Data & Civil Rights: Consumer Finance Primer

by Alex Rosenblat, Rob Randhava, danah boyd, Seeta Peña Gangadharan, and Corrine Yu
Produced for Data & Civil Rights Conference / October 30, 2014

Thanks in part to laws such as the Equal Credit Opportunity Act and the Community Reinvestment Act, our consumer finance system has come a long way in eradicating the kinds of openly discriminatory and "redlining" tactics that have historically been used to shut communities of color out of the economic mainstream. Yet today, those communities still face significant obstacles in obtaining fair and affordable credit, savings, and checking products & services that are essential to financial health. The issue has taken on additional complexity due to widespread changes in the industry, following the 2008 financial crisis and in the face of rapid technological growth.

Title VIII (the Fair Housing Act) of the Civil Rights Act of 1968 was the first federal legislation to prohibit institutions from considering some protected class information, like religion, in underwriting consumer mortgages.[1] This prohibition extended to include additional protected class categories, and to apply to other types of consumer credit transactions into the 1970s.[2] However, other characteristics that identify a particular demographic group, like address, in lieu of protected class information can still become unintentional proxies for protected status. New data analytics tools, predictive technologies, and an increasingly available range of data sources have enabled new financial instruments and services to be developed, but access to high-quality services remains restricted, often along racial and socio-economic class lines. How data is used and how algorithms and scores are designed have the potential to minimize or maximize discrimination and inequity. Yet, because of the complexity of many of these systems, developing mechanisms of oversight and accountability is extremely challenging. Not only is there little transparency for those being assessed, but the very nature of the new types of algorithms being designed makes it difficult for those with technical acumen to truly understand what is unfolding and why. This raises significant questions for those invested in making certain that finance and pricing are fair.

Background: Discrimination in Consumer Finance

Major civil rights concerns and themes in U.S. consumer finance:

- *Access to Sustainable Credit*: Before the housing crisis, people of color were often steered into risky subprime mortgages, even when they qualified for safer and less expensive loans. In recent years, the credit pendulum has swung the other way, with many creditworthy borrowers unable to obtain home loans at all.[3] Meanwhile, other expensive forms of credit – such as payday, car title, and subprime auto loans – are still heavily marketed to people of color. Improvements to consumer protection regulations and strong enforcement are essential. But eliminating predatory lending will remain challenging due to a dearth of affordable, safe alternatives in many areas. Those who rely on expensive credit (primarily people of color) are outside the mainstream scoring system, meaning they do not have a credit score or credit history, even if they are good credit risks. The financial system has been slow to adopt alternative scoring methods.

- *Lack of Savings*: One key driver of predatory lending is a lack of savings that people can tap into when they face unexpected expenses. Nearly two-thirds of households of color live in "asset poverty," with less than three months' worth of savings.[4] They are also less likely to have bank or credit union accounts, which is associated not only with lower savings rates but also with higher costs for check cashing and other basic financial services. While wage stagnation is obviously the biggest obstacle, financial education and policies to incentivize savings would help, as most people of color learn about money either through their families or "hard knocks."
- *Abusive Practices:* Consumers with mainstream bank accounts are not entirely safe from predatory or fraudulent practices. Most banks sell their customers on overdraft protection, which can be even more costly than payday loans. Some banks have also persuaded customers into buying expensive add-on products such as credit monitoring and payment protection services that they may not need. Banks and nonbanks alike have been connected to abusive debt collections, which include the use of "robo-filed" lawsuits and harassing or threatening conduct. These and similar types of practices are disproportionately likely to affect – and cause more harm to – consumers of color.
- *Emerging Technologies and New Providers:* In the wake of the financial crisis, many consumers are looking for new alternatives to traditional banking. Prepaid cards are gaining popularity, but they may have potential pitfalls such as confusing fee structures and high-cost credit features. The growth in mobile banking services is encouraging, but banks need to ensure that in-person branches remain available to all communities. And the growth in "big data" should not lead to higher costs or lower access for people of color. One idea worth exploring is using post offices to provide (or simply house) basic banking services in communities that are not currently being reached.

The Credit Score

Banks and other lending institutions rely heavily on an individual's credit score to assess whether to offer that person a loan, and at what rate. A consumer's credit score is used as a prediction of the chance that he or she will default on one or more loans within a fixed period, usually the upcoming two years. Individuals with a high credit score will have higher access to more money at a cheaper interest rate than individuals with a lower credit score, who are generally offered smaller loans at a more expensive interest rate.[5] Higher interest rates compensate lenders for the risk they take in loaning money to a less credit-worthy individual. The factors that go into a credit score, such as a missed payment, can create severe, cumulative penalties: low-income borrowers, who are less able to sustain a hit to their credit score, will have to repay future loans at higher rates that they are less able to afford than higher-income borrowers.[6] They can be sent into debt spirals and cut off from mainstream financial institutions, which makes them vulnerable to predatory alternatives like the payday loan industry. Reduced access to good credit is one of the driving factors of economic inequality; without access to substantial loans at low interest rates, individuals struggle to purchase a house, finance a business, or attend university.

Credit-scoring companies score consumers with reference to a series of attributes that statistically correlate with positive or negative outcomes. For example, Fair Isaac (now FICO) developed a mathematical product in 1962 with one of the biggest retailers in the U.S. at the time, Montgomery Ward, which used 12 factors to predict which credit applicants were a good or a bad credit risk.[7] The potential for all sorts of data to figure into predictive models fed into a desire for greater surveillance on credit applicants, although the analysis of that data was, and remains today, non-transparent to consumers. In the 1960s, credit bureaus generated reports on credit applicants that included a lot of stray and personal information, including any number of personal habits or attributes which could be used to assess credit-worthiness, like an individual's poorly kept yard or effeminate gestures.[8]

Data collection and analysis is not inherently fair, and credit bureaus were not transparent about what was or was not relevant data, which stoked public fears about what went into a credit score. To guard against arbitrariness and unfairness, Congress passed the Fair Credit Reporting Act (FCRA) in 1970, requiring credit bureaus to maintain files containing accurate and relevant information on people. FICO, for example, will describe how they calculate your credit score with reference to your payment history, amounts owed, length of credit history, new credit, and types of credit used.[9] However, the relative granularity of the data that they use and the calculation mechanisms are proprietary, so consumers know very little about how their various data are evaluated, except in broad, opaque terms. Although the FCRA gives consumers the opportunity to see one free credit report each year,[10] consumers can face practical obstacles when challenging inaccurate information in credit reports.[11]

In the 1970s, the Courts began applying anti-discrimination laws to credit-scoring mechanisms to prevent the use of protected class information, like religion, sex, age, marital status, and ethnic origin in the calculation of credit scores.[12] Lending institutions had to apply the same standards of risk mitigation equally, such that an African-American man and a Caucasian man with a similar credit history and score were equally eligible for the same type of credit.[13] The prevailing logic in political discussions asserted that businesses were, in any event, rational actors, and that they would not turn away profitable customers for arbitrary reasons, like race.[14]

Businesses have fought back against broad anti-discrimination regulation, citing to non-discriminatory intent. For example, in the 1970s, financiers argued that age should be excluded from the prohibition against using protected class information in credit-scoring assessments because banks often insured their lines of credit with credit life insurance; thus, they could not offer credit to people under the ages of 63-65.[15] Age is the only prohibited basis for discrimination that is permitted to be used in an empirically derived, demonstrably and statistically sound credit-scoring system as a predictive factor as long as it does not create a negative risk value for people over the age of 62.[16] During that same time period, Fair Isaac argued that removing relevant criteria, even if it seems politically or socially objectionable, would confound the accuracy of its mathematical scoring systems; they suggested that a clear policy mandate for changing the social make-up of the national credit pool would be a better, and less indirect, route to greater equality than forcing them to make less accurate models.[17]

Indeed, although they are important, laws that restrict the variables used in credit models may not be sufficient safeguards against discriminatory outcomes. The use of proxy variables may disfavor a particular protected class—even when discrimination or negative outcomes are unintended. A credit-scoring agency may defeat the spirit of a legal mandate to eliminate an "offensive" variable by relying on alternate data points, the analysis of which yields the same calculation predicated on the singular factor.

For instance, businesses owned by minority groups are less likely to be advanced credit than businesses owned by Whites.[18] While this can be the result of discriminatory actors, an alternative explanation is that since minority groups have lower rates of home ownership and home equity (as a result of discriminatory practices in housing markets), minority business owners may have fewer assets with which they can collateralize loans, so they have lower access overall to good credit.[19] Thus, an algorithm that excludes race from credit calculations will not automatically reduce economic inequality.[20]

Institutions designed automated credit scores in the 1970s to achieve full compliance with anti-discrimination laws, rather than to achieve fairness.[21] They rely on the rationality of algorithms[22] and their ability to sort and score people into well-ordered categories by methods that accurately identify who they are or how they should be ranked based on their data[23] to remove bias from the equation; the established presumption is that fairness is wrapped up in the accuracy and relevance of data, rather than in its analysis. In credit-scoring, unfavorable credit scores accurately and strongly correlate with protected statuses. However, they reflect and reinforce the history of prejudice that is encoded in the data they measure, and they can unfairly result in a disparate impact against protected classes. However, as long as the information and the data are accurate, there is no civil rights violation. Lenders are obliged to use accurate information, but they are not obliged to extend credit to consumers, according to Poon, who further observes that, "Non-discrimination simply means that if a lender lends to one person they cannot turn down another that expresses similar characteristics. In other words, the same data, processed in the same way, must lead to the exact same business decision."[24]

Aggregate Credit Scoring

While credit-scoring agencies assess creditors individually, data brokers also design and use predictive models based on aggregate (not individualized) data to identify or label blocks of customer bases. For example, a broker might determine a neighborhood's score on the basis of an average of individual credit scores of fifteen households in a given zip+4 geographical area and use that to make inferences about individual consumers in that area.[25] As with the case of individual credit scoring, some variables in aggregate scoring function as proxies for protected-class characteristics; neighborhoods, for instance, correlate strongly with race.[26]

Aggregate scores facilitate niche targeting of financial services or products. Although all consumers who apply for a given product are ultimately assessed based on their individual credit score, the selection of products targeted towards them may give consumers the appearance of a constrained set of options. "Fuzzy nudges"[27] prompt individuals to select options before them, rather than seeking out less visible options. In the grey zone between individual and aggregate scores, discrimination is not clearly prohibited or permitted in targeted marketing.[28] Once placed into an aggregate category, consumers can become prey to targeting for sub-prime financial products, like housing or auto loans.[29]

Aggregate customer groupings have broader implications for how individuals are assessed for financial risk based on the habits or actions of the people perceived to be in their networks. Consumers do not have access to aggregate scores, or to the factors that go into them, [30] and

FCRA applies primarily to the use of data for individual "eligibility" decisions in credit, employment, and tenancy, not to data collection.[31] Data brokers who gather information on consumers are not necessarily subject to the FCRA when they use that data to devise aggregate scores.[32] Meanwhile, lending institutions can and do rely on aggregate scoring systems in marketing and risk-mitigation.[33] For example, legal scholar Lori Andrews reported that a man whose credit limit had been $10,800 saw his credit reduced by Amex to $3,800 because he shopped at a store where other shoppers who purchased items there had a poor repayment history.[34] Of course, using common shopping habits as a predictor could be no more penalizing or rewarding than some other factor that is used to assess individual credit-worthiness, but the ways that aggregate scoring mechanisms influence individual credit scores is not readily transparent to the public. However, there is no way for an individual to effectively remedy a grouping they consider inaccurate or unduly influential.

Personalized Redlining

Data analytics is moving consumers away from democratic pricing schemes, where the price-tag on the shelf of a brick-and-mortar retail store outlet is generally the same for everyone, to a more dynamic model in which prices that can shifts daily or hourly based on the identity of the consumer and broader trends in the market, both online and in physical shops.[35] For example, Staples' online store used an algorithm that showed different prices for the same product to customers based on their location data.[36] Customers located far from competing stores see higher prices. The effect of algorithmic profiling has been described as 'personalized redlining.'[37]

This data-driven decision-making represents one increasingly common approach to market fairness. Given supply and demand, market fluctuations, and allocation of resources, such a data-driven approach is an incredibly valuable tool for companies seeking to compete on limited margins. This reduces the costs for some at the expense of others, while enabling companies to stay profitable. Yet, this data-driven approach raises significant but unanswered questions about the societal and individual impact of market-driven fairness.

Because dynamic pricing represents an opaque process to the consumer, market distortions may occur with great effect to low-income and low-education populations.[38] Customers with low levels of digital literacy will have less capacity to determine if and when they are being targeted for predatory or fair offers.[39] In the absence of better offers, they will remain largely unaware of price discrimination,[40] particularly because they may not have the means, resources, or knowledge to seek out alternative price tags. Offers that are tailored to consumers' specific, individual data points, such as their location at the time of purchase, or to the aggregate with which they are associated, can amplify discriminatory pricing overall while obscuring discrimination under the auspices of personalization. When a pricing algorithm determines who pays more for the same products, it is important to understand the factors that divide the people who are penalized or rewarded based on their data.

A New Data Divide

People who come from data-poor environments may face more significant challenges in accessing credit than those from data-rich environments.[41] According to the National Consumer Law Center (NCLC), approximately 64 million people in the U.S. have no credit history, or lack enough credit history to access credit from mainstream banks.[42] If they need a loan, they may have to use payday loan services, which are notoriously predatory.[43] In a data-centric system, the absence of a credit history can penalize a low-income consumer as much as a negative credit history.[44]

Individuals excluded from mainstream credit-scoring regimes may benefit from alternative credit scoring systems, which rely on other sources of data to assess their reliability as borrowers. For example, utility bill repayment histories could serve as an indication of credit-worthiness.[45] However, because of the leniency of utility companies (e.g., they do not traditionally shut off utilities in the event of short-term non-payment), low-income people may be delinquent on these bills, rather than other bills, when resources are scarce.[46] As a result, increasing the mechanisms of assessment may also unintentionally harm those that such a system is intended to empower.

Questions for Data, Civil Rights, and Consumer Finance

1. How can individuals control what information is gathered from them, what can be known about them, how accurate that information is, and how that information is used to increase or reduce their access to financial services?
2. What kinds of data are considered financial data? Who makes this determination?
3. Should the Fair Credit Reporting Act be expanded to cover data brokers and other types of marketing? What lessons can we learn from the weakness of the FCRA's remedies?
4. Should there be mechanisms for consumers to move beyond being able to correct data sources in order to monitor and confront algorithmic determinations?
5. How do we develop collective legal and regulatory solutions to protect individuals from new forms of inequity created by algorithmic determinations?
6. Does "big data" create new opportunities for consumers to increase their access to high quality credit? How would we go about assessing this?
7. What kinds of additional data would be predictive of the credit-worthiness of low-income communities? Does new data for underwriting put pressure on other key policies, such as public assistance programs that use unpaid bills as a criterion of eligibility?
8. How can data analytics be used to identify and remedy biases in the programs that are adopted by financial institutions?
9. How do data analytics intersect with differential pricing?
    a. What role does alternative scoring play in finance?[47]
    b. What constitutes market fairness?
    c. How do personalized pricing mechanisms intersect with civil rights concerns?
    d. How would the average consumer become aware of predatory pricing or price discrimination?

[1] Martha Poon, "Discriminating Without Discrimination," in "What Lenders See — a History of the Fair Isaac Scorecard," (PhD dissertation, Department of Sociology, Science Studies Program, UCSD, 2012), 1.
[2] Poon, "Discriminating Without Discrimination."
[3] *See, e.g.*: Urban Institute, "Where Have All the Loans Gone? The Impact of Credit Availability on Mortgage Volume," by Laurie Goodman, Jun Zhu, and Taz George (Washington, D.C., 2014). http://www.urban.org/publications/413052.html.
[4] Corporation for Enterprise Development, "Treading Water in the Deep End: Findings from the 2014 Assets & Opportunity Scorecard," (Washington, D.C., 2014). http://assetsandopportunity.org/scorecard/about/main_findings/.
[5] Frank Pasquale, *The Black Box Society* (Cambridge, MA: Harvard University Press, 2015), 32.
[6] Pasquale, *The Black Box*, 256-257.
[7] Poon, "Discriminating Without Discrimination," 19.
[8] Pasquale, *The Black Box*, 31.
[9] "FICO Credit Score Chart: How Credit Scores Are Calculated," myFICO, accessed November 11, 2014, http://www.myfico.com/crediteducation/whatsinyourscore.aspx.
[10] "Free Credit Reports," Federal Trade Commission, accessed November 11, 2014, http://www.consumer.ftc.gov/articles/0155-free-credit-reports.
[11] Alan Farnham, "Woman Wins $18.6M for Credit Report Errors," *ABC News*, July 30, 2013. http://abcnews.go.com/Business/equifax-loses-186-million-lawsuit/story?id=19803421.
[12] Poon, "Discriminating Without Discrimination," 28.
[13] Poon, "Discriminating Without Discrimination," 24.
[14] Poon, "Discriminating Without Discrimination," 5-6.
[15] Poon, "Discriminating Without Discrimination," 5-6.
[16] "Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit," The Federal Reserve Board, last updated April 30, 2008, http://www.federalreserve.gov/boarddocs/rptcongress/creditscore/demographics.htm.
[17] Martha Poon, "Discriminating Without Discrimination," in "What Lenders See — a History of the Fair Isaac Scorecard," (PhD dissertation, Department of Sociology, Science Studies Program, UCSD, 2012), 207, cited in Solon Barocas, "Panic Inducing: Data Mining, Fairness, and Privacy" (PhD dissertation, New York University, 2014), 244.
[18] Yue Hu, Long Liu, Jan Ondrich, and John Yinger, "Do Minority and Woman Entrepreneurs Face Discrimination in Credit Markets? Improved Estimates Using Matching Methods," (paper, Syracuse University, 2010). http://faculty.maxwell.syr.edu/jondrich/HLOY.pdf
[19] Emily Maltby, "Minority Entrepreneurs Still Face Bias, Prof Says," *WSJ Blogs - In Charge*, March 10, 2011. http://blogs.wsj.com/in-charge/2011/03/10/minority-entrepreneurs-still-face-bias-prof/.
[20] Barocas, "Panic Inducing," 232.
[21] Poon, "Discriminating Without Discrimination," 3.
[22] David Beer, "Power through the algorithm? Participatory web cultures and the technological unconscious," *New Media & Society* 11, 6 (2009): 985-1002; David Lyon, "Surveillance as Social Sorting: Computer Codes and Mobile Bodies," in *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*, ed. David Lyon. (New York, NY: Routledge, 2003); both cited in Solon Barocas, Sophie Hood, and Malte Ziewitz, "Provocation Piece - Governing Algorithms," (conference material, New York University, 2013). http://governingalgorithms.org/resources/provocation-piece/.
[23] Clive Norris, "From Personal to Digital," in *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*, ed. David Lyon. (New York, NY: Routledge, 2003), Kindle edition, 276.
[24] Poon, Martha. "Notes on civil rights in data-driven consumer finance." Email message to the authors. November 05, 2014, 2.
[25] Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," (Washington, D.C., 2014), 106. http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.
[26] Sewin Chan, Michael Gedal, Vicki Been, and Andrew Haughwout, "The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City," (paper, New York University, 2011). http://furmancenter.org/files/publications/Pathways_1_Newest_with_Figures_Working_Paper.pdf.
[27] Ashkan Soltani, "Emerging Trends in Online Pricing," (presentation transcript, Federal Trade Commission Spring Privacy Series, 2014), 56. http://www.ftc.gov/system/files/documents/public_events/182261/alternative-scoring-products_final-transcript.pdf.

[28] Federal Trade Commission, "Data Brokers," 106.
[29] Eileen Sullivan and Associated Press, "Discrimination Potentinal Seen in 'Big Data' Use," *23ABC News*, April 27, 2014. http://www.turnto23.com/news/state/discrimination-potentinal-seen-in-big-data-use-042714.
[30] Pamela Dixon, "Alternative Scoring Products," (panel discussion transcript, Federal Trade Commission Spring Privacy Series, 2014), 43. http://www.ftc.gov/system/files/documents/public_events/182261/alternative-scoring-products_final-transcript.pdf.
[31] Chris Jay Hoofnagle, "The Potemkinism of Privacy Pragmatism," *Slate*, September 2, 2014, 2. http://www.slate.com/articles/technology/future_tense/2014/09/data_use_regulation_the_libertarian_push_behind_a_new_take_on_privacy.2.html.
[32] Pamela Dixon, "Alternative Scoring Products," 54.
[33] Federal Trade Commission, "Data Brokers," 107.
[34] Lori Andrews, "Facebook Is Using You," *The New York Times*, February 4, 2012. http://www.nytimes.com/2012/02/05/opinion/sunday/facebook-is-using-you.html.
[35] Soltani, "Emerging Trends," 84-85.
[36] Jennifer Valentino-DeVries, Jeremy Singer-Vine, and Ashkan Soltani, "Websites Vary Prices, Deals Based on Users' Information," *Wall Street Journal*, December 24, 2012. http://online.wsj.com/news/articles/SB10001424127887323777204578189391813881534.
[37] Bill Davidow, "Redlining for the 21st Century," *The Atlantic*, March 5, 2014. http://www.theatlantic.com/business/archive/2014/03/redlining-for-the-21st-century/284235/.
[38] Nathan Newman, "How Big Data Enables Economic Harm to Consumers, Especially to Low-Income and Other Vulnerable Sectors of the Population," (public comments, Federal Trade Commission, 2014), 4-5. http://www.ftc.gov/system/files/documents/public_comments/2014/08/00015-92370.pdf.
[39] Newman, "Big Data Enables Economic Harm," 1.
[40] Aaron Pressman, "Big Data Could Create an Era of Big Discrimination," *Yahoo Finance*, October 14, 2013. http://finance.yahoo.com/blogs/the-exchange/big-data-could-create-era-big-discrimination-191444085.html.
[41] Center for Data Innovation, "The Rise of Data Poverty in America," by Daniel Castro (Washington, D.C., 2014). http://www2.datainnovation.org/2014-data-poverty.pdf.
[42] National Consumer Law Center, "Big Data, a Big Disappointment for Scoring Consumer Creditworthiness," by Persis Yu, Jillian McLaughlin, and Marina Levy (Boston, MA, 2014). http://www.nclc.org/issues/big-data.html.
[43] National Consumer Law Center, "Big Data, a Big Disappointment."
[44] Robert O'Harrow, Jr. and Michael Sallah, "Police Intelligence Targets Cash," *Washington Post*, September 7, 2014. http://www.washingtonpost.com/sf/investigative/2014/09/07/police-intelligence-targets-cash/.
[45] PERC, *"A New Pathway to Financial Inclusion: Alternative Data, Credit Building, and Responsible Lending in the Wake of the Great Recession,"* by Michael Turner, Patrick D. Walker, Sukanya Chaudhur, and Robin Varghese (Durham, NC, 2012), 20. http://www.perc.net/wp-content/uploads/2013/09/WEB-file-ADI5-layout1.pdf.
[46] National Consumer Law Center, "Full Utility Credit Reporting: Risks to Low-Income Consumers," (Boston, MA, 2012). http://www.nclc.org/images/pdf/energy_utility_telecom/consumer_protection_and_regulatory_issues/ib_risks_of_full_utility_credit_reporting_july2012.pdf; cited in David Robinson and Harlan Yu, "Civil Rights, Big Data, and Our Algorithmic Future," accessed November 11, 2014. http://bigdata.fairness.io/alternative-data/.
[47] Claudia Perlich, "Alternative Scoring Products," (panel discussion transcript, Federal Trade Commission Spring Privacy Series, 2014), 86. http://www.ftc.gov/system/files/documents/public_events/182261/alternative-scoring-products_final-transcript.pdf.