No Silver Bullet: Using Alternative Data for Financial Inclusion and Racial Justice

June 2022 (updated)

Background

Credit scores reflect dramatic and troubling disparities by race. For instance, a 2019 study by the Urban Institute found that over 50 percent of white households have credit scores over 700, but only 20 percent of Black households do. The Urban Institute study is the most recent report showing these startling racial credit scoring gaps; there are a multitude of older studies with similar results, listed in this report. The racial disparities in credit scores are due to a multitude of factors, including the racial wealth gap, decades of redlining and housing segregation, historical and present-day employment discrimination, and racially biased criminal justice practices.

Black and Latinx consumers are also more likely to lack a credit history or have too scant a history to generate a credit score, referred to as being "credit invisible." The Consumer Financial Protection Bureau (CFPB) found that about 15 percent of Black and Latinx consumers are credit invisible (compared to 9 percent of whites and Asians) and an additional 13 percent of Black and 12 percent of Latinx consumers have unscorable records (compared to 7 percent of whites).

Spurred by the Black Lives Matter protests of 2020 and subsequent racial reckoning, policymakers and advocates have focused on developing solutions to these racial disparities. One of the most touted solutions is using alternative data, *i.e.*, any data that is not traditionally included in credit reports issued by the Big Three credit bureaus (Equifax, Experian, and TransUnion). This includes relatively conventional financial data such as rent, telecom and utility payments as well as bank account cashflow data, which are the topic of this brief. It can also include less conventional data such as web browsing, social media, educational background or "Big Data."

But alternative data will not eliminate racial disparities in credit scores and is not a panacea for credit inequities. Alternative data may help some credit invisibles, including Black and Latinx consumers, but only if used carefully. Some forms are more promising than others. Most critically, any data that relies on financial information will still reflect racial disparities given the unequal economic positions of households of color and white households. And when financially-based data is fed into algorithms or artificial intelligence models, the results could replicate or amplify those disparities.

As for narrowing the racial homeownership gap, the impact may be limited. A November 2021 Government Accountability Office (GAO) study noted that while alternative data could "improve or generate scores for [credit invisible or low scoring] consumers, it is unclear whether the increases would be sufficient to qualify many additional consumers for lower-cost mortgages." The GAO also noted that nearly half (48 percent) of unscorable consumers were under 24 or over 65 years old, which the GAO characterized as "age groups less likely than most to be seeking mortgage credit."

The Details Matter

As discussed in an earlier NCLC Policy Brief, the “devil is in the details” with respect to alternative data. Some types and approaches to using it are helpful or promising, but others are harmful or pose risks. The following discussion elaborates on this and discusses the implications of each type of data in reducing racial disparities in credit scoring.

Important Consumer Protections for Alternative Data

To be most beneficial and avoid harming vulnerable consumers, alternative data should be:

- used only with the explicit and knowing permission or opt-in of the consumer
- exclusively consist of positive payment data
- be accessed by means other than the Big Three credit bureaus to avoid furthering entrenching their oligopoly position.

1. Rent Payment Data

Rent payment data is often aggressively promoted as a form of alternative data, and it can be promising. One pilot study of affordable housing residents found that 79% experienced an increase in their credit score due to rent payment reporting, by an average of 23 points.

However, rent payments must be reported in a manner that aids and empowers the renters who can benefit from them, but does not harm households struggling with housing costs. Rent payment reporting should always be with the consumer’s active permission – it should be *opt-in* only. Consumers should always have control over whether their data is shared as a basic principle. In addition, rent reporting should be limited to only positive payment information; programs that report negative or “full file” information have the potential to hurt the most vulnerable.

A consumer lawyer providing *pro bono* services at a homeless shelter sent NCLC a troubling example of the potentially devastating consequences from full-file rent reporting

> *I was shocked to find out that most of the problems had to do with credit reporting – right down my alley. Of the 10 or so folks I’ve met with so far, at least 6 were living in the shelter solely because a rent-related credit reporting issue had shut them out of the housing market. These are folks with enough income to pay market, non-subsidized rent, but were nevertheless forced to move their families into a homeless shelter simply because their credit history disqualified them from renting. And, it’s not a credit score problem. Folks with decent credit scores were automatically disqualified because of a past rent delinquency (even a single 30-day late). … I’m keenly aware of the outsized role credit reporting can play in keeping folks in difficult financial situations in general, but I was shocked to see a direct connection to something as extreme as homelessness.*

Email from Kenneth Riemer to Chi Chi Wu. April 7, 2022.

The pandemic has illustrated the potential harm of rent reporting if done incorrectly. At various times during the past two years, between six million to thirteen million households have been behind in rent. These households are disproportionately renters of color; for example, in September 2020, about 1 in 4 Black and Asian renters and 1 in 5 Latinx renters said they were not caught up on rent, compared to just 1 in 9 white renters. These tenants would all suffer significant credit harm to their credit reports if rent reporting is not positive-only and voluntary.

2. Utility payments

Consumer advocates have consistently opposed efforts to promote "full file" monthly reporting of gas and electric bill payment data to the Big Three credit bureaus, unless such efforts are voluntary for consumers. Such efforts have the potential to harm low-income consumers by adding reports about payments that are only 30 or 60 days late. 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 immediately pay seasonally high bills at the expense of other necessities.

Like rent payments, racial disparities exist in utility payment data. As an NCLC report details:

> African American and Latinx households are far more likely to experience a loss of heating or cooling due to an inability to pay bills and COVID-19 has only widened this gap. Survey data from Indiana University shows that low-income Black households are sent utility disconnection notices at two times the rate and have their electricity disconnected at five times the rate compared to low-income white households. Hispanic households receive notices at five times the rate and lost electricity at eight times the rate.

Report: More Must Be Done to Prevent Utility Consumers from Losing Service Due to COVID-Driven Arrearages, Nov. 2021 (p.10, footnotes omitted).

3. Bank Account Transaction & Cashflow Data

Bank account transaction and cashflow information holds great promise as a form of alternative data, but also great risk. Research by FinRegLab has shown that this data has the potential to help borrowers of color who might otherwise face constraints on their ability to access credit. Because it includes both income and expense information, bank account data can reveal whether a consumer has sufficient funds to afford taking on a loan.

But while it could open the doors to for some credit invisible consumers, the sharing of this data should only occur with the consumer's knowing and active permission. Bank account transaction data can be very sensitive and revealing, especially since a credit invisible consumer would not have a credit card and thus would likely be using a debit card for many transactions. The data might show when the consumer gets paid, where they shop, what advocacy organizations they support, or which healthcare providers they use. This sort of information is not only private, but information such as what neighborhood a consumer shops in could reflect racial and socio-economic disparities.

Consumers must never be required to provide this data, especially if they already have a thick credit file that qualifies them for credit. The nightmare scenario is a system where every consumer – thin or thick file, high FICO score or not – is forced to give up their privacy and allow creditors, employers, landlords, insurers, and government agencies a direct and permanent digital pipeline to their bank account data.

Furthermore, bank account data will almost certainly exhibit disparities by race. A key factor likely to be used by scoring models is overdrafts, and Black consumers are disproportionately affected by bank overdraft practices. A 2016 Pew study found that Black consumers are 12 percent of the US population, but account for 19 percent of the heavy overdrafters. Indeed, the ability of cash flow data to help minority and low-and moderate-income consumers will not bear fruit unless and until bank overdraft abuses are brought to an end.

Black consumers are also disproportionately unbanked – 13.8 percent of Black households were unbanked in 2019 (as well as 12.2 percent of Latinx households) compared to 2.5 percent of white households. Being unbanked, of course, excludes consumers from the possibility of using cashflow data. A significant portion of unbanked consumers leave or are forced out of traditional bank accounts because of high overdraft fees.

Stop Feeding the Credit Bureau Oligopoly

The Big Three credit bureaus have promoted the idea of alternative data as a panacea, claiming that it is a form of "credit inclusion." In a May 2021 Congressional hearing, the President of TransUnion Consumer Interactive touted the formation of that company's Racial Equity Task Force and emphasized how "alternative data is critical for credit inclusion."

However, one of the biggest criticisms of credit bureaus is that they constitute an oligopoly, indeed a functional monopoly, because consumers can't choose between the Big Three or walk away from them altogether. Including alternative data in credit reports has the effect of feeding even more of our financial information into the database of these three corporations, increasing their oligopoly power.

Consumers need alternatives to the credit bureaus. In this respect, the use of bank account cashflow information is the most promising form of alternative data, because it does not necessarily involve the credit bureaus but instead relies on third party "data aggregators." Thus, it could present a true alternative to the oligopoly of the credit bureaus.

Going Deeper and Better to Reduce Racial Disparities

In addition to using alternative data in a positive manner with consumer control, there will need to be other major efforts in order to minimize racial disparities and make scoring equitable. For one thing, credit scoring models need to be refined and improved with *intentionality*. Intentionality is key - the income disparities and wealth gaps reflected by credit scores were the product of centuries of intentional discrimination. They cannot and will not be reduced or resolved without the same level of intentionality.

Researchers have explored methods and tools to reduce these racial disparities and such efforts should be encouraged. An example is the process known as adversarial debiasing or adversarial learning, which uses two adversarial models to maximize the ability of the first model to predict a desired factor (e.g., probability of defaulting on a loan) while the second model that predicts the sensitive attribute (e.g., protected class) minimizes its disparities in the first.

At the same time, other phenomena may have disproportionate impacts on the credit records of Black and Latinx communities. A CFPB report documented double the rates of complaints to CFPB in majority Black communities, as well as a greater portion of complaints involving credit reporting in both Black and Latinx communities – 64% of those filing complaints in majority Black communities and 53% in majority Latinx communities had a credit reporting complaint versus 37% of those in predominantly white communities. The CFPB has suggested that higher rates of victimization for some kinds of identity theft may be part of the explanation. The Federal Trade Commission similarly noted that, in majority Black communities, the top problem reported to the FTC Consumer Sentinel database was about credit bureaus – about 21% of the total reports versus 5.2% of the reports in majority white communities and 12% in Latinx communities. Recent academic work explains why low-income and minority consumers may be more vulnerable to identity theft than conventional wisdom assumes, and some solutions for that problem.

Finally, there must be efforts to reduce the racial disparities in the underlying factors that create the divide in credit scores and other financial data – *i.e.*, reducing the racial wealth gap, combatting housing segregation, preventing employment discrimination, and implementing criminal justice reform. The experience during the COVID-19 pandemic is instructive. With two stimulus payments and expanded federal unemployment benefits, credit scores stayed stable and even went up for some consumers. The lesson is simple: if consumers have more financial resources, they can better pay their bills and their scores go up. Racism and its legacy drain precious resources from communities of color, which is directly reflected in credit scores and other financial data.