

Updated November 7, 2024

Alternative Data in Financial Services

Alternative data generally refers to information used to determine a consumer’s creditworthiness that the national consumer reporting agencies (CRAs)—Equifax, Experian, and TransUnion—have not traditionally used to calculate credit scores. These CRAs generally create consumer reports with historical information about repayment on credit products such as mortgages, student loans, credit cards, and auto loans. Credit applications, bankruptcies, and debts in collection are also regularly included.

New technology makes it possible for financial institutions to gather this alternative data, which include financial and nonfinancial data, from a variety of sources. Examples of alternative data can include payment history in telecommunications, rent, or utilities; checking account transaction information; educational or occupational attainment; how consumers shop, browse, or use devices; and social media information. Some of this information is now commonly used in certain types of credit underwriting and scoring. This alternative data can potentially enable CRAs and financial firms to score and underwrite credit to borrowers who would otherwise be denied credit, but it may present novel risks as well.

In the 118th Congress, legislation has been introduced to encourage increased usage of alternative data. Among the bills introduced are bills that would require federal agencies to incorporate alternative data in mortgage underwriting (H.R. 123 and H.R. 1266), enable/require reporting of rental or utility payments (S. 1654/H.R. 3418, S. 4944), and expand federal regulators’ use of sandboxes, which allow firms to experiment without regulatory action (H.R. 9309/S. 4951, H.R. 6584).

Different Uses of Alternative Data

Two prominent ways alternative data are used are in credit reporting and lending by financial technology (“fintech”) companies.

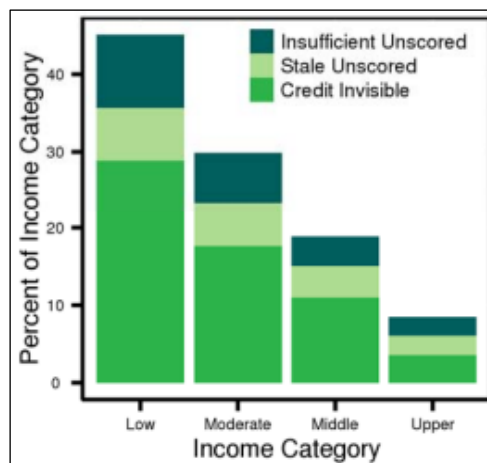
Credit Reporting

According to the Consumer Financial Protection Bureau (CFPB) in 2015, credit scores cannot be generated for approximately 20% of the U.S. population due to their limited credit histories. This statistic, the most recent provided by the CFPB, pre-dates many innovations in credit scoring that may have improved coverage.

Those without credit scores skew lower-income and younger and are more likely to be minorities. **Figure 1** shows the share of Americans by census tract income level that are considered (1) *credit invisible* with no information available, (2) *stale unscored* with no recent information, and (3) *insufficient unscored* with a lack of information for traditional scoring. The share of consumers in any of these

groups is highest in the low-income census tract (46%) and lowest in the upper-income category (7%).

Figure 1. Share of Consumers Unscored by Census Tract Income Level



Source: CFPB.

Incorporating alternative data into credit scoring can help score a broader swathe of the population. Research from Experian found that incorporating alternative data could help score 8.4 million previously unscorable borrowers. Further, alternative data could improve existing scores. Research from TransUnion found that when rent payments were included in consumers’ credit files, consumers “experienced an average increase of nearly 60 points to their credit score.” These increases were disproportionately for consumers with low credit scores.

The national CRAs have increasingly used alternative data in at least some of the models that they develop, but incorporating this data often requires a borrower to opt in. Specifically, Experian has developed a platform called “Experian Boost” that can consider the past two years of positive payment history for consumers’ rent, internet, utility, or phone payments that are made through their banks, credit cards, or other service providers. Data used by CRAs is regulated by the Fair Credit Reporting Act (15 U.S.C. §1681), meaning such data must still be “displayable, disputable and correctable.”

Fannie Mae and Freddie Mac have partnered with fintechs such as Esusu to encourage reporting of on-time rents in property management software. According to Freddie Mac, 10% of renters see their on-time rental payments reflected in their scores, so significant challenges remain in capturing rents or other measures of alternative data. Traditionally, such data was reported to credit bureaus only when delinquent, so such programs can encourage increased reporting of positive payment history.

Fintech Lending

Fintech lenders often use alternative data directly to make credit decisions in addition to or instead of information in the credit reporting system. Fintechs might have access to novel data through data brokers, prior relationships with borrowers, or borrowers sharing relevant information. Recent policy developments at the CFPB with the Section 1033 rule might streamline and accelerate the use of alternative data in fintech underwriting. Fintechs often pair this alternative data with artificial intelligence and machine learning technologies that can help identify relationships between that data and credit default risk.

Using this alternative data could allow them to increasingly lend to borrowers who might have otherwise been denied access to credit. For example, cash-flow data can give a better picture of consumers' finances. These lending models have benefited fintechs, as they play prominent roles in small business, mortgage, and consumer lending markets. There is existing debate in academic research if fintech firms that incorporate this alternative data are better at predicting credit default risk in the small business and consumer lending markets.

Financial Regulatory Policy Issues

Policymakers face a number of potential financial regulatory policy issues specific to alternative data.

Data Security and Privacy

The collection, storage, and distribution of alternative data might result in concerns over consumers' control over such data and their knowledge of how that data is being used in credit underwriting. The Gramm-Leach-Bliley Act (GLBA), requires financial institutions to safeguard this sensitive data and provide initial and annual disclosures, including the conditions in which they may share personal information. Currently, credit bureaus and fintechs generally require opt ins or applications to use alternative data in credit scoring or underwriting, meaning arguably consumers control their alternative data that is used in these spheres. This alternative data is under the storage and sharing requirements of the GLBA.

Fair Lending

The Equal Credit Opportunity Act prohibits discrimination against credit applicants. As a result, institutions utilizing alternative data in underwriting are prohibited from discriminating on the basis of protected class. Federal regulators, including the CFPB and the Department of the Treasury, have previously noted potential fair lending concerns associated with alternative data. For example, some consumer advocates have argued utilizing alternative nonfinancial data such as educational history that correlates with race can result in fair lending concerns. According to a 2021 Government Accountability Office study, 13 of 16 mortgage lenders interviewed said that fair lending was a concern when using alternative data. These regulatory risks could discourage financial institutions from utilizing alternative data.

Regulatory Sandboxes

Federal regulators can possibly encourage further alternative data usage through regulatory sandboxes, which

enable companies to test novel data and methods without fear of immediate adverse regulatory action. Sandboxes are often implemented in the form of "No-Action Letters" or NAL. The CFPB finalized its initial NAL policy in 2016, with further modifications in 2018 and 2019. In September 2022, the CFPB rescinded its previous NAL and sandbox policies.

The CFPB previously issued NALs for Upstart in 2017 and 2020 to utilize alternative data including education and work history to underwrite consumer loans. CFPB analysis found that Upstart's model including alternative data and machine learning had higher acceptance rates and lower costs relative to traditional models, while critics claimed that there were potential fair lending concerns with its modeling. This NAL was rescinded in June 2022, as Upstart wanted to add additional variables to its model. Separately, the CFPB approved applications from the Bank Policy Institute and Bank of America to experiment with utilizing cash flow in underwriting small dollar loans. Previous research by the CFPB has found cash flow to be useful in assessing credit risk. Despite no formal legal risk, in the case of the CFPB's sandbox, the CFPB generally reviewed and monitored these lending models.

Selected Legislation: 118th Congress

Legislation introduced in the 118th Congress would have implications for financial institutions' use of alternative data. S. 1654/H.R. 3418 would allow the Department of Housing and Urban Development and utility companies to report rental and utility payments to CRAs. S. 4944 would compel companies with federally backed multifamily real estate or residential loans to submit positive rental payments if tenants opt in to the program.

Other legislation targets the development of pilot programs to use alternative credit scoring data in underwriting. H.R. 1266 would authorize the Department of Veterans Affairs (VA) to use alternative data or credit scoring models for individuals who are otherwise eligible for VA loans but have insufficient credit history. These individuals would need to opt in to alternative data scoring. H.R. 123 would establish a pilot program for alternative data at the Federal Housing Administration, enabling potential borrowers with otherwise insufficient creditworthiness to opt in to that model.

In addition, some legislation is aimed at expanding federal regulators' sandbox policies. H.R. 9309/S. 4951 would create sandboxes at financial regulatory agencies to execute artificial intelligence projects, likely relying on alternative data. H.R. 6584 would restore the CFPB's sandbox and NAL policies to those present in 2019 with robust protections from CFPB legal action before their recent rescindment by the CFPB.

CRS Resources

CRS Report R47997, *Artificial Intelligence and Machine Learning in Financial Services*, by Paul Tierno

Karl E. Schneider, Analyst in Financial Economics

Disclaimer

This document was prepared by the Congressional Research Service (CRS). CRS serves as nonpartisan shared staff to congressional committees and Members of Congress. It operates solely at the behest of and under the direction of Congress. Information in a CRS Report should not be relied upon for purposes other than public understanding of information that has been provided by CRS to Members of Congress in connection with CRS's institutional role. CRS Reports, as a work of the United States Government, are not subject to copyright protection in the United States. Any CRS Report may be reproduced and distributed in its entirety without permission from CRS. However, as a CRS Report may include copyrighted images or material from a third party, you may need to obtain the permission of the copyright holder if you wish to copy or otherwise use copyrighted material.