits such as race, gender or religion; this suggests that fintech firms employing automated decision-making platforms are not likely to engage in intentional discrimination and therefore are less likely to violate antidiscrimination statutes. The operational mechanics of learning algorithms may, however, mask an algorithm's reliance on a trait that functions as a proxy for a legally protected trait.

Evidence demonstrates that incomplete or inaccurate data sets may influence the objectivity of learning algorithms. Perhaps even more alarming, learning algorithms are designed to identify the most expedient path or optimal variable for solving a problem or making a decision. Learning algorithms seek to identify variables that simplify and expedite the sorting, classifying and ranking of identified subjects. To that end, learning algorithms may rely on proxies or traits that are highly-correlated with protected traits.³³

This approach may result in the learning algorithm relying on facially-neutral variables in a manner that masks prohibited decision-making behavior.³⁴ In other words, the algorithm may make decisions using facially neutral variables that function as proxies in the decision-making process for prohibited criteria, violating antidiscrimination protections.

Even if developers expressly program algorithm's not to discriminate on the basis of a protected trait, the developers' biases may creep in and influence the algorithm's operation. Three examples illustrate concerns regarding biases in the data sets.

First, inaccurate, incomplete and otherwise flawed data sets may potentially amplify discrimination.³⁵ To illustrate this concerns, consider Amazon's attempt to use an automated decision-making platform to evaluate, score and rank job applicants for a software developer position.

Amazon created a resume review platform designed to identify and sort candidates with desirable attributes for a software developer position. The platform received facially-neutral instructions regarding educational or skill prerequisites and analyzed the resumes of employees recently hired for similar computer programmer positions. Beyond this

³³ Daniel Schwarcz and Anya Prince, *Proxy Discrimination In The Age Of Artificial Intelligence And Big Data*, IOWA L. REV. (Forthcoming 2020).

³⁴ Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 678 (2016).

³⁵ SAFIYA NOBLE, ALGORITHMS OF OPPRESSION (2018); VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE AND PUNISH THE POOR (2018); Margaret Hu, *Algorithmic Jim Crow*, 86 FORDHAM L. REV. 633 (2017).



initial data set and series of instructions, the platform taught itself to mimic human-like decision-making behavior. As the platform began to review real candidates' resumes, it operated independently, using cognitive analysis to decide which candidates to interview without specific instructions regarding the submitted resumes.³⁶

Amazon's goal was to identify best athletes in a competitive pools of applicants.³⁷ Notwithstanding programmers' intentions, the platform began to "penalize resumes that included the word 'women's,' as in 'women's chess club captain,'" and "downgraded graduates of. . . women's colleges."³⁸ Amazon's experiment illustrates the risk that an automated platform will inherit the biases that data sets and developers unknowingly introduce, leading to unanticipated and potentially prohibited discrimination against individuals who are members of a legally protected class.

Second, selecting and cleaning data sets involves human judgment. Data sets are often compiled by third party vendors and distributed to developers who utilize the data to create a training data set. A learning algorithm's successful analysis depends significantly on the data used to train the algorithm.

In order to achieve the predictive benefits of learning algorithms, data sets require a large number of observations. Even if a data set has a sufficient number of observations, the data must be subjected to several pre-processing steps including, among others, cleaning, partitioning, sampling, scaling and feature selection. These steps are necessary because datasets are rarely free from missing or inaccurate values. Data scientists must decide how to resolve missing values. The options for addressing these concerns may include removing the subjects with missing values from the data set and excluding them from the analysis. At each step from data collection decisions to the development of the algorithm, human judgment will influence how the algorithm operates. Finally, some commentators have demonstrated that underrepresentation, particularly of members of legally protected classes, may lead to digital discrimination.³⁹

One study suggests that fintech firms using AI based methodologies are replicating historic biases. According to the results of the study both fintech and traditional mortgage

³⁶ See Jeffrey Dastin, Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women, REUTERS (Oct. 9, 2018), <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>.

³⁷ *Id.*

³⁸ *Id.*

³⁹ Joy Buolamwini and Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROCEEDINGS OF MACHINE LEARNING RESEARCH, CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 1-15 (2018) (present an approach to evaluate bias present in automated facial analysis algorithms and datasets).



origination firms lending practices result in discrimination against Latinx and African-American borrowers.⁴⁰

Cyber Security Concerns

In addition to privacy and discrimination concerns, permitting fintech firms and CRAs to collect alternative data heightens cybersecurity concerns. The rising cost, frequency, and severity of data breaches now dominate risk management discussions. Over the last ten years, more than 4,000 known data breaches have shocked, debilitated, and even (temporarily) paralyzed markets. Commentators estimate that vast numbers of records containing confidential or sensitive data have been compromised. Experts suggest that data breaches cost the global economy more than \$ 400 billion dollars of losses annually.

As cyberattacks multiply, governments, corporations, and citizens scramble to mount a successful defense against cyber-intrusions. The size, sophistication, and diversity of styles of the cyberattacks renders these activities among the most perilous of emerging risk management concerns.

The cyberattacks against financial institutions threaten the stability of financial markets and create personal costs for consumers exposed during data breaches. As the New York State Department of Financial Services noted, “[c]yber hacking is a potentially existential threat to our financial markets.” Federal regulators have warned that cybersecurity threats may “wreak serious havoc on the financial lives of consumers.”

Financial transaction and social media data present particularly attractive targets for hackers. Pursuant to federal regulation and consistent with their business models, large financial institutions acquire, collect, and retain significant volumes of personal information. Collection, storage and transfer of this sensitive data renders financial institutions and retailers highly attractive targets for hackers.

Cyberattacks capture national and international attention because of their pervasive effects. For example, in December 2013, Target, a national retailer, announced that it was the target of a massive data breach. The hackers who orchestrated the data breach obtained the confidential credit and debit card information of more than 40 million customers. As investigations ensued, Target continued to adjust its estimate of the number of records accessed, ultimately reporting that hackers captured the personal data of as many as 110 million customers. In 2014, hackers invaded home improvement retailer Home Depot’s records and acquired 56 million customers’ credit and debit account information and 53 million customers e-mail addresses.

⁴⁰ See, e.g., Robert Bartlett et al., *Consumer-Lending Discrimination in the Era of FinTech* (October 2018) (unpublished manuscript), <https://faculty.haas.berkeley.edu/morse/research/papers/discrim.pdf> (an empirical study comparing discrimination in lending by traditional mortgage origination firms with face-to-face interaction with borrowers and decisions made by fintech platforms; the study finds that “lenders charge Latinx/African-American borrowers 7.9 and 3.6 basis points more for purchase and refinance mortgages respectively, costing them \$765M in aggregate per year in extra interest”).



Equifax's settlement this week illustrates the perils of cyberattacks against credit reporting agencies. Between mid-May 2017 and July 2017, Equifax, one of the country's largest CRAs suffered one of the largest known financial data breaches, exposing the personal information (names, addresses, dates of birth, Social Security numbers, and driver's license numbers) of more than 148 million Americans, 8,000 Canadians, and nearly 700,000 UK citizens.

Former Equifax CEO Richard Smith in testimony before Congress explained that the data breach resulted from hackers' exploitation of a flaw in "Apache Struts," an open source web application. While a patch was released during the first week of March 2017, Equifax failed to apply the security updates until two months later. Equifax should have addressed this vulnerability within forty-eight hours, but it did not.⁴¹ Equifax's information security scans also failed not detect the Apache Struts vulnerability.⁴²

On May 13, 2017, hackers exploited this vulnerability to access Equifax's systems and consumers' personally identifiable information.⁴³ Between May 13, 2017 and July 30, 2017, evidence suggests that the attackers continued to access sensitive information, exploiting the same Apache Struts vulnerability without being detected by Equifax's security tools.⁴⁴

Mr. Smith notified the Equifax board about the breach on August 22, 2017.⁴⁵ On September 7, 2017, Equifax disclosed the breach to the American public.⁴⁶ In other words, Equifax waited six weeks from the time they discovered the breach until they disclosed said breach to the American public. The Equifax settlement marks one of the largest data breach settlements and will provide up to \$425 million in consumer restitution. The settlement reflects a number of measures that Equifax will take to protect consumers' personal data and assist with fraud detection.

The Equifax data breach demonstrates the systemically important role of CRA in credit markets and US financial markets. As the universe of fintech firms expands, regulatory oversight of these entities must reflect the nature of the information that the firms will collect, store and transfer. Regulation must also reflect the significant role of these firms in the stability of consumer credit markets and broader financial markets.

As the Office of the Comptroller of the Currency ("OCC") and Federal Deposit Incorporation's ("FDIC") consider paths for granting fintech firms special purpose

⁴¹ *Oversight of the Equifax Data Breach: 7 Answers for Consumers*, Hearing Before the U.S. H. Comm. on Energy and Commerce Subcomm. on Digital Commerce and Consumer Protection, 115th Cong. 3 (2017) (statement of Richard F. Smith, CEO, Equifax), <https://democrats-energycommerce.house.gov/sites/democrats.energycommerce.house.gov/files/documents/Testimony-Smith-DCCP-Hrg-on-Oversight-of-the-Equifax-Data-Breach-Answers-for-Consumers-2017-10-03.pdf>.

⁴² *Id.*

⁴³ Stacy Cowley, *2.5 Million More People Potentially Exposed in Equifax Breach*, N.Y. TIMES (Oct. 2, 2017), <https://www.nytimes.com/2017/10/02/business/equifax-breach.html>.

⁴⁴ *Oversight of the Equifax Data Breach: 7 Answers for Consumers*, Hearing Before the U.S. H. Comm. on Energy and Commerce Subcomm. on Digital Commerce and Consumer Protection, 115th Cong. 3 (2017) (statement of Richard F. Smith, CEO, Equifax), <https://democrats-energycommerce.house.gov/sites/democrats.energycommerce.house.gov/files/documents/Testimony-Smith-DCCP-Hrg-on-Oversight-of-the-Equifax-Data-Breach-Answers-for-Consumers-2017-10-03.pdf>.

⁴⁵ *Id.*

⁴⁶ *Id.*



nonbank charters and Industrial Loan Corporation (“ILC”) charters concerns mount regarding careful monitoring of fintech firms’ privacy and cybersecurity measures and their ability to protect the collection, storage and transfer of alternative data.

Blockchain-Based Credit Scoring and Lending Models

For several years, fintech firms and conventional CRAs have integrated learning algorithms into credit scoring models. In more recent years, developers began to advocate for credit scoring models built on decentralized, distributed digital ledger protocols.

On January 27, 2018, Jesse Leimgruber, Alain Meier, John Backus published a whitepaper for Bloom Protocol, a “credit staking” decentralized credit scoring platform powered by Ethereum and the Interplanetary File System. According to the whitepaper, Bloom plans to offer three main services: Bloom ID (Identity Attestation), BloomIQ (Credit Registry) and BloomScore (Credit Scoring). According to Bloom, its model addresses the shortcomings of traditional credit scoring by transitioning the credit scoring process to the blockchain protocol. Touting its success as one of the first distributed ledger credit scoring and lending platforms in the world, Bloom promises to facilitate cross border credit scoring, accommodate users with no credit history, secure personal information, increase global access to credit development and provide greater competition in the credit risk evaluation market.

Bloom introduces three unique models: the BloomID, BloomIQ and Bloom Score. Using a peer assessment methodology, Bloom claims that consumers with thin, limited, impaired or no credit history may demonstrate creditworthiness and enjoy access greater access to credit. While the whitepaper clearly indicates that the model will evaluate conventional criteria such as loan and bill repayment history, Bloom relies heavily on social networking to assess a consumer’s eligibility to receive credit. A number of important details regarding Bloom’s methodology are not revealed in the whitepaper, but there is significant potential for a decentralized distributed ledger based credit scoring platform to assist invisible and unscorable consumers by offering greater transparency in the credit evaluation process, a more easily reviewable and correctible credit report and reduced incidents of fraud and data breaches.

For decades, consumer advocates, academics, regulators and state and federal legislators have recognized that low-income consumers pay remarkably more for basic financial services such as check cashing, money transfers and short-term loans. Nearly ten percent of American households continue to lack access to traditional savings and checking accounts.



Consumers with limited access to basic banking services, those living in financial services deserts (requiring them to commute significant distances to bank branches) have had too few options for obtaining access to credit on fair and reasonable terms. Check cashing storefronts, payday loan outlets and other predatory financial services providers exploited invisible or unscorable consumers' lack of access to conventional banking and credit services.⁴⁷

Fintech firms operating at the intersection of startup innovation and consumer credit evaluations raise a number of the normative questions.⁴⁸ As artificial intelligence increasingly influences the terms and availability of credit, this nascent technology and the firms adopting it will come to perform an important gatekeeping function, determining whose receives access to credit, and for those with access, learning algorithms will likely decide the most fundamental terms of any credit arrangement.

To be sure, the advent of artificial intelligence technology disrupts legacy banking, inspires a new market infrastructure and spurs development that *may* benefit unbanked and underbanked consumers. The successful expansion of access to credit may depend largely on regulators' effective supervision of the integration of alternative data and reliance on opaque, inscrutable and non-intuitive algorithms.

⁴⁷ Twenty percent of families with traditional bank accounts still rely on alternative financial services outlets. Jason Furman, *Financial inclusion in the United States*, June 10, 2013 available at <https://obamawhitehouse.archives.gov/blog/2016/06/10/financial-inclusion-united-states>.

⁴⁸ E. Gerald Corrigan, *Are Banks Special?*: ANNUAL REPORT 1982, FEDERAL RESERVE BANK OF MINNEAPOLIS (raising fundamental questions regarding the role of banks and prudential regulation).

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 unscorable 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/facebook-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.

NCLC

© 1990-2019 National Consumer Law Center, Inc. All rights reserved.
**NATIONAL
CONSUMER
LAW
CENTER***

2019-07-25 10:00:00 AM EDT
**Advancing Fairness
in the Marketplace for All**

Testimony before the

**U.S. HOUSE OF REPRESENTATIVES COMMITTEE ON FINANCIAL SERVICES
Task Force on Financial Technology**

Regarding

**“Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to
Credit”**

July 25, 2019

Chi Chi Wu

Staff Attorney

National Consumer Law Center

7 Winthrop Square, 4th Fl.

Boston, MA 02110

617-542-8010

cwu@nclc.org

Testimony of Chi Chi Wu, National Consumer Law Center
 Before the U.S. House of Representatives Committee on Financial Services
 Task Force on Financial Technology
 regarding
 “Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to
 Credit”
 July 25, 2019

Introduction and Summary

Chairman Lynch, Ranking Member Hill, and Members of the Financial Technology Task Force, thank you for inviting me to testify today regarding the use of alternative data in underwriting and credit scoring to expand access to credit. I offer my testimony here on behalf of the low-income clients of the National Consumer Law Center.¹

Alternative data has the potential to benefit millions of consumers, whether they are “credit invisible” or they have impaired records with the traditional Big Three nationwide consumer reporting agencies or “credit bureaus” - Equifax, Experian, and TransUnion. However, alternative data is not without its risks.

Much of the analysis of whether alternative data will benefit or hurt consumers depends on several key factors –

- What kind of alternative data is being used?
- How is the alternative data being used?
- What is the accuracy and predictiveness of the data?
- What level of disparate impact does the data have on protected groups, especially communities of color?

If alternative data is used for credit decisionmaking, its use must be regulated by the Equal Credit Opportunity Act (ECOA). Unless the data represents direct experience between the lender and the consumer, it should also be regulated by the Fair Credit Reporting Act (FCRA). Compliance with both these laws will be critical for the purposes of accuracy, predictiveness, transparency and minimizing disparate impact.

¹ The National Consumer Law Center is a nonprofit organization specializing in consumer issues on behalf of low-income people. We work with thousands of legal services, government and private attorneys, as well as community groups and organizations, from all states who represent low-income and elderly individuals on consumer issues. As a result of our daily contact with these advocates, we have seen many examples of the damage wrought by abuses from credit reporting agencies from every part of the nation. It is from this vantage point that we supply these comments. *Fair Credit Reporting* (9th ed. 2017) is one of the nineteen practice treatises that NCLC publishes and annually supplements. This testimony was written by Chi Chi Wu.

A. Introduction

Consumers are “credit invisible” if they either (1) do not have any credit history with Equifax, Experian, or TransUnion; or (2) if their histories are too scant or old (“thin”) to generate a credit score. According to the Consumer Financial Protection Bureau (CFPB), 26 million Americans (or about 1 in 10) do not have a credit history, and another 18 million are unscorable.²

There are significant racial disparities with respect to credit invisibility. About 15 percent of African American and Latinx consumers have no credit history compared to 9 percent of Whites. Another 13 percent of African Americans and 12 percent of Latinx consumers are unscorable, compared to 7 percent of Whites.³

For young consumers or those who are new to this country, credit invisibility may be a temporary status. Some consumers may deliberately avoid using credit and others, such as older consumers, may no longer need or use credit. A few consumers may even want to avoid the credit reporting system altogether, especially after the massive Equifax data breach of 2017.

For some purposes, the lack of a credit history could be better than a negative history. Employment and insurance are two uses of credit report information where no history may be better than solely negative information, and where invisibility may be a benefit. For example, a number of state insurance laws are designed to ensure that a consumer with no credit score is not treated worse than someone with an average credit score.⁴

However, credit invisibility poses real and significant problems for some consumers seeking affordable credit in order to buy a home, start a business, or simply open a mainstream credit card account. Policymakers, advocates, and industry members have all proposed alternative data as solution to credit invisibility. The term encompasses many different types of data, used in many different ways – examples range from monthly payment obligations like rent and cell phone bills to social media data. The CFPB has defined alternative data as information not in the traditional credit reports provided by Equifax, Experian, and TransUnion.⁵ The Government Accountability Office further distinguishes financial and non-financial alternative data.⁶ The latter is sometimes referred to as “Big Data.”

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

³ *Id.*

⁴ The Supreme Court observed the existence of these laws in *Safeco Ins. Co. of Am. v. Burr*, 551 U.S. 47, 73, 127 S. Ct. 2201, 2217, 167 L. Ed. 2d 1045 (2007) (“[a] number of States permit the use of such ‘neutral’ credit scores to ensure that consumers with thin or unidentifiable credit histories are not treated disadvantageously”).

⁵ Consumer Financial Protection Bureau, Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process, Docket No. CFPB-2017-0005, 82 Fed. Reg. 11183, May 19, 2017. [hereinafter “CFPB Alternative Data RFI”]

⁶ Government Accountability Office, Financial Technology: Agencies Should Provide Clarification on Lenders’ Use of Alternative Data, GAO-19-111, Dec. 2018, at 34.

Some types of alternative data and approaches to using it are helpful or promising, but others are harmful or pose risks. We have urged a cautious and thoughtful approach in developing solutions. As with so many aspects of credit and financial services, “the devil is in the details.”

B. The Type of Alternative Data Matters Greatly

Not all alternative data used to create a potential credit score is created equal. Some data shows promise, other data is a mixed bag, and some data is harmful enough that it should not be used.

I. Financial alternative data

Financial alternative data includes information about payment obligations that are periodic and recurring, similar to the loan payment obligations in a credit report, or reflect other mainstream financial information, such as bank account transactions. Some financial alternative data shows promise, such as bank account transaction or rental payment data. Other financial alternative data could be harmful, such as subprime loan or “no choice” gas & electric utility payment data.

- **Bank account transaction/cashflow data** – The analysis of a consumer’s bank account transactions and cashflow, *i.e.*, credits, debits, and balances, has shown significant potential as a source of financial alternative data. This same data is sensitive, private, and a tempting target for other uses. Because of its importance, there is a need for significant guardrails. We discuss bank account data further and set forth proposed principles for its use in Section C.
- **Rental data** – Traditionally, rental payment data was only reported when overdue rent was sent to a debt collector, usually after a tenant had been evicted or left the unit. Efforts to add positive data appear to be promising, especially those efforts that do not report late payments that are not sent to (and thus not reported by) a debt collector. Any efforts to include rental payment data should not penalize tenants who invoke their rights under state or local laws to withhold rent due to poor conditions.

The most difficult issue is probably getting sufficient amounts of rental payment data, since the vast majority of landlords do not report such data and do not have great incentive to do so.⁷ Many landlords are small property owners who might find it cumbersome to report payments regarding only one or two units.

A variation on using rental payments is to use address stability (*i.e.* number of times a consumer has moved) as a proxy for on-time payments. This is an overly crude measure of rental payment history, because consumers often move due to many reasons other than nonpayment of rent. As the CFPB notes, military servicemembers are one population that moves often.⁸ Another population that moves frequently are students. This lack of precision makes address stability a poor measure.

⁷ FICO, Risk & Compliance Truth Squad: Can Scoring Rental Data Vastly Improve Credit Access?”, May 10, 2017 (rental data makes up less than 1% of data reported to credit bureaus).

⁸ CFPB Alternative Data RFI, 82 Fed. Reg. at 11187.

- **Gas and electric utility data** – Most gas and electric companies currently only report accounts on traditional credit reports when they are very seriously delinquent. The push to add “full file” monthly reporting of gas and electric bill payment data to traditional credit reports, without giving consumers any choice, has the potential to hurt millions of low-income consumers by adding reports about payments that are only 30 or 60 days late. Data from utility companies and regulators shows that the percentage of low-income and energy assistance utility customers who are delinquent can range from 17% to 38%.⁹

The impact could be especially harsh on families who need time to pay off winter or summer bill spikes. Reporting of late payments could also undermine state consumer protections, such as prohibitions against wintertime shut offs for elderly or other vulnerable consumers, by compelling them to pay seasonally high bills even if it doesn't leave them enough money to pay for other critical expenses such as food.

In contrast, efforts to include utility data on a voluntary basis, for example by collecting the data from bank account transactions with consumer permission, could be useful. These efforts also show there is no need for utility credit reporting in which the consumer has no choice.

- **Subprime credit** – Payday loans and other forms of subprime credit are often not reported on traditional credit reports. Adding these types of credit could damage the credit records of these borrowers, even if they are current. High-cost credit is often designed to lead to a cycle of debt, and even merely using a subprime form of credit might negatively affect a credit score.¹⁰ Creating a database of consumers who have used high-cost credit could make them even more vulnerable. Adopting the use of subprime credit data would likely harm consumers who use such credit.

Furthermore, the necessity of using subprime credit as a form of alternative data is questionable, as most payday borrowers are not credit invisible. The authors of one

⁹ Consumer Groups' Comments in Response to the CFPB's Request for Information Regarding Use of Alternative Data and Modeling Techniques in the Credit Process, Docket No. CFPB-2017-0005, May 19, 2017, available at https://www.nclc.org/images/pdf/credit_reports/comments-alt-data-may2017.pdf, at 4.

¹⁰ Experian states:

- “When assessing “riskier” consumers, lenders want to know who is likely to pay them back and who is not. What are the red flags to determine the difference between the two?
- *Number of payday loans acquired within a year*
 - Number of first-payment defaults
 - Number of inquiries within the past 30–90 days
 - Income discrepancies reported by consumer.”

Experian, *The State of Alternative Credit Data*, 2018, <https://www.experian.com/assets/consumer-information/white-papers/alternative-credit-data-paper.pdf>, at 20 (emphasis added).

See also Ruth Lythe, *A Payday Loan Can Stop You Getting a Mortgage: Worrying News for Three Million Who Have Needed Fast Cash*, *The Daily Mail*, July 13, 2012, available at <http://www.dailymail.co.uk/money/mortgageshome/article-2173443/Payday-loans-Worrying-news-million-needed-fast-cash.html#ixzz20beuLWIm> (in United Kingdom, reporting that lenders are refusing to grant mortgages to payday loan borrowers).

research study analyzing payday loan applications noted that the results of their matching payday borrowers to credit bureau files “imply that nearly all of the payday loan applicants had a credit record at the time they applied for their first payday loan.”¹¹

- **Telecommunications data** – Over the past several years, there have been efforts to create scores using data from wireless phone and cable television providers. For example, FICO XD uses data from the National Consumers Telecom and Utility Exchange (NCTUE) as well as LexisNexis to generate an alternative score that can be used for otherwise credit invisible consumers.

Unlike regulated electric and gas service, telecommunications industries have fewer consumer protections that could be undermined by monthly reporting. Thus, they do not raise the same concerns as those sources of data. However, the level of accuracy of the data has not been assessed, nor the impact on consumers who lodge disputes over issues such as cramming, early termination fees, and questionable surcharges.

Transparency is also an issue; consumers may also not be aware that their cell phone and cable payment histories are being supplied to nationwide or specialty CRAs. This is important for low-income consumers because when they juggle bills, they often do so in consideration of multiple factors including credit reporting.

2. Non-financial alternative data or Big Data

Big Data encompasses a far-ranging array of information, including social media profiles, web browsing history and behavioral data. There are many unanswered questions about the predictiveness and accuracy of these data sources. For example, in one study in which NCLC employees ordered their own reports from four data brokers, these reports were found to be riddled with inaccuracies, ranging from the mundane (a wrong e-mail address or incorrect phone number) to the seriously flawed (wrong occupation or salary estimate).¹² Of the 15 records sought, some brokers had as many as 13 reports with errors in them.¹³

- **Educational or occupational attainment.** Use of education or occupation is quite troubling. There are obvious racial disparities in educational and occupational attainment. For example, 93% of non-Hispanic whites graduate from high school, but only 67% of Hispanics and 87% of African Americans do.¹⁴ Over 36% of non-Hispanic whites and

¹¹ Neil Bhutta, Paige Marta Skiba, & Jeremy Tobacman, Payday Loan Choices and Consequences, Vanderbilt Law and Economics Research Paper No. 12-30, October 11, 2012), <https://ssrn.com/abstract=2160947>, at 12.

¹² See Persis Yu, NCLC, Big Data: A Big Disappointment for Scoring Consumer Credit Risk, at 18, March 2014, available at www.nclc.org/images/pdf/pr-reports/report-big-data.pdf (reviewing the accuracy of reports by eBureau, ID Analytics, Spokeo, and Intelius).

¹³ *Id.*

¹⁴ Camille L. Ryan and Kurt Bauman, U.S. Census: Current Population Reports, Educational Attainment in the United States: 2015, March 2016, <https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf> (“In most

nearly 54% of Asians have a bachelor's degree, but less than 16% of Hispanics and 23% of African Americans do.¹⁵ As for occupation, 40% of non-Hispanic whites and 51% of Asians are employed in management, professional, and related positions, but only 30% of African Americans and 22% of Hispanics are similarly employed.¹⁶

The use of education and occupational attainment reinforces inequality, given that a consumer's educational attainment is most strongly linked with the educational level of his or her parents.¹⁷ Use of educational or occupational attainment would probably top the list of mobility-impeding data, and would ossify the gaping racial and economic inequality in our country.

- **Social media/friends & family data.** Use of social media profiles, particularly friend networks, raises serious concerns about racial disparities. Although not yet widely used in the United States, one company, Lenddo, which operates in twenty different countries, may deem a consumer to be less creditworthy if he or she is friends on Facebook with someone who was late paying back a loan to Lenddo.¹⁸ Location or geographic neighborhood is another way that creditors have based creditworthiness by association.¹⁹ Given the degree of residential housing segregation that exists in the U.S., location can function as a proxy for race and income and its use by creditors would reflect racial and socio-economic disparities. And most people's friends and family are likely to be of the same race, class, and cultural background.
- **Behavioral data, including web browsing.** Behavioral data includes information about how consumers interact with a web interface or answer specific questions, or data about how they shop, browse, use devices, or move about their daily lives. The use of behavioral data has also shown indications of racial bias, despite relying on seemingly

studies, parental education has been identified as the single strongest correlate of children's success in school, the number of years they attend school, and their success later in life.")

¹⁵ *Id.*

¹⁶ Bureau of Labor Statistics, BLS Reports, Report 1062: Labor Force Characteristics by Race and Ethnicity, 2015, September 2016, <https://www.bls.gov/opub/reports/race-and-ethnicity/2015/home.htm>.

¹⁷ Anna J. Egalite, How Family Background Influences Student Achievement, EducationNext, Vol. 16, No.2, Spring 2016, available at <http://educationnext.org/how-family-background-influences-student-achievement/>.

¹⁸ Katie Lobosco, Facebook friends could change your credit score, CNN.com (August 27, 2013) available at <http://money.cnn.com/2013/08/26/technology/social/facebook-credit-score/index.html>. See also Matt Vasilogambros, "Will Your Facebook Friends Make You a Credit Risk?" The Atlantic (August 7, 2015), <https://www.theatlantic.com/politics/archive/2015/08/will-your-facebook-friends-make-you-a-credit-risk/432504/>.

¹⁹ Jeffrey S. Morrison & Andy Feltovich, Leveraging Aggregated Credit Data and in Portfolio Forecasting and Collection Scoring, The RMA Journal, Oct. 2010, at 47, available at www.forecastingsolutions.com/publications/RMA_OCT2010.pdf (article written by Transunion researchers stating "...aggregated credit data is...helpful to [debt] collectors because it can identify local credit conditions clustered around common demographics. This is especially true for consumers with little or no credit history. For example, if the consumer is living in a ZIP code where the mortgage delinquency rates are climbing or always high, the chance for collection may be significantly less than for those in ZIP codes where the delinquency rate is relatively low and stable.").

racially neutral algorithms. In 2013, Latanya Sweeney, a professor of government at Harvard University, led a research project that concluded that Google searches of names more likely associated with black people often yielded advertisements for a criminal records search in that person's name.²⁰ On one website, searches of black-sounding names were 25 percent more likely to yield ads with offers to view the person's arrest or criminal record.²¹

C. Principles for use of bank account transaction data.

Recent efforts to use bank account transaction data have shown potential to help some thin file consumers. Bank account data can reveal whether a consumer has sufficient available funds to afford to take on a loan, since it includes both income and expense information, *i.e.* it can help show whether the consumer has ability to repay. Unfortunately, bank account information may not be able to help the 14 million consumers who lack a bank account.²²

Bank account data may avoid the need to rely on long historical timeframes and thus not consider negative marks from economic hardships from several years ago. Data suggests that many of the consumers with impaired credit were the victims of unfortunate events such as illness or job loss.²³ Bank account data can avoid the need to rely on long historical timeframes such as seven years, given that lenders review 12 months of statements at most when they manually review bank account activity.²⁴ Also, bank account data can show when there has been a healthy sustained recovery from an economic shock such as a job loss or illness.

Currently, the use of bank account data is voluntary, as consumers must actively give permission to a special type of company called a "data aggregator" to access to this information. The exception for the need for aggregators is the consumer's own bank, which obviously has access to the account information. Indeed, a CFPB study has speculated that that one of the primary "on ramps" to a credit report might be the consumer obtaining their first credit card from their

²⁰ Hiawatha Bray, Racial bias alleged in Google's ad results, *Boston Globe* (February 6, 2013) <https://www.bostonglobe.com/business/2013/02/06/harvard-professor-spots-web-search-bias/PtOgSh1ivTZMfyEGj00X4I/story.html>.

²¹ *Id.*

²² Federal Deposit Insurance Commission, 2017 FDIC National Survey of Unbanked and Underbanked Households, Oct. 2018, <https://www.fdic.gov/householdsurvey/2017/2017execsumm.pdf>, at 1.

²³ About 70 to 80% of consumers with impaired credit or a low score, such as a 600, will actually not default. These may be victims of extraordinary life circumstances who do not default again once they have recovered economically. See Chi Chi Wu, NCLC, Solving the Credit Conundrum: Helping Consumers' Credit Records Impaired by the Foreclosure Crisis and Great Recession, Dec. 2013, at 9-11, available at www.nclc.org/images/pdf/credit_reports/report-credit-conundrum-2013.pdf (summarizing research).

²⁴ For example, Fannie Mae requires lenders to review 12 months of bank account statements to establish payment activity. Fannie Mae Selling Guide, B3-5.4-03: Documentation and Assessment of a Nontraditional Credit History, August 30, 2016, available at <https://www.fanniemae.com/content/guide/selling/b3/5.4/03.html>. Anecdotally, we have heard that some lenders only require 3 to 6 months of bank account statements.

own bank.²⁵ The use of a data aggregator for account information allows this access even when a consumer does not have a deposit account at a large bank that also issues credit cards.

However, a data aggregator may access information far longer than consumers expect or may use or share the data in ways that the consumer did not affirmatively agree to. Thus, we urge that a consumer's permission to share bank account data for the purpose of credit underwriting always be time limited.

Another concern is that bank account data could be misused. Lenders and debt collectors could use it to figure out the best time to collect debts by analyzing when income comes in and can be grabbed. Bank accounts include sensitive information such as debit card purchases showing where the consumer shops or even which political causes they support. There will need to be methods to protect the privacy of this information while being able to use account information for credit decisionmaking. This could lead to discriminatory pricing not based on the individual's credit risk. Given how much sensitive information is in a bank account history, sharing should be strictly limited and disparate impacts carefully monitored.

Thus, we urge adoption of the following principles for use of bank account transaction data:

Consumer control

- **Consent must be real, knowing and meaningful.** It should never be buried in fine print. It must always be in a separate stand-alone document.
- **Consent should not be used to permit uses that consumers do not expect or understand.**
- **Consent must be limited by purpose.** A consent to use bank account data for credit underwriting should be limited to that use alone and should not permit the use of the data for other purposes such as marketing, debt collection, or government licensing.
- **Consent must be limited by data element.** A consumer should be able to choose sharing just cashflow information (credits, debits, balances) versus sharing cashflow plus the identities of merchants from debit card transactions or the identity of payors who make electronic deposits.
- **Consent should be time-limited and self-expiring.** A consent for credit underwriting should be a single use permission. A consent for account review for an open-end account should expire after one year and require renewal.

Consumer choice

- **Consumers should always have true choice in whether to share their bank account data.** There is too great a risk that creditors will require use of bank account transaction

²⁵ Consumer Financial Protection Bureau, Data Point: Becoming Credit Visible, June 2017, https://files.consumerfinance.gov/f/documents/BecomingCreditVisible_Data_Point_Final.pdf, at 33 (noting that “about 65 percent [of consumers studied], appear to have transitioned out of credit invisibility by opening an account by themselves despite their lack of a credit history” and that “perhaps some commercial banks are willing to lend to credit invisible consumers with whom they have existing deposit account relationships.”)

data for all consumers, including those who could have received credit without it. A consumer who already has a “fat file” and a good credit score should be able to rely on that alone without being required to share bank account information. Expansion into bank account information may benefit those consumers who have insufficient credit history information or lower credit scores, but could hurt or risk the privacy of consumers who already qualify for mainstream credit.

- **Consumers should never be required to share bank account transaction data for non-credit purposes**, such as employment, insurance, or government licensing. Needs-based government programs should be entitled to only a snapshot of current balances.

Meaningful benefit for consumers

- In order to ensure bank account transaction data benefits consumers, lenders should be urged to actually use consumer bank account transaction information supplied by data aggregators in a meaningful manner.

D. How should alternative data be used

The manner in which alternative data is used is important. One of the most important factors is consumer choice. If the use of alternative data is truly voluntary – that is, consumers make knowing and voluntary decisions to allow the use of the data and the data is used only for that limited purpose and in ways that consumers would expect – then it is much more likely to be helpful. This is why we would support the discussion draft of the Credit Access and Inclusion Act put forth by Representative Gottheimer, because unlike prior versions, it would permit the reporting of utility and rental payment information only when the consumer has provided written authorization, *i.e.*, only with consumer choice. We have in the past, and would in the future, oppose any versions that do not include a requirement for written consumer authorization, because of the potential for consumer harm.²⁶

Using alternative data to create special scores for otherwise unscorable consumers is preferable to the wholesale addition of the same data to traditional credit reports, where it might damage consumers who already have credit scores. A “second chance” score can give credit invisible consumers another shot to be seen, without potentially hurting some of the nearly 200 million consumers who do have a scoreable credit history.

A number of alternative scoring products have recently been unveiled that hold promise but must also be monitored. UltraFICO is a voluntary opt-in product that will rely on bank account transaction information from Finicity, a data aggregator working in partnership with Experian.²⁷ UltraFICO will only be used to enhance a consumer’s credit scores to see whether a denied application can be approved or a lower rate can be offered. ExperianBoost considers utility

²⁶ The discussion draft also does not preempt state consumer protection laws protecting the privacy of utility customers and hindering states from regulating tenant screening agencies, unlike prior versions. Such preemption is another reason we have opposed prior versions.

²⁷ FICO, Introducing UltraFICO, <https://www.fico.com/ultrafico/> (viewed July 21, 2019).

payments, but does so by reviewing bank account transactions that do not get included in traditional credit reports and is also voluntary opt-in.²⁸ FICO XD similarly is a second chance score using mostly telecom data from the National Consumer Telecom and Utilities Exchange, which is not included in traditional credit reports.²⁹

A significant issue with promising alternative products may be to get lenders to accept them. Many lenders have not even adopted FICO 9 or VantageScore models that simply lessen the impact of medical debt. There may need to be efforts to encourage lenders to consider alternative data when it is more predictive or beneficial to consumers than traditional credit reporting. A provision of Regulation B, which implements the Equal Credit Opportunity Act and is discussed further in Section E.2, may be helpful in this regard.

E. Existing Regulation of Alternative Data

When data is used in the decision whether or not to grant credit, its use should be regulated by the Equal Credit Opportunity Act – period. If that data is not the result of direct, firsthand experience between the lender and the consumer, it is also covered by the Fair Credit Reporting Act. These statutes provide the regulatory framework that governs the use of all data for credit decisionmaking, including alternative data.

The key requirements for any data used for credit decisions are that it be accurate, predictive and transparent. The regulatory schemes for the FCRA and ECOA speak to each of these requirements.

1. Fair Credit Reporting Act (FCRA)

The FCRA was intended to have a very broad scope of coverage. Information is a consumer report if it is:

- Used or expected to be used or collected in whole or in part to serve as a factor in establishing eligibility for consumer credit or other FCRA-covered purposes;
- Pertains to any of seven characteristics, which cover an extremely far-reaching range of information – credit worthiness, credit standing, credit capacity, character, general reputation, personal characteristics, and mode of living; and
- Issued by a third party that regularly assembles or evaluates such data for money or on a nonprofit cooperative basis.

Thus, almost all third-party data collected for credit decisionmaking purposes should be considered a “consumer report.” Unfortunately, several circuit courts have shown a reluctance to respect the plain language of the FCRA and its broad coverage.³⁰ We urge Congress to re-affirm

²⁸ Susan Henson, Experian, Introducing Experian Boost, a New Way to Instantly Improve Your Credit Scores, April 8, 2019, <https://www.experian.com/blogs/ask-experian/introducing-experian-boost/>.

²⁹ FICO, FICO Score XD, <https://www.fico.com/en/products/fico-score-xd> (viewed July 21, 2019).

³⁰ See *Kidd v. Thomson Reuters*, 925 F.3d 99 (2d Cir 2019) (CLEAR product was not a consumer report, despite state agency’s use for employment purposes, because Thomson Reuters had collected information

the broad scope of the FCRA and that it applies to any-third party data used for credit evaluation purposes.

FCRA protections are critical to protecting consumers when alternative data is used to evaluate them for credit. One of the key issues with alternative data is the level of accuracy of the data. As discussed above, our 2014 Big Data report revealed serious levels of inaccuracy in certain Big Data sources. The FCRA requires accuracy, in that Section 607(b) of the FCRA, 15 U.S.C. § 1681e(b), requires consumer reporting agencies (CRAs) to follow “reasonable procedures to ensure maximum possible accuracy.” Section 611(a) of the FCRA, 15 U.S.C. § 1681i(a), gives consumers the right to dispute any errors regarding information about them in a CRA’s files.

The FCRA also has specific notice requirements, which are intended to ensure transparency when information from a CRA is used. Mostly importantly, Section 615(a) and (h) of the Act, 15 U.S.C. § 1681m(a) and (h), require users of consumer reports to provide adverse action and risk-based pricing notices when information from a CRA has been used to deny them credit or charge them a higher price. This ensures that consumers are aware of the sources and types of information that are used against them in credit (and other) decisions, so that they are not left in the dark as to the reasons for decisions that may have critical consequences for their lives.

Furthermore, even if third party information is somehow not considered a consumer report, the FCRA includes a little-known provision that requires transparency in its usage. Section 615(b), 15 U.S.C. § 1681m(b), requires that lenders provide a specific notice if information that fits any of the seven characteristics listed in the definition of “consumer report” is obtained from a person other than a CRA and used to deny credit or charge more for it. This notice must inform the consumers of the right to make a written request for the reasons for the adverse action. Upon such a request, the user must disclose the nature of such information. Section 615(b) should apply to alternative data used for credit decisionmaking even if it somehow escapes the definition of a consumer report.

2. Equal Credit Opportunity Act (ECOA)

If alternative data is used for credit, there will be implications under the ECOA. Like the FCRA, the ECOA is a statute with a broad scope. It prohibits discrimination “with respect to any aspect of a credit transaction” on the basis of, *inter alia*, race, color, religion, national origin, sex or marital status, or age. 15 U.S.C. § 1691(a). “Credit” is broadly defined, as is the concept of

and intended it to be used only for non-FCRA purposes, expressly prohibited its sale or use for FCRA-related purposes, required users to make non-FCRA use certifications, and actively monitored compliance; entity must have a specific intent to furnish a “consumer report.”); *Zabriskie v. Fed. Nat'l Mortg. Ass'n*, 912 F.3d 1192 (9th Cir. 2019) (in a 2-1 decision, holding that Fannie Mae’s Desktop Underwriter program is not a CRA because its role is limited to providing software that allows lenders to assemble or evaluate information; majority ignored fact that it is DU itself which actually obtains information from various sources including nationwide CRAs and that DU itself issues a recommendation); *Fuges v. Southwest Title*, 707 F.3d 241 (3d Cir. 2012) (objectively reasonable for company that prepared reports on current owners of properties to interpret the reports as outside the FCRA because they allegedly pertained to the property and not to the consumer -- despite the fact the reports included information on judgments personally against the consumer).

“creditor,” which is not limited to banks or traditional lenders. 15 U.S.C. § 1691a(d) and (e). Finally, the ECOA is not limited to consumer credit but applies to certain types of business credit as well.

Most importantly for our purposes, Regulation B, which implements the ECOA, expressly notes that “legislative history of the Act indicates that the Congress intended an “effects test” concept ... be applicable to a creditor's determination of creditworthiness.” 12 C.F.R. § 1002.6(a). The effects test is another name for the disparate impact test, and the Official Staff Interpretations explain that the test:

may prohibit a creditor practice that is discriminatory in effect because it has a disproportionately negative impact on a prohibited basis, even though the creditor has no intent to discriminate and the practice appears neutral on its face, unless the creditor practice meets a legitimate business need that cannot reasonably be achieved as well by means that are less disparate in their impact.

Official Interpretations of Reg. B, 12 C.F.R. pt. 1002, supp. I, § 1002.6(a)-2.

As discussed in both Section B above and in Section G below, much of the data used for credit decisionmaking will have a disparate impact on protected classes. Thus, lenders must be extremely cautious in using alternative data given the effects test. This test has a three-step analysis which, teasing out the Official Staff Interpretations quoted above, consists of:

1. Does the practice have a disproportionately negative impact on a protected class even if it appears neutral on its face?
2. If so, does the practice meet a legitimate business need?
3. Can the same need be reasonably achieved using a less discriminatory alternative?

Like the FCRA, the ECOA also has specific notice requirements. It requires creditors to notify consumers of the action on an application. 15 U.S.C. § 1691(d)(1). If the creditor takes an adverse action, it must provide either a statement of reasons for the action or written notification of the right to such a statement. 15 U.S.C. § 1691(d)(2). This notice must be specific, and must meet the requirements of Regulation B and its corresponding Official Staff Interpretations.³¹

The notices required by the FCRA and ECOA raise one of the key issues with regards to the use of alternative models that rely on artificial intelligence or machine learning – transparency. Consumers are entitled to know not only *what* information is being used to assess them, but *how* that information is being used. Users of alternative modeling techniques will need to provide ECOA and FCRA adverse action notices, and to make sure the notices contain sufficient information to satisfy the ECOA standard that the notice disclose the specific reasons for the action taken.

³¹ Reg. B, 12 C.F.R. § 1002.9(b)(2); Official Interpretations of Reg. B, 12 C.F.R. pt. 1002, supp. I, § 1002.9(b)(2). See generally National Consumer Law Center, Credit Discrimination § 10.5.4.2 (6th ed. 2013), updated at www.nclc.org/library.

Having “black boxes” that evaluate creditworthiness should be a thing of the past, as a matter of both fairness and ensuring that consumers are fully educated about financial issues. Consumers should always be able know, access, and understand both what is the information being used in credit decisions and how it is used. They should be able to review the information for inaccuracies so they can dispute errors.

Finally, there is a seldom-used provision of Regulation B that requires lenders to consider creditworthiness information that is not part of a credit report, if requested by a consumer.³² This could be useful in encouraging lenders to consider alternative data. However, it is only a first step in that they are not required to treat such information with the same weight and in the same manner as credit report history information.

F. Alternative Data as Competition for the Credit Bureaus?

One of the biggest problems in credit reporting has been the lack of competition in that industry. The Big Three credit bureaus constitute an oligopoly that controls financial information vital for consumers’ economic lives. What’s worse, consumers aren’t even the customers – our data is the credit bureaus’ commodity – and we have no choice but to deal with these three companies.

Alternative data has the potential to provide competition to the credit bureaus by supplying new types of information that both lenders and consumers can choose to use instead of a credit report or score. But the credit bureaus have been on a buying spree, purchasing companies that gather alternative data. For example, Experian purchased Clarity while TransUnion purchased FactorTrust, both of which are consumer reporting agencies focused on subprime credit. In some cases, the credit bureaus form partnerships with alternative data providers to access their data. Equifax manages NCTUE, while Experian has a deal with Finicity, a bank account data aggregator. These trends could eliminate the possibility of real competition.

G. Algorithms as a Reflection on Society’s Inequities

Communities of color are not just more likely to be credit invisible - study after study has found that when minorities do have a credit score, as a group those scores are significantly lower than those of whites. A list of these studies is available in our policy brief, *Past Imperfect: How credit scores and other analytics “bake in” past discrimination and perpetuate it*.³³

³² 12 C.F.R. § 1002.6(b)(6)(ii)(creditor must consider any information the applicant presents to indicate credit history being considered does not accurately reflect applicant’s creditworthiness); Official Interpretations of Reg. B, 12 C.F.R. pt. 1002, supp. I, § 1002.6(b)(6)-1 (“On the applicant’s request, however, a creditor must consider credit information not reported through a credit bureau when the information relates to the same types of credit references and history that the creditor would consider if reported through a credit bureau.”)

³³ Chi Chi Wu, NCLC Past Imperfect: How Credit Scores and Other Analytics “Bake In” and Perpetuate Past Discrimination, May 2016, available at https://www.nclc.org/images/pdf/credit_discrimination/Past_Imperfect050616.pdf.

As discussed more in one of this our Policy Brief, the explanation for these disparities lies in the aftereffects of historical inequalities and the racial wealth gap. African American families own less than seven cents for every dollar in wealth owned by white families, while Latino households own less than eight cents for every dollar of white wealth. With fewer assets to draw on, people of color – and the friends and family to whom they might turn – are far less able to cushion the blows of financial catastrophes, such as job losses, income reductions, sickness, or unplanned expenses.

Despite the fact that credit scoring has a disparate impact, it is also predictive on an aggregate level. That is why it is legal under the ECOA – there is a legitimate business need for credit scoring given its predictiveness.

If alternative data is based on financial information, it will likely also exhibit racial disparities, for the same reasons.³⁴ In general, systems replicate themselves, and systemic racism keeps perpetuating itself whether or not individual animus-driven bias exists. This impacts the very foundations and institutions of our society – education, income, employment, housing, and criminal justice. This is why all the indicators discussed above – education, employment, income, and assets – exhibit such racial disparities.

Moreover, algorithms developed using data that exhibits racial disparities will themselves reproduce those disparities. There is an assumption that algorithms are automatically unbiased or judgment free, but recent research indicates otherwise.³⁵

Thus, because any economically based data is likely to show racial disparities, the Regulation B test for disparate impact will require a “legitimate business need” to use it for credit decisionmaking purposes. At a minimum, this means the data must be accurate and predictive.³⁶ But it should be more than just predictive – it should be more predictive than traditional data or should provide more access to credit (not less) to credit invisible or credit impaired consumers, and such credit should be affordably priced.

³⁴ See Carol Evans, Federal Reserve Board - Division of Consumer and Community Affairs, *Keeping Fintech Fair: Thinking about Fair Lending and UDAP Risks*, Consumer Compliance Outlook - Second Issue 2017 (2017), <https://consumercomplianceoutlook.org/2017/second-issue/keeping-fintech-fair-thinking-about-fair-lending-and-udap-risks/> (“[F]intech may raise the same types of fair lending risks present in traditional banking, including underwriting discrimination, pricing discrimination, redlining, and steering. Although some fintech trends may decrease certain fair lending risks, other trends could amplify old problems or create new risks.”) [hereinafter “Evans, *Keeping FinTech Fair*”]

³⁵ See Robert P. Bartlett, et al., *Consumer Lending Discrimination in the FinTech Era*, UC Berkeley Public Law Research Paper, December 7, 2017, <https://faculty.haas.berkeley.edu/morse/research/papers/discrim.pdf> (finding that fintech lenders discriminate, albeit 40% less than face-to-face lenders). See also, Evans, *Keeping FinTech Fair* (“while statistical models have the potential to increase consistency in decision-making and to ensure that results are empirically sound, depending on the data analyzed and underlying assumptions, models also may reflect and perpetuate existing social inequalities. Thus, big data should not be viewed as monolithically good or bad, and the fact that an algorithm is data driven does not ensure that it is fair or objective.”).

³⁶ *Id.* (“Generally, the more speculative the nexus with creditworthiness, the higher the fair lending risk”).

Also, even if predictive, there is certain data that is troubling because it entrenches and perpetuates inequality in such an obvious and stark way. One obvious example is educational attainment level, which as discussed in Section B.2 is highly correlated with parents' income and education. Another type of data would be the consumer's geographic neighborhood, given the obvious racial and economic segregation in housing in this country. Using these factors would only worsen the already herculean barriers that disadvantaged children face in making economic progress. And beyond race, we know that income and wealth inequality are already increasingly stark; using factors like educational level or neighborhood would just make it worse.

The danger of gathering massive amounts of data and using algorithms to analyze them is that they reinforce and entrench existing inequality, whether it be racial or economic. The American dream has always been that the child of a factory worker could be the CEO of a corporation, a Senator, or even President. That dream is already dying a slow death in this country. Using massive amounts of data, chopped up and analyzed, may be the nail in the coffin of that dream.

We know that credit reports and scores can reinforce existing inequality. The question is whether we treat new sources of data, *i.e.*, alternative data, in the same way or whether we develop algorithms and policies that allow the American dream to flourish once again. Data models, whether traditional or alternative, could be programmed to reduce racial and economic disparities while maintaining – and hopefully, improving -- predictiveness. This could address the third prong of the disparate impact test – “a less discriminatory alternative.”³⁷

We have a chance with new data sources and models to do better. The question is whether we will do so or whether we will contribute to the gaping inequality in our society.

* * * * *

Thank you again for the opportunity to provide my views to the Task Force today. I look forward to your questions.

³⁷ For example, one modification proposed by researchers would require including minority status as a “control variable” during the development of a credit scoring model. Michael LaCour-Little and Elaine Fortowsky, *Credit Scoring and the Fair Lending Issue of Disparate Impact* in *Credit Scoring for Risk Managers: The Handbook for Lenders* (Elizabeth Mays ed. South-Western Educational Pub. 2003).



July 23, 2019

The Honorable Stephen F. Lynch
Chairman
Task Force on Financial Technology
Financial Services Committee
U.S. House of Representatives
Washington, DC 20515

The Honorable French Hill
Ranking Member
Task Force on Financial Technology
Financial Services Committee
U.S. House of Representatives
Washington, DC 20515

Dear Chairman Lynch and Ranking Member Hill:

I would like to thank the Financial Services Committee for convening its Task Force on Financial Technology and organizing its hearing titled "Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit" on July 25, 2019. I am writing to express my views regarding the topic of that hearing.

My name is Dan Quan and I am an adjunct scholar at the Cato Institute's Center for Monetary and Financial Alternatives. I also advise high-growth, disruptive fintech companies. Previously, I led the Consumer Financial Protection Bureau's (CFPB) fintech office, Project Catalyst.

The United States has the most developed and competitive consumer credit market in the world. For example, the credit card industry has seen robust growth since the financial crisis, with \$1.07 trillion in outstanding balances as of May 2019.¹ Fintech lending has seen even greater growth, accounting for 38% of the \$138 billion unsecured personal loan market in 2018.² However, the inconvenient truth buried in those rosy numbers is that we still have a huge financial inclusion challenge. 45 million Americans, or 19.3% of the adult population, don't have access to credit.³ These consumers tend to be disproportionately African American, Hispanic, young, and low-income.⁴ Additionally, 53% of small businesses surveyed by the Federal Reserve could not obtain the full financing they sought.⁵

Now, however, the use of alternative data in credit scoring and underwriting holds great promise for bringing greater access to credit and capital to struggling consumers and small businesses.

¹ Board of Governors of the Federal Reserve System, "Consumer Credit G.19," July 8, 2019, <https://www.federalreserve.gov/releases/g19/current/>

² Kate Rooney, "Fintechs Help Boost US Personal Loan Surge to a Record \$138 Billion," *CNBC*, February 21, 2019, <https://www.cnbc.com/2019/02/21/personal-loans-surge-to-a-record-138-billion-in-us-as-fintechs-lead-new-lending-charge.html>

³ Consumer Financial Protection Bureau Office of Research, "Data Point: Credit Invisibles," May 2015, https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf

⁴ Consumer Financial Protection Bureau, "Who Are the Credit Invisibles? How to Help People with Limited Credit Histories," December 2016, https://files.consumerfinance.gov/f/documents/201612_cfpb_credit_invisible_policy_report.pdf

⁵ U.S. Federal Reserve, "Small Business Credit Survey: 2019 Report on Employer Firms," April 16, 2019, <https://www.fedsmallbusiness.org/medialibrary/fedsmallbusiness/files/2019/sbcs-employer-firms-report.pdf>

Underwriting models that include alternative data can increase lending volume, lower interest rates for borrowers, and improve the accuracy of default predictions. In short, alternative data can make lending more plentiful, more affordable, and sounder—with historically underserved borrowers and communities benefitting most.⁶

Alternative data can include anything that is not currently part of consumers' traditional credit reports. It can range from account transaction history (known as cash-flow underwriting), to educational and occupational information, to social media use and other online or mobile activities. The topic of alternative data sometimes raises concerns about whether it could increase discrimination against protected classes or intrude on borrowers' privacy. While these are legitimate concerns, it is counterproductive to prevent or tightly constrain the use of alternative data in lending. Restrictions would likely hurt, not help, marginalized borrowers who are overwhelmingly low-income and minorities. In fact, a recent study on the racial and ethnic disparities in credit access finds that the lending gap attributed to credit discrimination is much smaller for fintech firms than it is for traditional lenders.⁷ Discouraging innovation out of a concern for potential discrimination is therefore likely to undermine financial inclusion.

In my experience, not many firms use alternative data in credit scoring and underwriting today. Those that do mostly use cash flow data. For example, fintech lenders such as Oportun use cash flow data to provide credit to “unscorable” or “credit invisible” consumers. Other lenders, like payment processors Square and PayPal, use transaction histories to help them lend to their merchants effectively and efficiently. According to Square, its average loan size is \$6,000,⁸ an amount that most traditional financial institutions find unprofitable to finance. Individual consumers can opt to include data like their on-time bill payment history to boost their FICO scores.⁹ In my conversations with fintech lenders, none report any convincing evidence that social media data can predict consumers' repayment behavior. No lender is using it for underwriting purposes in the U.S.

A recent paper from the Federal Reserve Bank of Philadelphia agrees that underwriting with alternative data can better predict loan outcomes, resulting in improved terms for borrowers who, under traditional credit criteria, would receive higher-priced loans.¹⁰

There is no need for new laws or rules just for the use of alternative data. The same laws and rules that ensure fair credit access, privacy protections, and transparency in underwriting decisions equally apply to lenders who use alternative data. On the other hand, regulatory agencies should encourage the responsible use of alternative data by providing greater clarity to lenders in two key areas.

⁶ Julapa Jagtiani and Catherine Lemieux, “Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks?,” Federal Reserve Bank of Philadelphia Working Paper 18–13, March 2018, <https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2018/wp18-13.pdf>.

⁷ Robert Bartlett, Adair Morse, Richard Stanton, and Nancy Wallace, “Consumer Lending Discrimination in the Fintech Era,” National Bureau of Economic Research Working Paper No. 25943, June 2019, <https://www.nber.org/papers/w25943>.

⁸ Lalita Clozel, “The Story Behind Square’s Bank Charter Application,” *American Banker*, September 19, 2017, <https://www.americanbanker.com/news/the-story-behind-squares-bank-charter-application>.

⁹ Herb Weisbaum, “There’s a New Way to Instantly Boost Your Credit Score,” *NBC News*, February 5, 2019, <https://www.nbcnews.com/business/consumer/there-s-new-way-instantly-boost-your-credit-score-n967041>.

¹⁰ Julapa Jagtiani and Catharine Lemieux, “The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence of LendingClub Consumer Platform,” Federal Reserve Bank of Philadelphia Working Paper 18–15, April 2018 (Revised January 2019), <https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2018/wp18-15r.pdf>.

First, a transparent, secure, and frictionless data sharing ecosystem is necessary for increasing the use and reliability of alternative data. Regulators should support the Treasury Department's position that consumers have a right to permission their own financial data for third-party use.¹¹

Second, the CFPB needs to issue clear guidance on what responsibilities under the Fair Credit Reporting Act that third-party data aggregators and financial institutions have when consumers request to share their financial data through a secure mechanism such as an API. Uncertainty about these responsibilities creates confusion, friction, and can potentially result in consumer harm.

Additionally, the private marketplace, through industry and consumer group collaboration, can work together to ensure consumers fully understand their data rights, including how lenders will use their data, before they consent to sharing them.

The use of alternative data will enhance the affordability of credit and make our credit system more inclusive. While policymakers should remain vigilant, they must also stay open-minded about how to encourage further developments in market innovation that will benefit consumers, small businesses, and the economy at large.

I appreciate the opportunity to comment on this important hearing.

Sincerely,



Dan Quan
Adjunct Scholar
Center for Monetary and Financial Alternatives
Cato Institute

¹¹ U.S. Department of the Treasury, "A Financial System That Creates Economic Opportunities, Nonbank Financials, Fintech, and Innovation," Report to President Donald J. Trump, July 2018, https://home.treasury.gov/sites/default/files/2018-08/A-Financial-System-that-Creates-Economic-Opportunities---Nonbank-Financials-Fintech-and-Innovation_0.pdf.

Domino: A Blog About Student Debt

Educational Redlining?

The use of education data in underwriting could leave HBCU and MSI graduates in the dark

By Aryn Bussey | July 24, 2019

Access to credit can open doors to economic opportunity and determine consumers' ability to own a home, start a small business, or even purchase a vehicle. Millions of people, however, are shut out of today's credit system and the opportunities that come with it—particularly borrowers of color. According to the Federal Reserve (<https://www.federalreserve.gov/publications/2017-economic-well-being-of-us-households-in-2016-banking-credit.htm>), Black and Hispanic consumers within every income bracket are more likely than their white peers to be denied credit or offered less credit than requested. For years, policymakers have weighed the use of alternative data like cell phone or utility payment history to help expand marginalized communities' access to credit.

Tomorrow, the House Financial Services Committee's Fintech Task Force is holding a hearing (<https://financialservices.house.gov/calendar/eventsingle.aspx?EventID=404143>) on the use of alternative data to expand access to credit. It is critical that the Committee examine what I believe is a dangerous and discriminatory trend—the use of education data for determining creditworthiness under the guise of "innovation."

Recently, financial services companies have begun (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3349478) to explore using education data for determining creditworthiness. Education data like the type of institution attended (<https://lendedu.com/blog/student-loan-underwriting-processes>), a student's area of study (<https://www.bloomberg.com/news/articles/2019-04-09/college-grads-sell-stakes-in-themselves-to-wall-street>), a school's average SAT score, or even whether an applicant's parents completed college (<http://www.aei.org/wp-content/uploads/2017/02/Pricing-Without-Discrimination.pdf>) have been proposed, and even used, to determine credit eligibility and pricing. The use of this data is touted by industry as a way to expand access to credit for underserved communities. However, policymakers should consider how the use of education data in determining credit will further disadvantage the communities it is purported to help.

As a proud third-generation HBCU graduate, former legislative and executive branch education policy staffer, and a Black female consumer this approach deeply troubles me. America has a longstanding history of discrimination and inequality in education, ranging from access and affordability to predatory lenders (<https://www.reimanlaw.com/news-RSHSettlement>) and repayment (https://www.washingtonpost.com/education/2019/06/05/civil-rights-groups-urge-consumer-bureau-root-out-discrimination-student-loan-servicing-industry/?utm_term=.fc01246a3c7). And now, lenders are seeking to introduce another potentially discriminatory practice.

Use of education data promotes social stratification

The use of education data in credit decisions is troublesome given the history of disparate access to education in America. From standardized testing (<https://www.brookings.edu/research/race-gaps-in-sat-scores-highlight-inequality-and-hinder-upward-mobility/>), to degree attainment (<https://www.insidehighered.com/news/2018/06/14/states-struggle-close-degree-attainment-gaps-black-latino-students>), to exclusionary enrollment at "elite schools" (<https://www.nytimes.com/interactive/2017/08/24/us/affirmative-action.html>)" countless Americans have been deemed unworthy (<https://www.thenation.com/article/affirmative-action-state-universities/>) of an equitable education simply because of the color of their skin (<https://www.11alive.com/article/news/when-the-state-said-no-they-changed-history/85-cd0af775-e752-463e-ba4c-35efcd7d6818>) or the socioeconomic status of their families.

For example, although degree attainment is on the rise for many racial and ethnic groups, research shows (<https://www.americanprogress.org/issues/education-postsecondary/reports/2016/10/13/145098/closed-doors-black-and-latino-students-are-excluded-from-top-public-universities/>) there is a shortage of minority students, particularly African-American and Latino students, at selective institutions of higher education. Only nine percent of Black students, eight percent of Indigenous American students, and twelve percent of Latino students attend America's most elite public universities. When credit terms are tied to attendance at supposedly "elite" institutions, it can unfairly impact borrowers of color. Widespread adoption of educational criteria to determine creditworthiness will further stratify socioeconomic barriers to economic opportunity and mobility for Black and Brown consumers.

The use of education data in credit decisions also ignores the way most students actually select a college—more than 57 percent of incoming freshman attending public, four-year colleges enroll within 50 miles of home. Many of these students choose to stay close to home (<https://www.acenet.edu/news-room/Documents/Education-Deserts-The-Continued-Significance-of-Place-in-the-Twenty-First-Century.pdf>) for college because of family responsibilities, cultural norms, or factors related to working while enrolled in school. This is especially true for Latino, Black, and Native American (<https://www.acenet.edu/news-room/Documents/Education-Deserts-The-Continued-Significance-of-Place-in-the-Twenty-First-Century.pdf>) students. As a result, the use of institutional factors in credit decisions may have a disparate impact on Black and Brown communities.

Use of education data implicates fair lending law

Furthermore, given the gross disparities that exist across the education spectrum, the use of education data has significant fair lending implications. The Equal Credit Opportunity Act prohibits creditors from discriminating against applicants in any aspect of a credit transaction on the basis of characteristics such as race, color, and national origin.

A 2017 American Enterprise Institute report (<http://www.aei.org/wp-content/uploads/2017/02/Pricing-Without-Discrimination.pdf>) found that use of education criteria in underwriting decisions "tend to disparately affect protected classes of individuals," also noting that "the quality or selectivity of the institution has only a slight effect on earnings." Furthermore, the report notes that there is no available data that substantiates the necessity of including these factors in underwriting decisions or its relationship to the likelihood of repayment. This is cause for alarm where the risk of discrimination is so great.

For example, back in 2007 (https://www.nytimes.com/2007/06/19/us/19loans.html?_r=1&oref=slogin), then-Attorney General Andrew Cuomo warned Congress about this educational "redlining" (http://www.nbcnews.com/id/19316230/ns/business-personal_finance/t/cuomo-charges-redlining-student-loan-probe/#.XTaKRpNKjOR). An investigation found that private student lenders ranked colleges and universities by default rates and used this data to set interest rates. As part of the investigation, Cuomo found that borrowers with "excellent" credit attending non-selective institutions were given higher interest rates than borrowers with less than "stellar" credit attending selective institutions. Cuomo noted, "just as lenders in the mortgage industry once made judgments about credit lending in entire neighborhoods as a whole, so too are lenders making generalized judgments about student and parent credit risk based on a student's 'school neighborhood.'"

Note that multiple federal enforcement (https://files.consumerfinance.gov/f/201207_cfpb_Reports_Private-Student-Loans.pdf) agencies (<https://www.fdic.gov/news/news/press/2014/salliemae.pdf>) have now stated that certain educational based cohort-related factors create discriminatory ECOA risk or violations.

Lawmakers must scrutinize the use of this data before it's too late

Despite the buzz being generated by fintech companies, we must remember that all innovation is not inherently good for consumers. Subprime lending was first marketed in the early 2000s (https://www.federalreservehistory.org/essays/subprime_mortgage_crisis) as an innovative financial product to serve underserved borrowers. Before that, payday loans were touted as a necessary product (<https://www.pewtrusts.org/en/research-and-analysis/articles/2012/07/a-short-history-of-payday-lending-law>) to allow vulnerable consumers quick access to small dollar loans. Now, credit products using education data are being marketed as innovative products to expand access to credit.

We must also remember that this is broader than private student loans—this is broader than student debt. The ramifications of education data in underwriting extend across consumer financial products, and therefore across consumers' lives.

It is frightening to imagine a future where graduates of my alma mater and those alike could be thrown into a second class credit market simply because of the college we attended. We must not further the stratification of economic and racial disparities in this country in the name of purported innovation. It is important that our lawmakers seriously examine a system that could potentially penalize a second-generation Latina student who opts for the MSI (minority-serving institution) close to home so she can help care for her sick grandmother instead of going to an Ivy League institution on the other side of the country. We must ensure that a former foster youth is not penalized when buying his first home simply because he chose to become a teacher.

In closing, we cannot allow a future where borrowers—and in particular borrowers of color—pay more for, or are denied credit, simply based on who sits next to them in a classroom. I strongly urge lawmakers and regulators to explore the risk to consumers and potential biases that education data may introduce to the credit market. If we allow these practices to take root, we risk further marginalizing Black and Brown borrowers.

###

Aryn Bussey is an education policy expert and strategist. Aryn credits her efforts to create the White House Initiative on Educational Excellence for African-Americans as her greatest professional accomplishment to date.



<http://www.fdata.global/north-america>

VIA ELECTRONIC SUBMISSION

July 24, 2019

The Honorable Stephen Lynch
Chairman
Task Force on Financial Technology
House Committee on Financial Services
2129 Rayburn House Office Building
Washington, DC 20515

The Honorable French Hill
Ranking Member
Task Force on Financial Technology
House Committee on Financial Services
2129 Rayburn House Office Building
Washington, DC 20515

Dear Chairman Lynch and Ranking Member Hill:

The Financial Data and Technology Association of North America (“FDATA North America”) appreciates this opportunity to submit a letter for the record for the House Financial Services Committee Task Force on Financial Technology’s hearing “Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit.” As the leading trade association for permissioned data aggregation and the significant benefits that putting consumers and small businesses in control of their own financial data can bring to the market, FDATA North America and its member companies strongly believe in the ability of technological innovation to empower consumers and to provide for greater financial inclusion and access.

FDATA North America was founded in early 2018 by several firms whose technology-based products and services allow consumers and small businesses to improve their financial wellbeing. We count innovative leaders such as Axxess, Betterment, Envestnet Yodlee, Flinks, Intuit, Kabbage, Lendified, Mogo, Moven, Morningstar, MX, Petal, Plaid, Questrade, Quicken Loans, The ID Co., TransUnion, Wealthica, and others as our members. We are a regional chapter of FDATA Global, which was the driving force for Open Banking in the United Kingdom and which continues to provide technical expertise to regulators and policymakers and to regulatory bodies internationally contemplating, designing, and implementing open banking frameworks.

There are many components to cash-flow underwriting. Though the task force is principally focused, for the purposes of this hearing, on alternative data, FDATA North America would respectfully recommend a related but separate issue to the task force’s attention. As technology and innovation in financial services have advanced over the last several years, numerous fintech companies, many of whom are FDATA North America members, now utilize

permissioned financial transaction data, accessed through data aggregation, into cash-flow underwriting practices, which can allow a lender to see a more holistic financial picture of a credit applicant and, critically, to provide or expand credit access and to lower the cost of credit for individuals or small businesses. With as many as 45 million Americans who are either credit invisible or unscorable¹, this process provides individuals or small businesses with more opportunities to safely access more affordable credit, even in the absence of a traditional credit score or history, based on an analysis of their own transaction data.

The adoption of the Fair Credit Reporting Act (“FCRA”) in 1970 sought to ensure the accuracy, fairness and privacy of consumer information when used for financial decision-making. The goal of the FCRA was to provide transparency to consumers in an opaque system of credit data and a means to allow consumers to have incorrect data contained in their credit file corrected. The FCRA provided important tools that allowed consumers to take increased control of the accuracy of their credit data by providing more visibility into the credit reporting process. To achieve this outcome, a regime was created to ensure that financial institutions, data furnishers, creditors, and the credit bureaus ensured that the data being collected, transmitted and used for a credit decision about a consumer is accurate.

The rise of financial technology, fueled by consumer-permissioned data aggregation, and the notion that the consumer is in full control of their financial information, was not considered when the FCRA was developed and implemented. A key element of cash-flow underwriting is the ability of the consumer to opt-in to sharing elements of their financial transaction data in addition to or instead of traditional credit reporting data. Transaction data, to which the consumer or small business has, under existing law, full access to and transparency into at all times, is materially different, both in process and substance, than the data elements collected about a consumer from their creditors for the purposes of credit reporting. The FCRA, however, is silent on this issue.

As more lenders, financial technology firms, and consumers recognize the value of using their own transactional data to demonstrate their creditworthiness, there is a significant need for clear guidance that supports the ability of end users to access and share their own financial data and, importantly, prevents the misapplication to this market of regulations written for consumer reporting agencies.

Already some financial institutions are seeking to block consumers from sharing their transactional data with lenders unless it is through a credit reporting agency. But permissioned transaction data is substantively different than credit reporting data. Importantly, end users are continuously in control of their financial transaction data, can monitor it in real time, have the ability to correct erroneous information quickly, and are providing their consent to a technology-powered lender to obtain their data. For all of these reasons, FDATA North America submits that the objectives of the FCRA are already embedded in the cash flow underwriting market and respectfully suggests that the task force should ask the Consumer Financial Protection Bureau

¹ https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf

("CFPB") to provide clarity on this point across the market, so that no market participant withholds from its customers the ability to take advantage of cash-flow underwriting applications.

In today's financial market, consumers are progressively more aware of the technology available to help them increase their financial wellbeing and grow their credit. The FCRA, which originally sought to ensure transparency, never envisioned a time when permissioned data could be used in cash-flow underwriting and was intended only to apply to financial data collected about consumers from their creditors. The absence of clear guidance on end users' right to access and share their data to obtain a loan creates barriers for lenders to serve their customers and provides restrictions for consumers and businesses alike wishing to take advantage of cash-flow underwriting tools, despite their explicit consent. For the sake of greater financial inclusion and access, and more competitive pricing and lending products, FDATA North America believes guidance is required from the CFPB to ensure that consumers continue to have the right to use their own financial transaction data for underwriting tools.

Thank you for once again for holding this important hearing.

Sincerely,

A handwritten signature in black ink, appearing to read "S. Boms", with a long horizontal line extending to the right.

Steven Boms
Executive Director
FDATA North America



Fair Isaac Corporation
200 Smith Ranch Road
San Rafael, CA 94903 USA
T 415 446 6000
F 415 492 9381
www.fico.com

August 8, 2019

The Honorable Stephen Lynch
Chairman, Task Force on Financial Technology
2109 Rayburn House Office Building
Washington, DC 20515

The Honorable French Hill
Ranking Member, Task Force on Financial Technology
1533 Longworth House Office Building
Washington, DC 20515

Dear Chairman Lynch and Ranking Member Hill:

On behalf of FICO, thank you for holding the recent Task Force on Financial Technology hearing entitled "Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit." We believe this is an important topic and that the use of alternative data in credit underwriting holds great promise for a wide range of stakeholders.

As you may know, FICO is not a credit bureau, like Equifax, TransUnion, or Experian, nor is it owned by the credit bureaus, like VantageScore. Rather, we are an independent analytics company. Since the company's founding in 1956, FICO has been a pioneer in the credit scoring market developing new innovations that provide benefits to lenders, investors and consumers through the facilitation and expansion of access to credit in a safe and responsible manner. We are especially proud of our financial inclusion initiatives both here in the U.S. and in 30 countries around the world. Much of this work is focused on leveraging new data that lies outside traditional credit bureau files. It is this information that is widely defined as "alternative data."

FICO introduced the FICO® Score in the United States in 1989. For the past thirty years, FICO has improved its predictive power by creating new versions that employ innovative analytic techniques to extract additional insights out of the data residing at the credit bureaus. The latest version, FICO® Score 9, is the most predictive FICO Score ever developed, providing lenders, consumers, investors, and other stakeholders with a myriad of benefits not available elsewhere. FICO Score 9 was built with the latest scoring technology, including our leading edge multi-faceted modeling technique to enhance risk prediction in mortgage originations, and has incorporated a more consumer friendly, predictive, and differentiated approach to the treatment of medical collections. Paid collections are not considered by FICO Score 9.

However, after three decades of enhancing and improving the FICO Score, we believe that the real promise for expanding access to credit is by responsibly drawing on reliable data outside the credit bureaus. FICO has made a significant, years-long investment in developing a sound and innovative approach to using alternative data to expand access to credit.

In 2016, FICO introduced a new product that leverages Fair Credit Reporting Act ("FCRA")-compliant alternative data sources to generate reliable FICO Scores for consumers who cannot be scored using credit bureau data alone. The introduction of this new product is novel, as it considers bill payment data and select public records data, in addition to whatever traditional credit bureau data exists, to score

The Honorable Stephen Lynch and French Hill
August 8, 2019
Page 2

more than 26.5 million previously unscored consumers—more than 50 percent of those previously unscored.

It is important to note that when discussing alternative data as it relates to telecommunication and utility payment information, the FICO Score, since its inception in 1989, has always considered this data when present in the consumer files at the credit bureaus. However, there is very little of this information present at the credit bureaus today. Our latest research reveals that only 5% of consumers have telco payments in their credit bureau files and even less, 2.5%, have utility information. Rental data, which FICO Score 9 considers and is often discussed as a differentiator, is present for only 2.3% of consumers with credit bureau files. These facts underscore the importance of leveraging new FCRA-compliant data sets that reside outside of the credit bureaus.

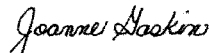
In 2018, FICO took another step in identifying new ways to expand access to credit in a safe and responsible manner through a new innovation that enables consumers to opt-in and permission access to their bank account information (checking, savings or money market) which can enhance their credit score. In a typical scenario, a lender will receive an applicant's FICO Score which may not qualify the applicant for credit or, alternatively, provide for a favorable interest rate. A lender using FICO's newest innovation can offer the applicant the ability to permission the applicant's bank account information which can be used to gain additional insights by the lender in the form of a new FICO credit score which combines the existing credit bureau data with the additional bank information.

To date, FICO's work with alternative data has resulted in our analytic scientists examining many data sets for potential use. FICO is focused on not only incorporating information that improves the predictive value of the model, but that also results in a broad-based, FCRA-compliant model that meets all the requirements of highly regulated financial institutions. As a result, FICO employs a rigorous six-point test to determine whether alternative data is worthy of inclusion in a FICO® Score model.

Finally, as the Task Force and the House Financial Services Committee continue to examine alternative data and its use, FICO believes that efforts that further the accurate and compliant furnishing of this data, like the Credit Access Inclusion Act of 2019, merit support.

Again, thank you for convening the hearing and for the opportunity to submit this letter for the record. Please do not hesitate to contact me or Daniel Nestel, FICO Senior Director of Government Relations, should you have any questions or need additional information.

Sincerely yours,



Joanne Gaskin
Vice President, Scores and Analytics



Barrett Burns President & CEO barrettburns@vantagescore.com

August 2, 2019

The Honorable Stephen F. Lynch
Chairman, Task Force on Financial Technology
Committee on Financial Services
U.S. House of Representatives
Washington, D.C. 20515

The Honorable French Hill
Ranking Member, Task Force on Financial Technology
Committee on Financial Services
U.S. House of Representatives
Washington, D.C. 20515

Dear Chairman Lynch and Ranking Member Hill:

The hearing held on July 25th by the House Financial Services Committee's Task Force on Financial Technology focused on many critically important issues regarding the use of alternative data in credit scoring and underwriting to expand access to credit. There are many definitions of alternative data covering a broad range of categories. There seems to be an emerging consensus that the use of at least certain categories of alternative data can facilitate greater access to credit for millions of consumers who are, in fact, creditworthy, but nonetheless fail to qualify for affordable mainstream credit if credit scores derived from the conventional credit scoring models are used.

Private sector researchers, as well as those in government agencies such as the Consumer Financial Protection Bureau ("CFPB"), often refer to these individuals as the "credit invisibles." There is also a significant subset of consumers with "thin" credit files (i.e., with less than three credit trades on their file) who can also benefit from the reporting of additional data such as rent, cell phone and utility payments. Use of these categories of data may move such consumers to "thick" file status. Since many lenders will not consider applicants with thin file credit histories, some consumers can enjoy the double benefit of additional predictive data and another trade line.

One of the primary reasons why VantageScore Solutions, LLC, was established in 2006 was to develop a consumer credit scoring model that would be more predictive, score more consumers and be more inclusive than models previously or then available. In pursuing that goal, through each generation of VantageScore we have built in the capacity to consider rent, utility and cell phone payments; we do consider such data when it is present in the credit file. Small quantities of such data are currently contained in the core credit files at each of the national Credit Reporting Companies (CRCs) and, as such, are considered by all versions of VantageScore. We have found these data to be useful in predicting consumer behaviors, particularly in assessing experience with or "depth of" credit. We do not, however, consider such data to be "alternative" as they are included as part of the core credit file of the consumer reported through the CRCs, represent concrete payment obligations and because VantageScore has found them predictive since building its first model thirteen years ago.



We work only with credit file data precisely because of the stringent regulations (i.e., ECOA and FCRA) and data standards (e.g., quality, accuracy, standardization, and universality) that apply to such data, but do not regulate other information that is not contained in credit files, such as “social” data or certain other types of alternative data.

Our approach recognizes the concern alluded to by one of the witnesses at the Task Force hearing, Lawrence Evans of the GAO, who in his written testimony referred to a serious concern raised by Carol Evans, Associate Director of the Federal Reserve Board’s Division of Consumer and Community Affairs, when she wrote:

... it’s been reported that some lenders consider whether a consumer’s online social network includes people with poor credit histories, which can raise concerns about discrimination against those living in disadvantaged areas. Instead of expanding access to responsible credit, the use of data correlated with race or national origin could serve to entrench or even worsen existing inequities in financial access. Finally, it is important to consider that some data may not appear correlated with race or national origin when used alone but may be highly correlated with prohibited characteristics when evaluated in conjunction with other fields.¹ [Internal footnote omitted]

By consistently following this modelling protocol relating to rent, utility and cell phone payments outlined above, VantageScore credit scoring models are used in every part of the “credit process” where credit scores are used. Moreover, we have pursued the goal of scoring as many credit files as can responsibly and predictably be scored by leveraging new modeling techniques and moving beyond legacy thinking. Our latest two models, VantageScore 3.0 and 4.0, generate a score for approximately 96% of consumers with a file at one (or more) of the three CRCs.

Just as it is extremely important to be accurate and precise in developing credit scoring models and computing consumer credit scores, it is also very important to use language about credit score model development and credit scores with accuracy and precision. Consider the following:

- Frequently throughout the Task Force hearing participants seemed to overlook the distinction between “credit score models” and “credit scores” and conflated the two into one or used the terms interchangeably.
- VantageScore Solutions is solely a “model developer”; we do not provide a “score” but rather an algorithm or credit scoring model to which the CRCs then append credit file data to determine consumer credit scores. The algorithm is built on historical data samples that do not contain any personal identifiable information. We do not have access to any personal identifiable information, nor any data links to live CRC consumer credit files, and therefore we do not have the data needed to run through our models to

¹ <https://www.consumercomplianceoutlook.org/2017/second-issue/keeping-fintech-fair-thinking-about-fair-lending-and-udap-risks/>



calculate an individual's score. In short, we are a model developer, and not a credit bureau that delivers a credit score when credit file data is appended to a model.

- It is also critically important for policymakers and witnesses to use the all-encompassing generic term "credit score" when it is their intention to describe all credit scores, rather than using inaccurately the name of any single proprietary brand of credit score. If a specific brand is used to refer to all credit scores, they inadvertently both exclude from the discussion a significant subset of the universe they intend to include and simultaneously engage in what might be considered a "brand endorsement."

That last point is particularly significant in a marketplace which once was the domain of a single credit score model developer (the Fair Isaac Corporation) and the term "FICO score" could then be used interchangeably with the term "credit score"; but that certainly is no longer the case. For 13 years we have had a competitive credit score marketplace in which millions of consumers who are unscorable using the legacy FICO credit scoring models are, in fact, now scoreable using VantageScore credit scoring models. To illustrate that point, consider the statement contained on page 2 of the National Consumer Law Center's written testimony, where NCLC states: "According to the Consumer Financial Protection Bureau (CFPB), 26 million Americans (or about 1 in 10) do not have a credit history...." That may have been true more than a decade ago, and even today, such consumers may lack a credit history when viewed through the prism of the legacy FICO model.

However, in today's competitive credit scoring marketplace, the reality is that VantageScore Models 3.0 and 4.0 are able to accurately score the overwhelming majority of those designated by researchers at the CFPB as "credit invisibles" if legacy FICO models are used as the sole arbiter of creditworthiness and scoreability.

Extensive and continued testing demonstrates that the scores assigned to these consumers using VantageScore credit scoring models are predictive. Indeed, tests made utilizing VantageScore credit scoring models show that there is nearly an identical alignment of default rates between consumers with limited credit histories and consumers who have conventional credit behaviors falling into similar credit score ranges.

Over the course of the past 13 years, VantageScore credit scoring models have been very well received in the consumer lending marketplace and are viewed as highly predictive and inclusive. The numbers are impressive. More than 10.5 billion credit scores calculated with VantageScore's models were used in the twelve-month period from July 2017 through June 2018; including 4.4 billion credit scores used by credit card issuers; and there were more than 2,800 unique users of which 2,500 were lenders.

Why this growth in marketplace acceptance? Because the VantageScore 4.0 model allows lenders to accurately assess approximately 40 million more consumers than conventional models **without** lenders having to loosen their credit quality standards to reach a larger audience. That also makes the credit markets more accessible to creditworthy consumers previously locked-out of the credit



markets (often minorities and members of the armed forces returning from deployment) while creating a corresponding opportunity for lenders. It's important to note that among these 40 million additional scoreables that are produced, there are approximately 10 million consumers who score at a 620 or above. Of that 10 million, approximately 2.4 million are African Americans or Latinos who are otherwise unable to obtain a conventional score. Given the heavy testing that the 2,500 lenders and their regulators perform on scoring models used in the lending processes, those lenders would not have moved to using credit scores calculated with VantageScore if those credit scores were less predictive or inclusive than their incumbent models.

One final point for your consideration: During the Task Force hearing some suggested that consumers be given the opportunity to "opt-in" or "opt-out" of having certain data (including rent, utility and cell phone payments) considered when calculating their scores. In addressing that issue, it is important not to overlook the fact that accurate scoring relies on the principle of rank ordering the individual whose score is being computed against the total pool of scoreables. To the extent that individuals on a large scale should opt-out of having certain data considered, then a scoring model could become skewed if only part of the population were being scored. In addition, if a significant number of consumers were to "opt-in" with one CRC but not the other CRCs, then that CRC's data would perform differently than the others. Further, any biases introduced in the data by self-selection of consumers opting-in to having their data considered would need to be carefully assessed.

Chairman Lynch and Ranking Member Hill, we commend you for the way you are approaching the critically important questions regarding the use of alternative data in credit scoring and underwriting to expand access to credit. America has sensible and time proven consumer protection laws and regulations in place. Data that risks being discriminatory, that cannot be corrected, or that could be manipulated for certain outcomes clearly needs to be considered with extreme caution because at the end of the day that type of data may indeed harm, rather than benefit, the very consumers, lenders and marketplaces they are intended to benefit. Yet we nevertheless need to be innovative in discovering and testing different data streams that will in turn permit even greater access to prudent, sustainable credit.

We assure you of VantageScore Solutions' willingness to work cooperatively with you and your staffs as you move forward in addressing these issues and any other matters regarding consumer credit scores that come before the Task Force on Financial Technology and/or the House Committee on Financial Services. If you have any questions or if we can be of further assistance please don't hesitate to contact our Washington Counsel, Bill Donovan, at wjdonovan@wjdolanlaw.com or by calling (703) 254-6633.

Sincerely,



July 23, 2019

The Honorable Stephen F. Lynch
Chairman
Task Force on Financial Technology
Financial Services Committee
U.S. House of Representatives
Washington, DC 20515

The Honorable French Hill
Ranking Member
Task Force on Financial Technology
Financial Services Committee
U.S. House of Representatives
Washington, DC 20515

Dear Chairman Lynch and Ranking Member Hill:

I would like to thank the Financial Services Committee for convening its Task Force on Financial Technology and organizing its hearing titled “Examining the Use of Alternative Data in Underwriting and Credit Scoring to Expand Access to Credit” on July 25, 2019. I am writing to express my views regarding the topic of that hearing.

My name is Dan Quan and I am an adjunct scholar at the Cato Institute’s Center for Monetary and Financial Alternatives. I also advise high-growth, disruptive fintech companies. Previously, I led the Consumer Financial Protection Bureau’s (CFPB) fintech office, Project Catalyst.

The United States has the most developed and competitive consumer credit market in the world. For example, the credit card industry has seen robust growth since the financial crisis, with \$1.07 trillion in outstanding balances as of May 2019.¹ Fintech lending has seen even greater growth, accounting for 38% of the \$138 billion unsecured personal loan market in 2018.² However, the inconvenient truth buried in those rosy numbers is that we still have a huge financial inclusion challenge. 45 million Americans, or 19.3% of the adult population, don’t have access to credit.³ These consumers tend to be disproportionately African American, Hispanic, young, and low-income.⁴ Additionally, 53% of small businesses surveyed by the Federal Reserve could not obtain the full financing they sought.⁵

Now, however, the use of alternative data in credit scoring and underwriting holds great promise for bringing greater access to credit and capital to struggling consumers and small businesses.

¹ Board of Governors of the Federal Reserve System, “Consumer Credit G.19,” July 8, 2019, <https://www.federalreserve.gov/releases/g19/current/>

² Kate Rooney, “Fintechs Help Boost US Personal Loan Surge to a Record \$138 Billion,” *CNBC*, February 21, 2019, <https://www.cnbc.com/2019/02/21/personal-loans-surge-to-a-record-138-billion-in-us-as-fintechs-lead-new-lending-charge.html>

³ Consumer Financial Protection Bureau Office of Research, “Data Point: Credit Invisibles,” May 2015, https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf

⁴ Consumer Financial Protection Bureau, “Who Are the Credit Invisibles? How to Help People with Limited Credit Histories,” December 2016, https://files.consumerfinance.gov/f/documents/201612_cfpb_credit_invisible_policy_report.pdf

⁵ U.S. Federal Reserve, “Small Business Credit Survey: 2019 Report on Employer Firms,” April 16, 2019, <https://www.fedsmallbusiness.org/medialibrary/fedsmallbusiness/files/2019/sbcs-employer-firms-report.pdf>

Underwriting models that include alternative data can increase lending volume, lower interest rates for borrowers, and improve the accuracy of default predictions. In short, alternative data can make lending more plentiful, more affordable, and sounder—with historically underserved borrowers and communities benefitting most.⁶

Alternative data can include anything that is not currently part of consumers' traditional credit reports. It can range from account transaction history (known as cash-flow underwriting), to educational and occupational information, to social media use and other online or mobile activities. The topic of alternative data sometimes raises concerns about whether it could increase discrimination against protected classes or intrude on borrowers' privacy. While these are legitimate concerns, it is counterproductive to prevent or tightly constrain the use of alternative data in lending. Restrictions would likely hurt, not help, marginalized borrowers who are overwhelmingly low-income and minorities. In fact, a recent study on the racial and ethnic disparities in credit access finds that the lending gap attributed to credit discrimination is much smaller for fintech firms than it is for traditional lenders.⁷ Discouraging innovation out of a concern for potential discrimination is therefore likely to undermine financial inclusion.

In my experience, not many firms use alternative data in credit scoring and underwriting today. Those that do mostly use cash flow data. For example, fintech lenders such as Oportun use cash flow data to provide credit to “unscorable” or “credit invisible” consumers. Other lenders, like payment processors Square and PayPal, use transaction histories to help them lend to their merchants effectively and efficiently. According to Square, its average loan size is \$6,000,⁸ an amount that most traditional financial institutions find unprofitable to finance. Individual consumers can opt to include data like their on-time bill payment history to boost their FICO scores.⁹ In my conversations with fintech lenders, none report any convincing evidence that social media data can predict consumers' repayment behavior. No lender is using it for underwriting purposes in the U.S.

A recent paper from the Federal Reserve Bank of Philadelphia agrees that underwriting with alternative data can better predict loan outcomes, resulting in improved terms for borrowers who, under traditional credit criteria, would receive higher-priced loans.¹⁰

There is no need for new laws or rules just for the use of alternative data. The same laws and rules that ensure fair credit access, privacy protections, and transparency in underwriting decisions equally apply to lenders who use alternative data. On the other hand, regulatory agencies should encourage the responsible use of alternative data by providing greater clarity to lenders in two key areas.

⁶ Julapa Jagtiani and Catherine Lemieux, “Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks?” Federal Reserve Bank of Philadelphia Working Paper 18–13, March 2018, <https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2018/wp18-13.pdf>.

⁷ Robert Bartlett, Adair Morse, Richard Stanton, and Nancy Wallace, “Consumer Lending Discrimination in the Fintech Era,” National Bureau of Economic Research Working Paper No. 25943, June 2019, <https://www.nber.org/papers/w25943>.

⁸ Lalita Clozel, “The Story Behind Square’s Bank Charter Application,” *American Banker*, September 19, 2017, <https://www.americanbanker.com/news/the-story-behind-squares-bank-charter-application>.

⁹ Herb Weisbaum, “There’s a New Way to Instantly Boost Your Credit Score,” *NBC News*, February 5, 2019, <https://www.nbcnews.com/business/consumer/there-s-new-way-instantly-boost-your-credit-score-n967041>.

¹⁰ Julapa Jagtiani and Catharine Lemieux, “The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence of LendingClub Consumer Platform,” Federal Reserve Bank of Philadelphia Working Paper 18–15, April 2018 (Revised January 2019), <https://www.philadelphiafed.org/-/media/research-and-data/publications/working-papers/2018/wp18-15r.pdf>.

First, a transparent, secure, and frictionless data sharing ecosystem is necessary for increasing the use and reliability of alternative data. Regulators should support the Treasury Department's position that consumers have a right to permission their own financial data for third-party use.¹¹

Second, the CFPB needs to issue clear guidance on what responsibilities under the Fair Credit Reporting Act that third-party data aggregators and financial institutions have when consumers request to share their financial data through a secure mechanism such as an API. Uncertainty about these responsibilities creates confusion, friction, and can potentially result in consumer harm.

Additionally, the private marketplace, through industry and consumer group collaboration, can work together to ensure consumers fully understand their data rights, including how lenders will use their data, before they consent to sharing them.

The use of alternative data will enhance the affordability of credit and make our credit system more inclusive. While policymakers should remain vigilant, they must also stay open-minded about how to encourage further developments in market innovation that will benefit consumers, small businesses, and the economy at large.

I appreciate the opportunity to comment on this important hearing.

Sincerely,



Dan Quan
Adjunct Scholar
Center for Monetary and Financial Alternatives
Cato Institute

¹¹ U.S. Department of the Treasury, "A Financial System That Creates Economic Opportunities, Nonbank Financials, Fintech, and Innovation," Report to President Donald J. Trump, July 2018, https://home.treasury.gov/sites/default/files/2018-08/A-Financial-System-that-Creates-Economic-Opportunities---Nonbank-Financials-Fintech-and-Innovation_0.pdf.