BIG DATA

A BIG DISAPPOINTMENT FOR SCORING CONSUMER CREDIT RISK

March 2014

NCLC®
NATIONAL CONSUMER LAW CENTER®

Revised March 14, 2014.

ABOUT THE AUTHORS

Co-author **Persis Yu** is a staff attorney at the National Consumer Law Center (NCLC) and works in the Student Loan Borrower Assistance Project and on other consumer advocacy issues. Prior to joining NCLC, she was a Hanna S. Cohn Equal Justice Fellow at Empire Justice Center in New York. Her fellowship project focused on credit reporting issues facing low-income consumers, specifically in the areas of accuracy, housing, and employment. Persis is a graduate of Seattle University School of Law, and holds a Masters of Social Work from the University of Washington and a Bachelor of Arts from Mount Holyoke College. She is a contributor to NCLC's *Student Loan Law* and *Fair Credit Reporting*.

Co-author **Jillian McLaughlin** is a Master of Public Policy candidate at the Harvard Kennedy School of Government with a concentration in Business and Government Policy. She holds a Bachelor of Arts from Kalamazoo College and is a former researcher at NCLC.

Contributing author **Marina Levy** conducts research and assists the advocacy team at NCLC. She completed her undergraduate degree in International Affairs and Applied Legal Studies at Suffolk University, where she worked as a research assistant for the Government Department.

ACKNOWLEDGMENTS

This report is a release of the National Consumer Law Center. The authors thank David Robinson, Harlan Yu, and Aaron Rieke of Robinson & Yu, LLC; Ed Mierzwinski of the U.S. Public Interest Research Group; and many others for providing their valuable time and expertise on this subject. We also thank NCLC colleagues Carolyn Carter, Jan Kruse, Lauren Saunders, Margot Saunders, Olivia Wein, and Chi Chi Wu for valuable comments and assistance, and the many colleagues who requested their data as a part of our accuracy study.

The findings and conclusions presented in this report are those of the authors alone. This report was completed on February 14, 2014; information on the chart was fact checked as of Dec. 11, 2013.

NCLC®
NATIONAL CONSUMER LAW CENTER®

ABOUT THE NATIONAL CONSUMER LAW CENTER

Since 1969, the nonprofit National Consumer Law Center® (NCLC®) has used its expertise in consumer law and energy policy to work for consumer justice and economic security for low-income and other disadvantaged people, including older adults, in the United States. NCLC's expertise includes policy analysis and advocacy; consumer law and energy publications; litigation; expert witness services, and training and advice for advocates. NCLC works with nonprofit and legal services organizations, private attorneys, policymakers, and federal and state government and courts across the nation to stop exploitive practices, help financially stressed families build and retain wealth, and advance economic fairness.

7 WINTHROP SQUARE, BOSTON, MA 02110 ■ 617-542-8010 ■ WWW.NCLC.ORG

BIG DATA

A BIG DISAPPOINTMENT FOR SCORING CONSUMER CREDIT RISK

TABLE OF CONTENTS

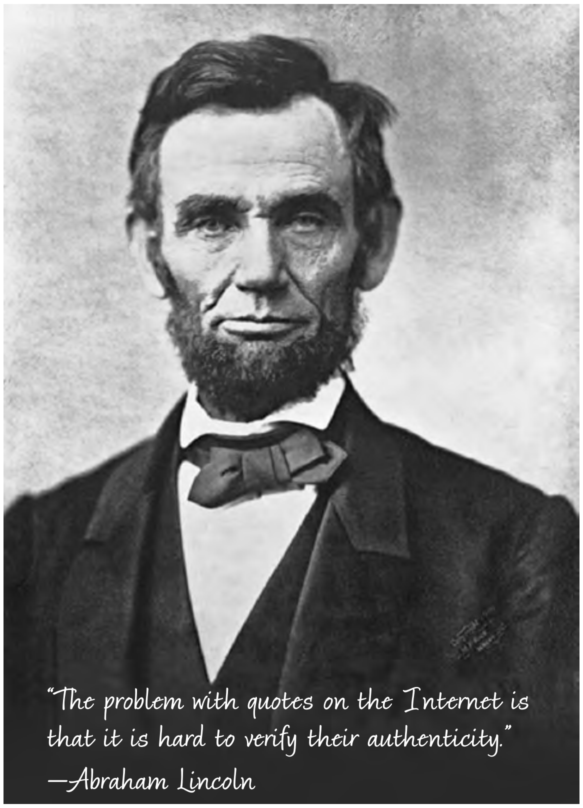


Photo used under a Creative Commons license. Originally posted on Flickr by Emory Maiden, www.flickr.com/photos/44551921@N04/7301369296

EXECUTIVE SUMMARY

Approximately 64 million consumers in the United States have no credit history or lack sufficient credit history to generate a credit score, cutting off access to traditional banking services. Finding a way of getting affordable access to credit is of vital importance to the economic well-being of this population. It also represents an untapped market with the potential for big profits. So it is unsurprising that in this era of big data, information culled from Internet searches, social media, and mobile apps would be put to use towards that goal. However, it is unclear as to whether doing so will be beneficial for the low-income consumer. These products may fill a void and provide affordable access to credit to these underserved populations *or* they may be a means of preying on vulnerable communities.

Big data makes big promises. It promises to make better predictive algorithms that in turn can make better products available to the unbanked and underbanked. But can big data live up to this big promise?

When analyzing this use of big data, consumers and policy makers should be concerned with these questions:

1. Are the decisions based upon accurate data?
2. Can the algorithms, when fed with good data, actually predict the creditworthiness of low-income consumers?
3. Does the use of big data in reports used for credit, employment, insurance, and other purposes comply with consumer protection laws?
4. Is there the potential for a discriminatory impact on racial, geographic, or other minority groups?
5. Does the use of big data actually improve the choices for consumers?

Answering these questions has been especially challenging given the secretive and proprietary nature of the products examined. Therefore, the National Consumer Law Center (NCLC) did its own investigation of the information data brokers had on its staff and reviewed products using big data analytics.

NCLC's Study of Big Data Accuracy

Big data proponents argue that multiplying the number of variables will expand access to borrowers with thin credit files. Expanding the number of data points also introduces the risk that inaccuracies will play a greater role in determining creditworthiness. Given these indications of accuracy problems, we conducted our own survey for this report of the data maintained on consumers by big data brokers. Even given our initial skepticism, we were astonished by the scope of inaccuracies among the data brokers we investigated.

In general, obtaining the data was challenging and the reports our volunteers received were riddled with inaccuracies or included little or incomplete information. Errors ranged from the mundane—a wrong e-mail address or incorrect phone number—to seriously flawed. Interestingly, eBureau touts its ability to estimate income based on its advanced models and offer insights based upon the consumer's education. Despite that claim, seven of the fifteen consumer reports generated by eBureau contained errors in estimated income, nearly doubling the salary of one participant and halving the salary of another, and eleven of the fifteen reports incorrectly stated the volunteer's education level. Reports purchased from Intelius and Spokeo had the most inaccuracies while the reports from Acxiom, eBureau, and ID Analytics contained very little information.

Applying the Fair Credit Reporting Act

An analysis of the Fair Credit Reporting Act, shows that many big data brokers could be considered consumer reporting agencies (CRAs) and subject to the FCRA. The FCRA imposes substantial duties on a CRA. Three of the most important functions of the FCRA deal with accuracy, disclosure, and the right to dispute items on the report. It is highly unlikely, given the size of the data set and the sources of information, that the companies that provide big data analytics and the users of that data are meeting these FCRA obligations.

Evaluating the Discriminatory Impact

Because big data scores use undisclosed algorithms, it is impossible to analyze the algorithm for potential racial discriminatory impact. According to the companies' marketing materials, consumers are judged based upon data generated from their Internet usage, mobile applications, and social media. However, access and usage of these sources vary by race and socioeconomic status, and thus any algorithm based upon them may have racial disparities.

Different races also use the Internet differently. For example, according to Nielsen spokesman Matthew Hurst, "Black consumers are also 30 percent more likely to visit Twitter using mobile phones than the average customer." These different ways of accessing the Internet leave a digital data trail. Yet, despite these known differences, little is known about how each of these variables is weighted or used by big data analytics.

Big Data, Better Products?

Finally, proponents of big data underwriting argue that by using a constellation of factors to price credit, the cost of credit will be reduced for low-income borrowers, thus enabling lenders to provide lower-cost small loans as alternatives to payday loans. We evaluated seven loan products that are based on big data underwriting, six of which present themselves as payday loan alternatives. Some of the features of these loans are arguably "less bad" than those offered by traditional payday lenders, but these products still fail to meet the requirements to be considered genuine, better alternatives. They still feature three-digit APRs.

Even more troubling is that all of the lenders except Presta and MySalaryLine require borrowers to provide sensitive banking information (i.e. bank name, routing number, and account number). A lender could potentially use this information to reach into a bank account and take the funds if the consumer fails to make a payment. The requirement for electronic information is of concern and may be an attempt to obtain access to the consumer's account while evading the important protections of the Electronic Funds Transfer Act. The requirement that the borrower provide bank account information could ensure that the lender will be repaid, even if the borrower is unable to afford the loan without neglecting other expenses (like rent or food) or falling into a cycle of debt.

Conclusion and Recommendations

Unfortunately, our analysis concludes that big data does not live up to its big promises. A review of the big data underwriting systems and the small consumer loans that use them leads us to believe that big data is a big disappointment. More and more, consumers are leading robust lives online. However, as data about consumers proliferates, so does bad data.

Key Federal Policy Recommendations

- The Federal Trade Commission (FTC) should continue to study big data brokers and credit scores testing for potential discriminatory impact, compliance with disclosure requirements, accuracy, and the predictiveness of the algorithms.
- The FTC and the Consumer Financial Protection Bureau (CFPB) should examine big data brokers for legal compliance with FCRA and Equal Credit Opportunity Act (ECOA).
- The CFPB should create a mandatory registry for consumer reporting agencies so that consumers can know who has their data.
- The CFPB, in coordination with the FTC, should create regulations based upon the FTC's research that:
  a. Define reasonable procedures for ensuring accuracy when using big data;
  b. Specify a mechanism so that consumers can do a meaningful review of their files including all data points that can be linked to that consumer (not just those that identify the consumer explicitly); and
  c. Define reasonable procedures for disputing the accuracy of information.
- The CFPB should require all of the financial products it regulates to meet Regulation B's requirements for credit scoring models.

Analysis of Big Data Loan Products

PRODUCT	PROVIDER	STATE	COSTS		TERMS	APR WITH FEES	INSTALLMENT PAYMENTS	COLLECT ELECTRONIC BANK INFORMATION	FINANCIAL EDUCATION
Great Plains Lending	ThinkFinance	Nat'l	Varies by amount From $91.68 to $2386.84		Bi-weekly payments	Varies by amount 349.05% to 448.76%	✓	✓	✓
LendUp	LendUp	CA	Varies by loan amount and length From $10.70 to $44		30 days	Varies by loan amount and length 199.53% to 748.77%	Not available to first time borrowers.	✓	✓
MySalaryLine	ThinkFinance	AZ, MO	$150	AZ: $7.50 plus 14¢ daily MO: 55¢ daily	Next Pay Date	MO: 134%			
			$300	AZ: $15 plus 29¢ daily MO: $1.10 daily					
			$500	AZ: $25 plus 48¢ daily MO: $1.83 daily					
Plain Green	ThinkFinance	Nat'l	Varies by amount From $189.52 to $1979.84		Bi-weekly payments	Varies by amount 299.17% to 378.95%	✓	✓	✓
Presta	ThinkFinance	Nat'l	Varies depending on monthly payment (For an iPad 4*, $23 weekly payment, $64 initial payment, effective fees of $738)		Weekly payments	Varies by product	✓		
RISE (Formerly Payday One)	ThinkFinance	CA, DE, ID, LA, MO, NM, OH, SC, SD, TX, UT, WI	Varies by state, plus interest: Up to $735 in TX, $693 in OH		Bi-weekly payments	Varies by state 299.16% to 358.85%	✓	✓	✓
Spotloan	ZestFinance	All states except MA, MO, ND, and WV	Varies by loan amount and length From $206.04 to $1572.69		Bi-weekly payments	390%	✓	✓	

The information on this chart is based upon publicly available information found on the following products' websites on Dec. 11, 2013.

INTRODUCTION

Approximately 64 million consumers in the United States have no credit history or lack sufficient credit history to generate a credit score, cutting off access to traditional banking services. Finding a way of getting affordable access to credit is of vital importance to the economic well-being of this population. It also represents an untapped market with the potential for big profits. So it is unsurprising that in this era of big data, information culled from Internet searches, social media, and mobile apps would be put to use towards that goal.

Big data makes big promises. It promises to make better predictive algorithms that in turn can make better products available to the unbanked and underbanked. But can big data live up to this big promise?

Big data products claiming to hold the key to unlocking the mystery of low-income consumers' creditworthiness must be able to show that they actually do what they claim to do. Some have suggested that big data is merely noise. As Nate Silver writes in *The Signal and the Noise*:

> *If the quantity of information is increasing by 2.5 quintillion bytes per day, the amount of useful information almost certainly isn't. Most of it is just noise, and the noise is increasing faster than the signal. There are so many hypotheses to test, so many data sets to mine— but a relatively constant amount of objective truth.*[1]

According to Tomaso Poggio, an MIT neuroscientist who studies how our brains process information, the problem is that evolutionary instincts lead us to see patterns where there are none–"finding patterns in random noise."[2]

Big data products must also show that they can meet not just the goals but also the ideals of consumer protection laws. They should operate with transparency, accuracy, and relevancy. Despite existing consumer protection laws giving consumers easy access to their credit reports, traditional credit reports are known to have high rates of error. Adding to the number of data points with data of questionable quality seems unlikely to result in higher rates of accuracy for consumers.

Finally, big data products must operate in a way that is fair and free from discrimination. Different communities use and access technology in different ways. The data that is mined often has different implications for different populations. Big data must not lay the groundwork for lending that discriminates against vulnerable consumers— whether intentional or unintentional.

Companies are starting to use big data to make decisions about whether to offer loans to consumers and on what terms. When analyzing this use of big data, consumers and policy makers should be concerned with these questions:

1. Are the decisions based upon accurate data?
2. Can the algorithms, when fed with good data, actually predict the creditworthiness of low-income consumers?

3. Does the use of big data in reports used for credit, employment, insurance, and other purposes comply with consumer protection laws?
4. Is there the potential for a discriminatory impact on racial, geographic, or other minority groups?
5. Does the use of big data actually improve the choices for consumers?

The public literature reveals surprisingly little about how big data brokers and users of big data operate. Unfortunately, our investigation, detailed in this report, found that big data turns out to be a big disappointment. The data brokers we investigated provided very little data and the data they did provide had many errors. Moreover, the products we reviewed failed to provide more affordable products for low-income consumers.

DIGITAL DEMOGRAPHICS

Historically, issues related to technology and privacy were seen as middle-class consumer issues. However, now that the Internet is increasingly a requirement for social and economic inclusion, these issues impact low-income consumers to a much greater extent. As low-income consumers use the Internet more, lenders and data brokers have more tools to analyze the credit potential of more low-income consumers.

The Pew Internet & American Life Project catalogs the Internet habits of individuals and families. In the lowest-income demographic surveyed, 76 percent of adults used the Internet.[3] However, disparities still exist in how low-income consumers access the Internet. For example, 65 percent of consumers making less than $25,000 a year lack access to broadband in the home.[4] Lower-income households with a member who owns a Smartphone are more likely than higher-income households to access the Internet primarily using a mobile device.[5] Of adults that earn less than $30,000 a year, 41 percent own a Smartphone.[6]

Social media use among lower-income consumers is also widespread. Of households that make under $30,000 per year, 77% frequent social media sites.[7]

Of adults that earn less than $30,000 a year, 41% own a Smartphone; 77% frequent social media sites.

To date, these communities have been underserved by traditional lenders, so there is an opportunity for lenders to use big data to provide credit products to them. However, it is unclear as to whether doing so will be beneficial for the low-income consumer. These products may fill a void and provide affordable access to credit to these underserved populations *or* they may be a means of preying on vulnerable communities.

BACKGROUND ON BIG DATA

The rapid evolution of technology has ushered in the rise of what some industry analysts dub "the Decade of Big Data." The McKinsey Global Institute defines big data as "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze."[8] However, in common usage (and for the purposes of this report), big data means the massive amounts of data that consumers generate in everyday life–through business transactions, e-mail messages, photos, surveillance videos, web traffic, activity logs stored in giant structured databases, or unstructured text posted on the web, such as blogs and social media.[9] In the last decade, the amount of data generated has grown exponentially, partially due to the rise of web tracking techniques and the increasing use of Internet-enabled mobile devices. As the amount of available data has grown, innovations in computing capability, the falling cost of data storage, and advances in statistical analysis make it easier to interpret and monetize data.

The private sector, government agencies, and nonprofits are taking advantage of the proliferation of data to transform the way they operate. Private industry has harnessed the power of big data to develop sophisticated advertising campaigns. Companies target potential customers whose interests and demographic information they have identified through social networking data, web browsing history, and purchase information. Target, for example, can reliably predict which shoppers are pregnant based on the history of products purchased at the store, combined with other demographic information purchased from third-party data brokers.[10] Overall, business customers spend $45 billion a year for data.[11]

It's unsurprising amidst all this digital noise that lenders would seek to capitalize on big data to drive credit decisions. Douglas Merrill is the former chief information officer (CIO) at Google and founder of ZestFinance. At Google, Merrill managed the rise of one of the world's largest data companies. Now, he's deploying the analytical tools and technological savvy he cultivated at the search engine behemoth to transform subprime credit underwriting.

According to Merrill, "[a]ll data is credit data."[12] Merrill founded ZestCash in 2009 but re-named the company ZestFinance after switching its focus from directly lending small dollar loans to selling the data analysis it provides to other lenders of subprime products. Instead of evaluating potential borrowers based on a FICO score, which uses 10-15 variables to arrive at its score, ZestFinance renders a credit decision after analyzing thousands of variables.[13] The company runs the variables through ten different models. By expanding the number of variables, the company argues, the credit decision will more accurately reflect the risk a person presents. Subprime borrowers, who typically have poor FICO scores and therefore pay much higher interest rates on loans, may actually turn out to be good credit risks. In conjunction with the algorithms using big data, new lines of financial products have been introduced targeting unbanked and underbanked populations. However many of these products are very expensive and may not be beneficial.

THE BIG DATA ECOSYSTEM—HOW DOES IT WORK?

Thousands of companies specialize in data, but three different functions exist: data collection, data aggregation, and data analysis.

Data Collection

To understand data collection, it's important to understand how data is created. With the introduction of Internet-enabled devices (computers, mobile phones, and tablets), the amount of data that a consumer generates is enormous. Between 2006 and 2011, the amount of data generated increased by a factor of nine to 1.8 zettabytes (1.8 trillion gigabytes).[14] Each time a consumer visits a website, makes a purchase, or indicates a preference on Facebook or other social networking sites, data is created.

For example, a woman interested in purchasing a mystery novel will sit at her computer and open a web browser. She types "Amazon.com" into the URL line. By typing in the URL, her computer requests the page from Amazon's server. The computer transmits its Internet Protocol (IP) address to the webpage. An IP address is similar to a brick-and-mortar address, in that each address is unique. Based on the woman's IP address, the website's server can predict her zip code (with varying degrees of accuracy).[15] Amazon's server sends the webpage and downloads a "cookie" (line of text) onto the woman's hard drive. Several other third-party marketing firms that contract with Amazon may also download cookies. A cookie can contain various types of information, including (but not limited to) the time of her visit, the subpages she visited, and the items she purchased. Cookies also typically designate a unique ID to one's computer. By assigning a unique ID, third-party tracking companies can see other pages a person visits, intuiting preferences.

This data is incredibly valuable to marketers and there are few restrictions on such data in the U.S. This data can be bought and sold at will.

Third-party tracking companies also may embed a piece of software called a "web beacon" which not only can track which webpages a person visits, but also record the text typed in. For example, if a webpage has a beacon on it, then when a person uses the "search" function on a webpage, such as Amazon's, that information is relayed to a third-party marketer.[16] Subsequent pages that the person visits are summarily tracked. If the woman purchases a few mystery novels from Amazon and then books a flight for a family vacation, surfs the web for the latest political gossip, and "Googles" the best rates for car insurance, a third-party tracking company may capture every single move she makes.

A rich portrait of individuals emerges from the ability to track their online behavior. From purchase histories to search topics, a completely unedited and unmediated version of a person emerges. This data is incredibly valuable to marketers and there are few restrictions on such data in the U.S. This data can be bought and sold at will.

Web crawling is another technique that companies can use without developing a relationship with a host page. Web crawling involves the duplication and categorization of information from websites, typically by automated means. Programmers can write software that scans websites and sorts posts. Rapleaf, a tech company, used web crawling techniques to scan posts from Facebook.[17] Social Intelligence Corp. collects data on individuals by deploying web crawlers to analyze Facebook profiles and pictures. Individuals can be categorized based on groups that they "like" or comments posted.[18] Based on this data, the company sells information in the form of a background check report that prospective employers may use to determine the consumer's eligibility for employment. Photos tagged of the individual by other users may also be included in the report.

Data Aggregation

The ability to combine and cross-reference this data with other data creates an enormous opportunity to expand the information available about a particular individual. Data aggregation is the process of combining an array of data or data sources to compile a comprehensive portrait of an individual, behavior, or characteristic. ZestFinance, for example, combines data from alternative credit bureaus with data gleaned from web crawling to make a decision about whether to loan money to individuals.[19]

Companies have sought to make data aggregation easier by creating platforms that reformat data to make it uniform. Zoot Enterprises, for example, buys data from fourteen major databases and allows business clients to conduct searches across all fourteen databases.[20]

Data Analysis

Data analysis is completed by running either raw data or aggregated data through a series of models (usually called algorithms) to reveal patterns or test hypotheses.

While the collection, aggregation, and analysis are all distinct steps in using big data, they are not necessarily performed by separate actors. ZestFinance, for example, buys data from data brokers but also collects its own data through web crawling.[21] It combines the data and runs it through ten separate models before rendering a credit decision. Most companies use a hybrid model where they perform their own proprietary analysis on data obtained from multiple data brokers, aggregators, or other sources. As discussed in detail in the next section, depending on the structure of the company, many of the activities of the actors performing these three steps are subject to the regulations of the Fair Credit Reporting Act.

SUPERSIZE IT: IS BIGGER ALWAYS BETTER?

Big data proponents argue that multiplying the number of variables will expand access to borrowers with thin credit files. Thus, they claim that big data will be used to generate a credit score that gives creditors a fuller picture of a consumer and therefore gives a more accurate and robust predication of the consumer's ability to repay. While that potential may exist, it is unclear that this is what actually occurs. Big data only generates better results if the algorithm is predictive and if the data that feeds it is accurate. At present, there is no mechanism in place to ensure the integrity of credit scores generated by big data.

Certainly, problems exist with the traditional credit scoring system. First, credit scores cannot predict if any particular person will actually engage in the behavior. In fact, often the probability is greater that a particular low-scoring person will not engage in the behavior. Second, many low-income consumers have low credit scores simply because they have either a "thin file" or "no file." This means that they have very little reported credit history—often because low-income consumers are less likely to access the types of financial services that report to the traditional credit bureaus. A denial of credit to these consumers is based on the absence of credit history rather than anything negative in their credit histories.

Big data credit scoring models attempt to address both of these critiques of traditional credit scoring. They claim that by expanding the data points in their algorithms, they can create a more refined predictive score. Also, by expanding the type of data analyzed, they claim that they enable lenders to extend access to credit to traditionally underserved populations.[22]

Creating better credit scores and increasing access to credit for the estimated 64 million consumers[23] who have little or no information in traditional reports at the major credit bureaus (Equifax, Experian, and TransUnion) are laudable goals. However, expanding the type of information used also carries risks.

Credit Score Basics

A credit scoring system is one that numerically weighs or ''scores'' some or all of the factors considered in the underwriting process. Factors are developed based on data about past borrowers from their files at consumer reporting agencies and sometimes from other sources. Examples of factors used in a traditional credit scoring system include:

- history of payment of past obligations,
- amounts owed,
- length of credit history, and
- types of credit already held.

The number of points received often determines whether the consumer is offered credit, how much credit is granted, and at what price.

Credit scores are used to predict the probability that consumers will engage in a particular behavior, e.g., miss a payment, default, or file for bankruptcy. Because a credit score is generated based on information in a consumer's credit file, it will change as the information in the consumer's credit file is updated. The leading creator of models is FICO, formerly known as Fair Isaac & Co. Even though FICO develops other types of credit scores, a credit risk score is sometimes referred to as a "FICO score." The "Big Three" credit reporting agencies (Equifax, Experian, and TransUnion) have developed their own credit risk score model to compete with FICO, called the "VantageScore."

Like promoters of big data, promoters of "full file" utility credit reporting claim that using utility data will assist thin file or no-file consumers to build credit histories and gain access to credit. However, full file utility credit reporting could end up harming consumers' credit scores or give them low scores instead of no scores. Many low-income customers would receive negative marks for a 30- or 60-day late payment during months when utility costs are high, even though they eventually catch up when costs are lower (e.g., in summer months for cold winter states). Financially distressed consumers could (and should) be prioritizing payment based upon whether their utilities will be shut off so they can afford to pay for food or other critical items and defer their utility payments until later. Also, the federal assistance Low Income Heating Energy Assistance Program (LIHEAP) actually requires consumers to receive a shut-off notice before they can get assistance in some states. Reporting utility information means that these consumers cannot access LIHEAP assistance without damaging their credit records. In this way, even if "full file" utility credit reporting is predictive of creditworthiness, it unfairly punishes vulnerable consumers for making the best financial decision for their families.

Given the breadth of personal and potentially sensitive information big data brokers collect, credit scoring models based upon big data must be analyzed to determine their true impact on low-income consumers. Specifically, consumers and policy makers should be concerned with the integrity of the data in three ways:

1. accuracy of the data used;
2. verifiable predictiveness of the algorithm; and
3. potential discriminatory impact.

DATA ACCURACY: GARBAGE IN, GARBAGE OUT?

While big data enthusiasts highlight important flaws in the current credit scoring system, critics have identified drawbacks to expanding credit variables from fifteen to thousands. Instead of minimizing the impact of an unimportant credit signal, a big data approach could amplify the significance of a completely irrelevant signal. Nassim Taleb, a risk engineering professor at New York University, is a vocal critic of big data. Taleb said, "…if I generate…a set of 200 variables—completely random and totally unrelated to each other—with about 1,000 data points for each, then it would be near impossible *not* to find in it a certain number of 'significant' correlations of sorts."[24] With a thousand different variables—from prepaid cell phone payments to rental payments to social media histories—the correlation among variables can confuse instead of clarify.

Expanding the number of data points also introduces the risk that inaccuracies will play a greater role in determining creditworthiness. More data does not necessarily mean better data. In a 2013 study by the Federal Trade Commission, researchers found that 20 percent of traditional credit reports had errors; 5 percent of credit reports contained errors that could result in a lower credit score, making credit inaccessible or costlier.[25] The traditional credit bureaus are highly centralized and finite—there are only three—but hundreds of other consumer data brokers exist that provide alternative credit information and other types of consumer data. Even with a relatively centralized system, it can be difficult to get mistakes corrected among the "Big Three" credit bureaus (Equifax, Experian, and TransUnion). If data aggregators or data analysts harvest data from dozens of sources, inaccuracies are harder to detect and the source of the error can be difficult to identify.

A small research study published in 2005 suggests that these accuracy problems are not merely hypothetical. Researchers requested files from two major data brokers—Acxiom and ChoicePoint—that provide lenders with alternative credit data. Eleven of eleven ChoicePoint reports contained at least one error. Although eleven testers requested their files from Acxiom, the company only mailed six. Four of the six reports reported incorrect information.[26]

In addition, the National Consumer Law Center's (NCLC) 2012 report *Broken Records*, reviewed the inaccuracies endemic to the criminal background check industry and found criminal background checks to be rife with errors.[27] Many of these background check agencies rely on unverified, incorrect, or outdated data available on the Internet, rather than doing the more difficult or expensive research to track down more accurate information.

Given these indications of accuracy problems, we conducted our own survey for this report of the data maintained on consumers by big data brokers. Even given our initial skepticism, we were astonished by the scope of inaccuracies among the data brokers we investigated.

More data does not necessarily mean better data. In a 2013 study by the Federal Trade Commission, researchers found that 20% of traditional credit reports had errors; 5% of credit reports contained errors that could result in a lower credit score, making credit inaccessible or costlier.

NCLC's Study of Big Data Accuracy

In an attempt to learn more about big data brokers, fifteen volunteers, all whom were NCLC employees, attempted to retrieve their information from four different data brokers: eBureau, ID Analytics, Intelius, and Spokeo. By either purchasing or requesting the consumer file from each of these companies, NCLC hoped to illuminate the type of data collected, the accuracy of the data, and the ease of obtaining a consumer report. Report authors Persis Yu and Jillian McLaughlin also requested reports from Acxiom. Other volunteers did not request reports from Acxiom because of the financial cost and because, in light of a *New York Times* article published prior to our study, no information was expected (see p. 17). Though they were all NCLC employees, the volunteers ranged in age, work experience, income, education, and social media presence.

The five big data companies chosen. We chose the five companies based on several factors, including representativeness of the industry, the variety of users likely to buy the data, and the relative ease with which information could be requested. Data from two of the five companies (Spokeo and Intelius) was purchased via a subscription or individual report.

Acxiom is believed to have amassed the world's largest commercial database on consumers.[28] It claims to provide insight into consumers' preferences and behaviors with its data products and services.[29] One of Acxiom's products lets corporate clients purchase hundreds of details about individuals or households, such as whether the household size, income, and even whether it has concerns about allergies, diabetes, or "senior needs."[30] As mentioned previously, only co-authors Persis Yu and Jillian McLaughlin attempted to obtain reports from Acxiom.

eBureau also generates a "score" for use by clients in determining a consumer's credit risk. eBureau claims to analyze inputs from thousands of databases, storing billions of records in its warehouses.[31] Interestingly, eBureau touts its ability to estimate income based on its advanced models. According to its promotional materials:

> *eBureau's Income Estimator utilizes dozens of different predictive data sources, including independently compiled sources of demographic data, real asset information, and spending behavior. Income Estimator's sophisticated scoring model factors in 375 discrete individual, household, and neighborhood variables to produce a highly accurate estimate. It has been validated against hundreds of thousands of self-reported income records from across the US…*[32]

ID Analytics primarily serves clients using "identity intelligence," which monitors the name, address, Social Security number, phone number, and date of birth consumers disclose. The company uses this information to create an alternative credit score; in theory, the alternative score predicts financial stability. For example, if consumers change addresses or phone numbers frequently, they are likely to be less stable than consumers who consistently report the same number and address.[33] ID Analytics collects consumer information from cable providers, cell phone companies, and checking account history.[34]

Intelius appeals to a broader market segment than eBureau and ID Analytics, which primarily target financial services providers. Intelius stores more than 20 billion records and compiles data on individuals for a variety of purposes—hiring, dating background checks, and fraud prevention. Intelius specifically markets its data to businesses and consumers.[35]

Spokeo focuses on a consumer audience and specifies on its site that the information collected should not be used for any purpose listed under the Fair Credit Reporting Act (FCRA), including employment screening, credit decisions, and insurance eligibility (see page 26).[36] Spokeo suggests that its site can be used to identify the holder of a telephone number, help families reunite with loved ones, and assist nonprofits and small businesses in identifying potential donors or customers.

The five companies represent a range of services, markets, and tactics. Four of the five companies are on the radar of the FTC. In December of 2012, the FTC ordered Acxiom, eBureau, ID Analytics, and Intelius to disclose to the agency their methods of data collection and privacy practices.[37]

Obtaining the reports. In general, obtaining the reports was challenging. At the time, Acxiom,[38] purportedly the world's largest data company, required consumers to submit an online request form and then physically mail a personal check for $5 to cover the cost of processing the request.[39] When an NCLC researcher asked if Acxiom would accept a money order instead of a personal check, Acxiom stated that it would accept a money order only with a notarized copy of the consumer's identification document.

On their websites, ID Analytics and eBureau provide users with a standardized process for requesting individual consumer reports. However, after our volunteers submitted the voluminous information identified on its website, eBureau sent a letter requiring each volunteer to verify his or her Social Security number by sending a copy of the card, W-2, or other official document.[40] ID Analytics required two volunteers to submit additional information.

Two volunteers were unable to find their information when searching Spokeo, and one of those was also unable to find information from Intellius. The volunteer who received no Spokeo or Intellius information has a unique name and numerous social media accounts.

The data obtained. The reports from Acxiom, eBureau, and ID Analytics contained very little information. In contrast, the reports from Intellius and Spokeo were more robust.

Acxiom. Our experience was similar with that of a journalist from the *New York Times* who requested her report from Acxiom in 2012. She expected to see most of the 1,500 data points the company claims to amass per consumer. Instead, the company sent a report listing the journalist's previous residential addresses. She wrote, "For a corporate client, the company is able to match customers by name with, say, the social networks or Internet providers they use, but it does not offer consumers the same information about themselves."[41] In our experience, one co-author's Acxiom report included current and previous residential addresses as well as a very incomplete voting history; the other co-author's Acxiom report also included an incorrect middle initial of her name, current and previous residential addresses (current address was incorrect), as well as an inaccurate and incomplete voting history.

eBureau claims to utilize "vast amounts of predictive data to offer instant insights across multiple industries, from higher education to financial services to automotive and insurance marketers."[42] Yet, the reports by eBureau only had nine fields: first name, last name, address, phone number, Social Security number, date of birth, income, education, and length of residence.

ID Analytics claims to help companies "optimize[] credit decisions about individuals to maximize revenue opportunities and reduce risk"[43] through its data and algorithms. Yet, the reports it gave to our volunteers contain just ten fields: name, address, social security number, phone, date of birth, name variations, date of birth variations, Social Security number variations, address history, and phone number history.

Intelius claims to be "a confidential way to find people so you can reconnect or just get more info on a person." Its People Search reports include phone numbers, address history, age and date of birth, relatives, and social media profiles.

Study Participants with Incorrect Information in Their Data Reports

Of the 15 participants' reports, many contained errors (see light red figure) or there was no record found (asterisk). Acxiom was not included in this study.

*No records found

Spokeo reports provide the most information. They include personal data such as, name, address, phone number, email address, age, marital status, education level, family members, and social media profiles. The reports also included demographic information about the volunteer who requested the file, such as area home values, occupations, median incomes, race and gender statistics, and average age.

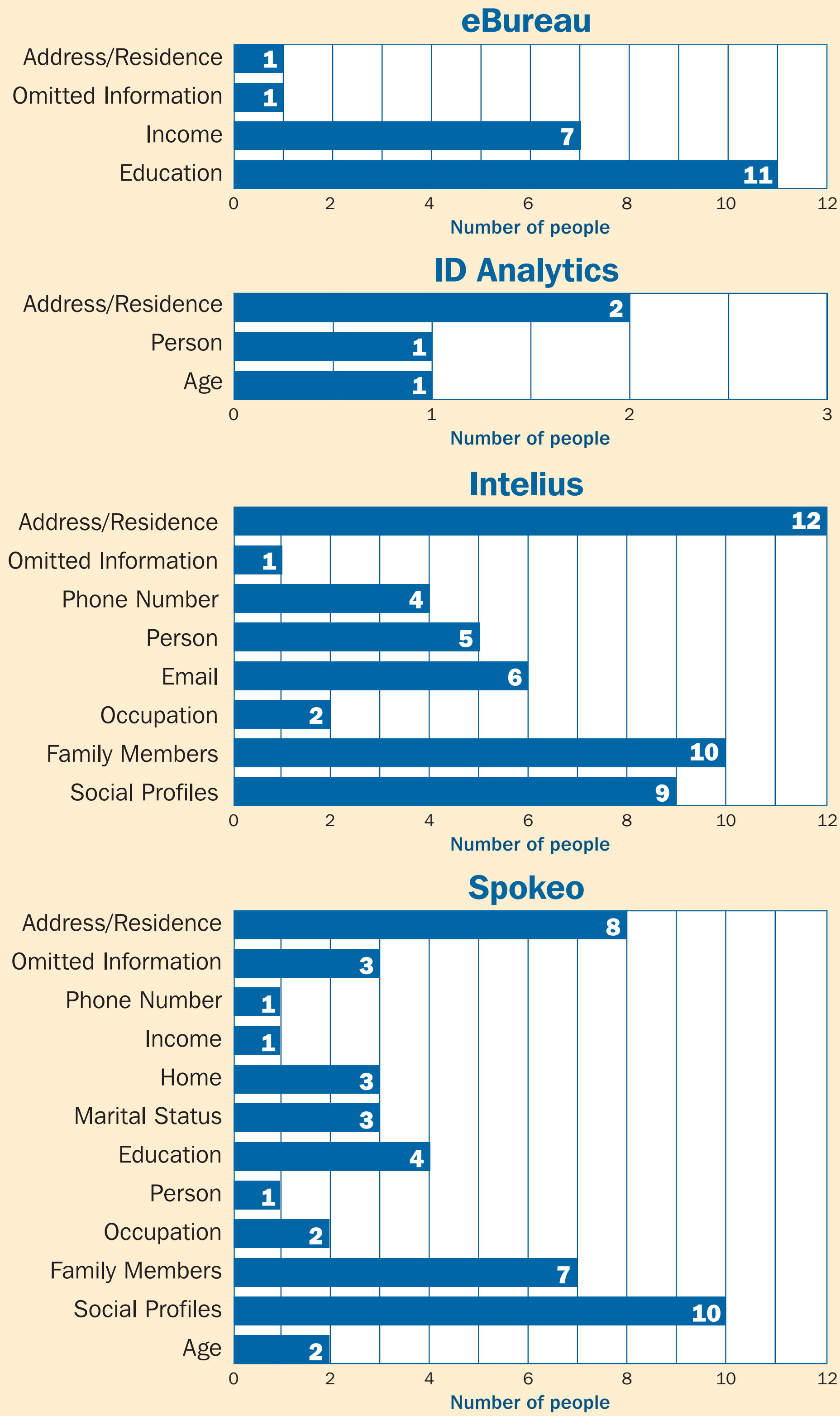
Given the claims by each of these data broker companies, the sparse information they produced for our volunteers may only be a fraction of the data the company stores about them. Such an omission may violate the Fair Credit Reporting Act.

Accuracy. The reports our volunteers received were riddled with inaccuracies. Errors ranged from the mundane—a wrong e-mail address or incorrect phone number—to seriously flawed. One of the reports combined information about our volunteer with information about two other individuals; other reports listed wrong addresses, relatives, and occupations. Some reported home addresses in states in which the volunteer never resided. Interestingly, eBureau touts its ability to estimate income based on its advanced models[44] and offer insights based upon the consumer's education. Despite that claim, seven of the fifteen consumer reports generated by eBureau contained errors in estimated income, nearly doubling the salary of one participant and halving the salary of another, and eleven of the fifteen reports incorrectly stated the volunteer's education level.

Reports purchased from Intelius and Spokeo had the most inaccuracies. The most common errors in both were wrong address (twelve out of fourteen volunteer reports and eight out thirteen volunteers respectively), added or omitted immediate family members (ten and seven respectively), and added or omitted social media accounts (nine and ten respectively).

Study Participants with Mistakes in Their Data Report (per Company and Category)

Of the 15 study participants' reports, there were a number of errors. Acxiom was not included in this study.



Verifying the Predictiveness of Big Data Credit Scores

Aside from the unverified claims by companies profiting from the algorithms used to analyze big data, consumers have no way to know whether the algorithm accurately predicts their creditworthiness. Though a similar critique is certainly true of FICO and other traditional credit scores, consumers are given a general understanding of how those scores work and roughly how different variables are weighted. This gives consumers a guidepost for predicting their own creditworthiness and potentially adjusting their behavior in order to improve their creditworthiness.

With big data, there is no independent source confirming the accuracy or reliability of the algorithms used to generate a predictive score. Nor is there transparency regarding how the score is calculated. Consumers obtaining loans based upon this score have no real way of knowing whether the loan really is tailored for them or whether this is an elaborate marketing scam.

As discussed next, the FCRA does not explicitly require credit scores to be predictive of creditworthiness. However, Regulation B,[45] the implementing regulation for the Equal Credit Opportunity Act (ECOA), does address predictiveness.[46] Regulation B requires that a credit scoring system satisfy four criteria:[47]

1. The data used to develop the system must constitute either the entire pool of applicants or an appropriate sample of applicants who applied for credit within a reasonable preceding period of time;[48]
2. The system must have the purpose of predicting applicants' creditworthiness with respect to "legitimate business interests" of the creditor using it;[49]
3. The system must be "developed and validated using accepted statistical principles and methodology";[50] and
4. The system should be periodically reviewed and re-validated as to its predictive ability and adjusted accordingly.[51]

Regulation B itself makes limited use of this definition of a credit scoring system, referring to it only with respect to when creditors may consider information about age and public assistance status.[52] However, the practical importance of this definition is much greater, as some of the banking regulators have required the banks they regulate to meet Regulation B's requirements for credit scoring models.[53] The Consumer Financial Protection Bureau (CFPB) could similarly require products that it regulates to meet these requirements.

Applying the Fair Credit Reporting Act (FCRA)

One of the most powerful tools consumers and regulators have to ensure a fair and accurate credit reporting system is the FCRA. As the use of big data purports to address some of the deficiencies of traditional credit reporting, it must also be held to the same consumer protection standards.

FCRA Background

The FCRA is a federal statute that first became effective April 25, 1971. It regulates the activities of consumer reporting agencies (CRAs), the users of reports, and those who furnish information to CRAs (furnishers). The Act also provides remedies to consumers affected by such reports.

The FCRA attempts to protect consumers' privacy and reputations by placing various obligations on persons who use or disseminate credit information about consumers. For example, CRAs must adopt reasonable procedures to ensure that the information they disseminate is accurate and up-to-date and that it is furnished only to users with certain permissible purposes. The Act also imposes disclosure obligations for both CRAs and users. These are designed to ensure that consumers will know when a consumer report has been used as the basis of action adverse to their interests, and that consumers will know about the information being disseminated about them. CRAs also must reinvestigate information that consumers dispute and inform users of the dispute. Those who furnish information to CRAs must participate in an agency's reinvestigation and are subject to other duties. Importantly for consumers, the Act provides consumers with a civil remedy (meaning the right to sue for damages) for most violations of the Act.

The term "consumer reporting agency" refers not just to credit bureaus, but also to many other entities that meet the statutory definition. This may include creditors, data brokers, employment screening companies, check approval companies, alternative credit bureaus, and others.

Under the FCRA, "consumer reporting agencies" (CRAs) are companies or nonprofits that provide consumer reports to third parties for the purposes of determining eligibility for credit, insurance, employment, or other business transactions.[54] The definition of a "consumer report" is fairly broad. It is a written, oral, or other communication of any information by a CRA bearing on one of seven factors:

- a consumer's creditworthiness,
- credit standing,
- credit capacity,
- character,
- general reputation,
- personal characteristics, or
- mode of living.[55]

The FCRA also has a special definition for a "credit score."[56] However, credit scores also fall within the general definition of a consumer report.[57]

Since a report need bear on only one of the seven factors listed in the statute, a wide variety of information about a consumer satisfies this part of the definition of a consumer report, including most of the information collected by big data brokers.[58] For example, these reports claim to assemble details about property values and ownership; likely income and assets; an applicant's educational background; professional licenses; phone service history; subprime credit information, such as use of a payday loan; and ownership of boats and airplanes as a way of assessing credit risk. Other information assembled in these reports from social media and spending patterns deals with personal characteristics or mode of living, and would be considered a consumer report if used to determine whether to extend credit, employment, or other purpose under the FCRA. For example, creditors believe that those who buy birdseed, snow rakes, and felt pads for furniture are good credit risks, while those who buy chrome-skulls are not. Reports of these types meet the definition of consumer report.[59]

The second prong of the definition of a consumer report narrows the broad scope of the first prong. It requires that the information must be "used or expected to be used or collected in whole or in part" for certain purposes— determining eligibility for credit, insurance, employment, or certain other business transactions.[60]

Databases that Do Not Name the Consumer

With some of the big data databases, data may be stored without specifically identifying an individual consumer by name. This does raise questions about whether these data brokers are CRAs since a consumer report is one "bearing on a *consumer's* credit worthiness" and other factors.[61] Consumer is defined in the FCRA as an individual.[62] Thus, at a minimum, a consumer must be "an identifiable person." Therefore, a report on an anonymous computer username is not a consumer report.[63]

However, citing advancements in technology and the public availability of a broad range of data about consumers, the FTC clarified in its 2011 Staff Summary that information may constitute a consumer report even if it does not identify the consumer by name if it could "otherwise reasonably be linked to the consumer."[64]

In some cases, the consumer's name may actually be irrelevant. Hypothetically, if an online lender uses an analysis of the websites a potential borrower views based upon the cookies on the computer applying for the loan, then the most important piece of identifying information for that consumer may be the IP address and not the potential borrower's name. Still, since the lender is using the IP address as a proxy for an individual, a report about that IP address should be considered a consumer report because it can reasonably be linked to the consumer who will be repaying the loan.

What Consumer Reporting Agencies Must Do Under the FCRA

If the FCRA does apply to a big data database, it imposes substantial duties on the CRA. Three of the most important functions of the FCRA deal with accuracy, disclosure, and the right to dispute items on the report. It is highly unlikely, given the size of the data set and the sources of information, that the companies that provide big data analytics and the users of that data are meeting these FCRA obligations.

Accuracy

The FCRA requires CRAs to use reasonable procedures to ensure maximum possible accuracy of the information in a report.[65] Under the FCRA, the requirement of accuracy does not mean merely that the information must be technically accurate. Rather, accuracy encompasses the completeness of the information, the relevance, and the interpretation. In most jurisdictions, it is not sufficient for information to be literally true; it also cannot be misleading or incomplete.[66]

Three of the most important functions of the FCRA deal with accuracy, disclosure, and the right to dispute items on the report. It is highly unlikely, given the size of the data set and the sources of information, that the companies that provide big data analytics and the users of that data are meeting these FCRA obligations.

The accuracy requirement will be a challenge to many big data brokers. The first challenge will be ensuring that data is technically accurate. The Internet is full of bad information. Moreover, ensuring that information matches the correct consumer may be nearly impossible when consumers are only identified by their first and last names. Also, consumers using shared computers may have data improperly attributed to them. Our survey of the big data companies' reports showed a remarkable level of inaccuracy.

Completeness will also be a challenge with big data. Depending on the data source, many pieces of information will be snapshots in time. For example, a lender wanting to analyze patterns of online shopping may do so by using cookies embedded in the consumer's web-browser. However, those cookies will not include items that were returned or that were purchased as gifts.

Finally, courts consider a consumer report to be inaccurate when it is "misleading in such a way and to such an extent that it can be expected to [have an] adverse [effect]."[67] Information can be technically accurate but misleading in a number of ways. For example, a data broker might have information on a borrower's educational status and list the consumer's educational status as "completed high school." If the consumer has also completed a four-year college program, this may be technically correct, but it implies that the consumer's education ended after high school.

The way that information is displayed may also be misleading. NCLC's report *Broken Records* described employment background checks that used line spacing that made the list of offenses appear to be longer than it actually was.[68] A scoring model that is not actually predictive of creditworthiness could also be considered misleading. The FCRA does not explicitly require credit scores to predict what they claim to predict. However, scoring models that claim to predict creditworthiness but fail to do so are misrepresenting the consumer's information. A score that incorrectly claims to predict a consumer's creditworthiness portrays the consumer's data in a light that is not true.

Disclosure

To know whether information on a consumer report is accurate, the consumer must know that such a report exists and what information it contains. For this reason, the FCRA gives consumers the right to find out what is in the files maintained by the CRA.

The law gives consumers the right to request the information in their file at a CRA–often for free.[69] Moreover, when a user takes an adverse action relating to credit, insurance, or employment, in whole or in part because of information in a consumer report, the user must provide an adverse action notice to the consumer.[70] This notice identifies the CRA that supplied the consumer report and gives instructions on how to obtain a report. It alerts the consumer to the existence of the report and that it contains some adverse information that the consumer should probably check. Unfortunately, compliance with this notice requirement is sparse with non-traditional consumer reports.[71]

CRAs must clearly and accurately disclose to the consumer all information in the consumer's file at the time of the request.[72] A CRA violates the FCRA by refusing to provide this information or by providing only partial disclosure. The responses to the requests that our volunteers made for their consumer reports show that big data brokers are likely to fail to comply with the requirement to disclose all information in the consumer's file. As discussed, even though data may not be explicitly identified to a certain consumer, because it can be linked to the consumer, the CRA should disclose that information.

The right to dispute information is an important safeguard necessary to ensure the accuracy of the consumer's data and is one of the most important functions of the FCRA. However, compliance with this part of the statute may prove to be challenging, if not impossible, to data brokers collecting big data.

However, even if data brokers were to provide this disclosure, the information may not be comprehensible for consumers due to its sheer volume. Therefore, meaningful disclosure may not be possible when using big data. This may prove to be a fundamental flaw with using big data for determining eligibility for credit or other FCRA-covered purposes.

Right to Dispute Inaccuracies

One of the most critical protections provided by the FCRA is the consumer's right to dispute the accuracy or completeness of any item of information in his or her file. Upon receiving a dispute, the FCRA requires the CRA to conduct a reinvestigation, reviewing all relevant materials and contacting the source of any information. Any information that cannot be verified must be deleted.[73]

The right to dispute information is an important safeguard necessary to ensure the accuracy of the consumer's data and is one of the most important functions of the FCRA. However, compliance with this part of the statute may prove to be challenging, if not impossible, to data brokers collecting big data. Most of the information collected is gathered from the consumer without his or her knowledge. Based upon the way that data is collected and gathered, it may be difficult for a consumer to dispute and for the data broker to verify whether it really was the consumer visiting a certain website or making a certain purchase.

Given that unverifiable data should be deleted, theoretically, a consumer should be able to dispute all information in his or her file, and the CRA would be required to delete it if unverifiable. Given the reluctance of consumer reporting agencies historically to delete any information, whether a big data broker would comply is another matter.

Evasion of the FCRA

The FCRA's definition of "consumer report" covers a broad range of information (see page 21). This broad scope is narrowed by purposes to which the information is used. Thus, some data brokers have attempted to avoid liability under the FCRA by claiming that their products are not used for consumer reporting purposes and that the information is not assembled for the purpose of furnishing consumer reports.

One way that these companies have attempted to evade the FCRA is by including boilerplate language in their agreements stating that the information is not a consumer report and telling users that it may not be used for any permissible purpose under the FCRA.[74] However, this type of boilerplate should not be sufficient to exclude information from the definition of a "consumer report." (See examples of disclaimers on page 26.)

For example, one company, Spokeo, was sued by the FTC in 2012 for marketing its products to companies in the human resources, background screening, and recruiting industries without taking the steps to protect consumers that are required under the FCRA. Spokeo settled the suit for $800,000.[75] However, in its terms of use, Spokeo still attempts to disclaim its obligations under the FCRA.[76]

The Spokeo disclaimer should not comply with the law. Regardless of the ultimate use of the information, if an entity providing consumer information reasonably expects or reasonably should expect that the information might be used for FCRA purposes, and the entity does not have reasonable procedures in place to limit the uses to which the information can be put, then the entity should qualify as a CRA.[77]

The FTC has warned companies that the presence of a disclaimer stating that reports should not be used for FCRA purposes is not sufficient to avoid FCRA coverage.[78] Boilerplate language in an agreement is not sufficient to defeat the expectation that some users who have access to reports bearing on creditworthiness or other FCRA factors might use the information for FCRA purposes. As the FTC has stated to several data brokers:

> *If you have reason to believe that your [mobile application] reports are being used for employment or other FCRA purposes, you and your customers who are using the reports for such purposes must comply with the FCRA. This is true even if you have a disclaimer on your website indicating that your reports should not be used for employment or other FCRA purposes.*[79]

Examples of Data Broker Disclaimers to Sidestep the FCRA*

Accurint® for Collections does not constitute a "consumer report" as that term is defined in the federal Fair Credit Reporting Act, 15 USC 1681 et seq. {FCRA). Accordingly, Accurint for Collections may not be used in whole or in part as a factor in determining eligibility for credit, insurance, employment or another permissible purpose under the FCRA.

Intelius FCRA Restrictions. Intelius is not a consumer reporting agency as defined in the Fair Credit Reporting Act ("FCRA"), and the information in the Intelius databases has not been collected in whole or in part for the purpose of furn ishing consumer reports, as defined in the FCRA. You shall not use any of our information as a factor in (1) establishing an individual's eligibility for personal credit or insurance or assessing risks associated with existing consumer credit obligations, (2) evaluating an individual for employment, promotion, reassignment or retention (including employment of household workers such as babysitters, cleaning personnel, nannies, contractors, and other individuals), or (3) any other personal business transaction with another individual {including, but not limited to, leasing an apartment).

Rapleaf Compliance with Fair Credit Reporting Act. Rapleaf is not a consumer-reporting agency ("Consumer Reporting Agency") as defined by the Fair Credit Reporting Act, 15 U.S.C. § 1681 et seq. ("FCRA") and Data Services and reports do not constitute "Consumer Reports" as that term is defined in the FCRA. You agree to not use or provide the Data Services Data for any purposes enumerated in the FCRA in lieu of obtaining a Consumer Report.

Specifically, you agree not to use or provide the Data Services Data, or authorize anyone else to use or provide the Data Services Data, for the following purposes:

a. in connection with establishing a consumer's eligibility for credit or insurance to be used primarily for personal, family or household purposes, or in connection with assessing risks associated with existing credit obligations of a consumer;
b. for the purpose of evaluating a consumer for employment, promotion, reassignment or retention as an employee;
c. for any tenancy verification or in connection with any application to rent real property;
d. in connection with a determination of a consume(s eligibility for a license or other benefit that depends on an applicant's financial responsibility or status;
e. as a potential investor or servicer, or current insurer, in connection with a valuation of, or assessment of credit or prepayment risks associated with, an existing credit obligation;
f. in connection with any information, service or product sold or delivered to a "Consumer" (as that term is defined in the FCRA) that constitutes or is derived in substantial part from a Consumer Report;
g. for any other purpose covered under the FCRA; or
h. for the preparation of a Consumer Report or in such a manner that may cause such data to be characterized as a Consumer Report. You agree not take any "Adverse Action" (as that term is defined in the FCRA). which is based in whole or in part on Data Services or data, against any Consumer.

Spokeo You may not use Spokeo.com or any information acquired from Spokeo.com:

i) to engage in activities that would violate applicable local, state, national or international law, or any regulations having the force of law, including the laws, regulations, and ordinances of any jurisdiction from which You access Spokeo.com;

ii) to send any commercial email or text message that does not comply with CAN-SPAM, the Telephone Consumer Protection Act or any other applicable state law;

iii) to evaluate a consumer's eligibility for credit or insurance to be used primarily for personal, family, or household purposes, to evaluate a person's eli gibility for employment or volunteering purposes, to evaluate a person's eligibility for a government li cense or benefit, to evaluate a person for renting a dwelling property, or for any other purpose specified in the Fair Credit Reporting Act (15 U.S.C. § 1681b);

iv) in any manner that may violate any local, state, federal, or international privacy law to which You may be subject on the basis of Your location or the location of the person searched.

* Examples found on select data brokers' websites; not an inclusive list.

Furthermore, while it is sufficient if the CRA anticipates a listed use, this is not even necessary. It should be enough if, in the usual course of events, one would expect the report to be used for an FCRA purpose.[80]

EVALUATING THE DISCRIMINATORY IMPACT OF BIG DATA SCORES

Because big data scores use undisclosed algorithms, it is impossible to analyze the algorithm for potential racial discriminatory impact. According to the companies' marketing materials, consumers are judged based upon data generated from their Internet usage, mobile applications, and social media. However, access and usage of these sources vary by race and socioeconomic status, and thus any algorithm based upon them may have racial disparities.

Non-Hispanic white households have greater rates of broadband adoption than other socioeconomic groups. The adoption gap is widest between non-Hispanic whites and Hispanics (14 percent difference in adoption).[81] Thirty percent of whites use their mobile phone as their sole Internet connection compared to roughly 48 percent of Latinos and 39 percent of blacks.[82] Households that access the Internet solely using a mobile device are also more likely to be low-income.

Different races also use the Internet differently. For example, research by Pew and the Federal Reserve Board show that blacks and Latinos use their smartphones to do their banking more than any other race or ethnicity,[83] while whites are more likely to bank using a traditional desktop or laptop.[84] Additionally, according to Nielsen spokesman Matthew Hurst, "Black consumers are also 30 percent more likely to visit Twitter using mobile phones than the average customer."[85] These different ways of accessing the Internet leave a digital data trail. Yet, despite these known differences, little is known about how each of these variables is weighted or used by big data analytics.

Data gathered via the Internet is coded with an IP address. An IP address can be predictive of the consumer's zip code or even latitude and longitude.[86] Significantly, based upon census data, zip code and location can function as a proxy for race and income. There is already evidence that some companies target different zip codes differently. An investigative report by the *Wall Street Journal* found that the office supply store Staples priced its merchandise differently depending on the zip code gleaned from the consumer's IP address.[87]

While the discriminatory pricing of staplers may not be the gravest of injustices, the potential pitfalls of this type of pricing scheme raises concerns. If instead of office supply stores, banks and lenders engaged in this analysis, low-income and communities of color could be given higher interest rates and other less favorable terms.

Another potential concern relates to creditworthiness by association. Creditors have based a consumer's creditworthiness on the characteristics of others. For example, American Express lowered a customer's credit limit from $10,800 to $3,800, not based on his payment history with the company, but because "[o]ther customers who have used their card at establishments where you recently shopped have a poor repayment history with American Express."[88] With this type of analysis, low-income consumers with pristine credit histories could find their big data credit scores lowered simply because they save costs by shopping at low-end outlets whose customers include people who have trouble paying their bills.

> American Express lowered a customer's credit limit from $10,800 to $3,800, not based on his payment history with the company, but because "[o]ther customers who have used their card at establishments where you recently shopped have a poor repayment history with American Express."

There is already evidence that location is being used as a proxy for a consumer's ability to repay a debt. A recent report by TransUnion highlights this ominous trend:

> *…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.*[89]

Does using location and type of device in calculating a credit score violate federal credit discrimination laws? The answer to this question is complex, and depends on the product at issue.

There are two main types of discrimination theories under civil rights law: disparate treatment and disparate impact (or the "effects" test). Disparate treatment occurs when a business or employer treats a person differently on the basis of race or another prohibited basis (gender, age, religion, etc.). Disparate impact occurs when a business's policy or practice, neutral on its face, has a disproportionate negative impact on a protected group. Under this theory, the business's motive in treating applicants differently might not be race or another prohibited basis, but the effect is to adversely impact a particular protected class. The classic example of disparate impact is where an employer only hires people over a certain height. Because women are, on average, shorter than men, this policy would likely result in fewer women getting hired than men.

The Equal Credit Opportunity Act (ECOA) prohibits racial discrimination in the granting of credit, and is the federal anti-discrimination law that would likely apply to companies that use big data credit scores.[90] It prohibits not just disparate treatment, but also policies or practices that have a disparate impact.

In order to make out a "prima facie" (initial) case for disparate impact, the plaintiff must:

- **Identify** a specific policy (e.g., use of location) that has a discriminatory effect;
- **Show a disparate impact** of the policy on a group protected by anti-discrimination laws; and
- **Show causation**, i.e. a link between the policy and the disparate impact.

Making out a prima facie case of disparate impact does not necessarily mean that a practice violates the ECOA. Under the disparate impact analysis, a creditor or company can defend its policy by showing a "business necessity." Courts have articulated a number of different tests and definitions of "business necessity," including "compelling need," "manifest relationship," "legitimate, non-discriminatory rationale," and "demonstrably necessary."[91]

With respect to ECOA, regulatory interpretations of this Act state that creditors can defend a policy that produces disparate impact by showing "a demonstrable relationship between" the challenged policy and "creditworthiness."[92] Thus, if a variable or factor in a credit scoring model causes a disparate impact, but is "demonstrably related" to creditworthiness, it may be permissible under fair lending laws. The variable or factor, however, must be related to creditworthiness and not some other reason, such as generating maximum profit.

The business necessity analysis may differ for scoring models using large amounts of aggregated data as opposed to traditional credit reports. Traditional credit scores are based on credit histories, and supposedly measure the consumer's likelihood of repaying a loan. There is an understandable connection between timely repayment of past obligations and the likelihood of timely repayment of future obligations, so a "demonstrable relationship" argument can be easily made. While there might be some correlation between web searches, IP address, or social media posts and the likelihood of repayment, there has been no definitive understandable reason provided as to why those data points are a good measure of creditworthiness.

Finally, one should not rule out the possibility of a disparate treatment analysis.[93] Given the amount of personal information available online, it is possible, if not likely, that users of big data can discover the consumer's race, gender, religion, national origin, or other characteristics that lenders are prohibited from considering.

The Equal Credit Opportunity Act (ECOA) prohibits racial discrimination in the granting of credit, and is the federal anti-discrimination law that would likely apply to companies that use big data credit scores. It prohibits not just disparate treatment, but also policies or practices that have a disparate impact.

BIG DATA, BETTER PRODUCTS?

Proponents of big data underwriting argue that by using a constellation of factors to price credit, the cost of credit will be reduced for low-income borrowers, thus enabling lenders to provide lower-cost small loans as alternatives to payday loans. However, our analysis of loans priced according to big data underwriting challenges this assumption.

Elements of an Affordable Loan Versus a Payday Loan

Payday loans are very high-cost, short term loans that ensnare borrowers in a debt trap. The finance charge for a payday loan typically ranges from $10 to $30 for every $100 borrowed. Loans typically cost 400% annual interest or more.[94] The dangers of payday loans are well documented.[95] Payday loans lead to repeat borrowing and escalating cost. Taking out a payday loan increases the likelihood that the borrower will lose a bank account, file for bankruptcy, be subject to eviction, delay medical care, face a utility shut-off, and become delinquent on a credit card.[96]

In 2010, the National Consumer Law Center released a report, *Stopping the Payday Loan Trap: Alternatives that Work, Ones that Don't*, comparing different alternatives to payday loans. According to that report, a truly affordable alternative product that avoids the pitfalls of traditional payday loans must:

- Have an annual percentage rate (APR), including fees, of 36% or less;
- Have a term of at least 90 days, or one month per $100 borrowed;
- Require multiple installment payments rather than a single balloon payment;
- Not require that the borrower turn over a post-dated check or electronic access to a bank account; and
- Be issued after an evaluation of the borrower's ability to repay the loan.

In addition, many of the best payday loan alternatives had features that helped borrowers get on a path to financial security, such as including a savings component to the loan or offering financial education.

NCLC Analysis of Big Data Loan Products

Using these standards, we evaluated seven loan products that are based on big data underwriting, six of which present themselves as payday loan alternatives.[97] Think Finance provides the technology for underwriting five of the seven loan products. Think Finance works with tribal payday lenders to provide two of those products (i.e. Great Plains Lending and Plain Green). ZestFinance provides the underwriting technology to a tribal payday lender for Spotloan and LendUp uses its own big data infrastructure.

According to their own materials, all products charge triple-digit APRs, including fees, for first-time customers. Different products offer different APRs depending on the loan amount, where the borrower obtained the loan, and the repayment schedule (see page 7). The APRs ranged from about 134% to 748%, more typical of payday loans and far more than 36%. The materials did not state how these APRs were calculated; therefore, it is not clear whether the APRs include all fees and could be even higher.

Five of the seven products require weekly or biweekly installment payments. The other two, MySalaryLine and LendUp, require full payment after a set number of days. Borrowers repay MySalaryLine loans on the next pay date while LendUp gives up to thirty days from the loan start date for the borrower's first loan.

All of the lenders except Presta and MySalaryLine require borrowers to provide sensitive banking information (i.e. bank name, routing number, and account number). However, some of the lenders may not use this electronic information in every case.

- **LendUp** automatically deducts the owed amount from the borrower's account that was used to deposit the loan originally.
- **MySalaryLine** works in conjunction with an employer's payroll provider to debit the amount automatically from an employee's next paycheck through payroll direct deposit.
- **Plain Green** and **Great Plains Lending** are also enabled to perform automatic withdrawals from a borrower's bank account on her payday, but it is unclear whether they will deduct an amount if a borrower asks to provide payment through other means.
- **Presta** automatically charges customers' debit or credit cards according to their payment schedule.
- **RISE** will electronically debit a payment from a borrower's checking account unless alternate arrangements are made.
- **Spotloan** will automatically deduct payments from customers' checking accounts, but claims it will also accept checks.

Four out of the seven products did offer financial education resources in the form of online courses. RISE, Plain Green and Great Plains Lending all offer the same "Financial U" online learning center, which is available only to borrowers. In the case of RISE, one of the ThinkFinance products, a $10 reward is offered upon successful completion of the program. Plain Green and Great Plains Lending also offer a reward for program completion, but one must borrow money to obtain information about what the reward is and how to claim it. LendUp provides its own financial education materials available to the public, which consist of informational videos followed by quizzes. Successful completion of credit courses allows LendUp borrowers to accumulate points, which can be used to achieve higher status levels, in line with LendUp's gamification of lending. These status levels claim to open the door to better terms for future loans, including higher loan amounts and installment payments.

Some of the features of these loans are arguably "less bad" than those offered by traditional payday lenders, but these products still fail to meet the requirements to be considered genuine, better alternatives. They still feature three-digit APRs. With the exception of LendUp and MySalaryLine, all products accept installment payments; however, many of them require weekly or biweekly payments rather than monthly ones.

As mentioned previously, all but two of the lenders require borrowers to provide sensitive banking information (i.e. bank name, routing number, and account number). While some may not make use of this information in all cases, the same is true of traditional payday lenders, which typically allow borrowers to repay in person or by other means. The Electronic Funds Transfer Act generally prohibits conditioning an extension of credit on the consumer's repayment of that debt by preauthorized electronic fund transfers.[98] Many payday-type lenders structure their loan products to evade the important protections of this Act while still maintaining a high degree of access to the consumer's account, and we are concerned that the lenders discussed in this report that require bank account information may be following this pattern. It may also violate the FTC's Credit Practices Rule which prohibits creditors from using certain contract provisions that the

FTC found to be unfair to consumers.[99] RISE, Plain Green, Great Plains, and Spotloan do allow consumers to repay their loans by alternatives means, but require consumers to provide sensitive banking information (i.e. bank name, routing number, and account number). A lender could potentially use this information to reach into a bank account and take the funds if the consumer fails to make a payment. The requirement that the borrower provide electronic information could ensure that the lender will be repaid, even if the borrower is unable to afford the loan without neglecting other expenses (like rent or food) or falling into a cycle of debt.

> Loan terms for these seven products appear to be an improvement over their traditional payday lending counterparts only in that some allow installment payments and some allow repayment periods of 90 days or longer. The differences are not enough to consider the products as safe or genuine alternatives to payday loans.

More importantly, it is unclear whether these lenders actually evaluate a borrower's ability to repay, precisely because there is no information available concerning the specific methods big data uses to underwrite loans, nor is there any information available about the default rates for any of these products. The ability to repay a loan must consider more than a credit score or predictive set of algorithms. It must consider the income and assets a consumer has, in addition to the consumer's debts and obligations. Without an understanding of the data points that go into these lenders' underwriting algorithms, it is not possible to determine if the risk of default is being properly evaluated.

In short, loan terms for these seven products appear to be an improvement over their traditional payday lending counterparts only only in that some allow installment payments and some allow repayment periods of 90 days or longer. The differences are not enough to consider the products as safe or genuine alternatives to payday loans.

CONCLUSION AND POLICY RECOMMENDATIONS

With the advances in technology, big data is more and more likely to find its way into lending decisions. As stated above, there is a need for improving affordable access to credit to low-income consumers. However, access by itself is not the ultimate goal; affordable access is the goal.

Although innovation should be encouraged, it should not go unfettered. The good intentions driving new products that aim to broaden access for low-income consumers are laudable—but they are no substitute for strong consumer protections, which remain vitally important. As new financial products emerge—especially those targeted towards low-income consumers and the unbanked or underbanked—the integrity of those products must be examined. As described, the framework is:

1. Are the decisions based upon accurate data?
2. Can the algorithms, when fed with good data, actually predict the creditworthiness of low-income consumers?

3. Does the use of big data in reports used for credit, employment, insurance, and other purposes comply with consumer protection laws?
4. Is there the potential for a discriminatory impact on racial, geographic, or other minority groups?
5. Does the use of big data actually improve the choices for consumers?

Answering these questions has been especially challenging given the secretive and proprietary nature of the products examined. Without voluntary disclosure of the methods that data brokers use to collect and analyze data, there is no way for consumer advocates to answer the first question in our framework. Fortunately, the FTC has taken an interest in the use of big data. Hopefully its analysis will give us answers to some of these questions.

Unfortunately, our analysis concludes that big data does not live up to its big promises. Consumers have the right to be judged based upon accurate and relevant information. But even our small sample found that consumer information housed by data brokers was riddled with errors—and this is just the data they were willing to give us. We suspect that error rates are actually higher. Big data brokers do not provide consumers with a meaningful way to verify the accuracy of their information, nor is there any way that inaccurate information can be disputed.

Big data brokers do not provide consumers with a meaningful way to verify the accuracy of their information, nor is there any way that inaccurate information can be disputed.

Furthermore, the use of big data in the lending arena does not appear to result in more affordable products for low-income consumers. While some loans are marginally better, for the most part, credit products using alternate data are just as expensive as payday loans.

The credit reporting system is far from perfect. It is possible that new technologies could play a role in providing low-income consumers better access to affordable credit. However, the products we reviewed do not meet that promise.

Despite the big promises, a review of the big data underwriting systems and the small consumer loans that use them leads us to believe that big data is a big disappointment. More and more, consumers are leading robust lives online. However, as data about consumers proliferates, so does bad data.

Key Federal Policy Recommendations

- The FTC should continue to study big data brokers and credit scores testing for potential discriminatory impact, compliance with disclosure requirements, accuracy, and the predictiveness of the algorithms.
- The FTC and the CFPB should examine big data brokers for legal compliance with FCRA and ECOA.
- The CFPB should create a mandatory registry for consumer reporting agencies so that consumers can know who has their data.

- The CFPB, in coordination with the FTC, should create regulations based upon the FTC's research that:
  a. Define reasonable procedures for ensuring accuracy when using big data;
  b. Specify a mechanism so that consumers can do a meaningful review of their files including all data points that can be linked to that consumer (not just those that identify the consumer explicitly); and
  c. Define reasonable procedures for disputing the accuracy of information.
- The CFPB should require all of the financial products it regulates to meet Regulation B's requirements for credit scoring models.

ENDNOTES

1. Nate Silver, *The Signal and the Noise*, 13 (2012).
2. Id. at 12.
3. Pew Internet & American Life Project, "Who's Online: Internet User Demographics," www.pewinternet.org/data-trend/internet-use/latest-stats/ (last visited Feb. 14, 2014). The lowest income bracket surveyed was adults with an income of less than $30,000 per year.
4. Dharma Dailey, Amelia Bryne, Alison Powell, Joe Karaganis, and Jaewon Chung, Social Science Research Council, "Broadband Adoption In Low-Income Communities" (Mar. 2010) available at www.ssrc.org/programs/broadband-adoption-in-low-income-communities/.
5. Aaron Smith, Pew Internet & American Life Project, Smartphone Adoption and Usage (July 11, 2011), available at www.pewinternet.org/2011/07/11/smartphone-adoption-and-usage/.
6. Pew Internet & American Life Project, "Mobile Technology Fact Sheet" (Dec. 27, 2013), www.pewinternet.org/fact-sheets/mobile-technology-fact-sheet/ (last visited Feb. 14, 2013).
7. Joanna Brenner, Pew Internet & American Life Project, Pew Internet: "Social Networking" (Dec. 27, 2013), available at www.pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx.
8. James Manyika, McKinsey Global Institute, "Big Data: The Next Frontier for Innovation, Competition, and Productivity" (May 2011), at 1.
9. P.C.Mag.com, "Encyclopedia: Big Data," www.pcmag.com/encyclopedia/term/62849/big-data (last visited Feb. 18., 2014).
10. Kashmir Hill, "How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did," *Forbes*, Feb. 16, 2012, www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did.
11. Acxiom, Things We Are Working On (June 2012).
12. Zestfinance.com, How We Do It, www.zestfinance.com/how-we-do-it.html (last visited Feb. 18, 2014).
13. "Crunching the numbers: Banks know a lot about their customers. That information may be valuable in more ways than one," *Economist*, May 19, 2012, www.economist.com/node/21554743.
14. John Gantz and David Reinsel, IDC, "Extracting Value from Chaos" (June 2011) available at www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf.
15. Jennifer Valentino-DeVries, Jeremy Singer-Vine, & Ashkan Soltani, "Websites Vary Prices, Deals Based on Users' Information: A Tale of Two Prices," *Wall Street Journal*, Dec. 24, 2012, http://on.wsj.com/XenAEy.
16. "What They Know," *Wall Street Journal*, available at http://blogs.wsj.com/wtk/ (last visited Feb. 18, 2014).
17. Christopher Heine, "Rapleaf Agrees to Leave Facebook Alone," *ClickZ*, Nov. 1, 2010, www.clickz.com/clickz/news/1866122/rapleaf-agrees-leave-facebook
18. Kashmir Hill, "Social Media Background Check Company Ensures That Job-Threatening Facebook Photos Are Part Of Your Application," *Forbes*, June 20, 2011, www.forbes.com/sites/kashmirhill/2011/06/20/now-your-embarrassingjob-threatening-facebook-photos-will-haunt-you-for-seven-years.
19. Penny Crosman, "ZestFinance Aims to Fix Underwriting for the Underbanked," *American Banker*, Nov. 19, 2012, www.americanbanker.com/issues/177_223/zestfinance-aims-to-fix-underwriting-for-the-underbanked-1054464-1.html.

20. Zoot Enterprises, Data Aggregation, www.zootweb.com/data_aggregation.html.
21. Ari Levy and Dakin Campbell, *Bloomberg News*, "Startups assess loans without using FICO scores," *San Francisco Chronicle*, Jan. 23, 2012, www.sfgate.com/business/article/Startups-assess-loans-without-using-FICO-scores-2674457.php.
22. GigaOM, Underwriting for the Underbanked Through Data Mining, www.livestream.com/gigaombigdata/video?clipId=pla_b2aac565-50ce-4430-8693-eb2ba0729c5e.
23. Experian, Thin File—Redefined, www.experian.com/consumer-information/thin-file.html (last visited Feb. 18, 2014).
24. Nassim N. Taleb, "Beware the Big Errors of 'Big Data'", *Wired*, Feb. 8, 2013, www.wired.com/opinion/2013/02/big-data-means-big-errors-people/.
25. Kate Rogers, "FTC: 1 in 5 Have Credit Report Errors, Do You?," *Fox Business*, Feb. 12, 2013, available at http://fxn.ws/UaPO6Y.
26. Deborah Pierce & Linda Ackerman, "Privacy Activism, Data Aggregators: A Study of Data Quality and Responsiveness" (May 19, 2005), available at www.csun.edu/~dwm3265/IS312/DataAggregatorsStudy.pdf.
27. Persis S. Yu & Sharon M. Dietrich, Nat'l Consumer Law Ctr, *Broken Records: How Errors by Criminal Background Checking Companies Harm Workers and Businesses* (April 2012), available at www.nclc.org/issues/broken-records.html.
28. Natasha Singer, "You For Sale," *N.Y. Times*, June 16, 2012, www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html?pagewanted=all&_r=0.
29. Acxiom, Data Solutions, www.acxiom.com/data-solutions/#database-services.
30. Natasha Singer, "You For Sale," *N.Y. Times*, June 16, 2012, www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html?pagewanted=all&_r=0.
31. eBureau, eScores: Data Sheet, www.ebureau.com/sites/default/files/file/datasheets/ebureau_escore_datasheet.pdf.
32. eBureau, Income Estimator: Data Sheet, www.ebureau.com/sites/default/files/file/datasheets/ebureau_income_estimator_datasheet.pdf.
33. ID Analytics, White Paper, Obtaining a Unique Perspective on Credit Risk with Identity Information, (Mar. 2009), www.idanalytics.com/assets/whitepaper/IDAnalyticsCreditOpticsWhitePaper.pdf.
34. ID Analytics, Leverage Deep Insight into Consumer Identity Behavior, www.idanalytics.com/technology.
35. Intelius, Products, http://corp.intelius.com/products.
36. Spokeo, About, http://www.spokeo.com/about.
37. Fed. Trade Comm'n, "FTC to Study Data Broker Industry's Collection and Use of Consumer Data," (Dec. 18, 2012), available at http://www.ftc.gov/opa/2012/12/databrokers.shtm. The other five data brokerages included Acxiom, Corelogic, Datalogix, Peekyou, Rapleaf and Recorded Future.
38. Since this study was conducted, Acxiom has created an online portal, www.AboutTheData.com, for consumers to view their information. This new portal has not been analyzed as a part of this report.
39. Acxiom, U.S. Reference Info Report, www.acxiom.com/about-acxiom/corporate-governance/us-reference-info-report/. As of Feb 14, 2014, this page no longer exists. Archives of this page are on file with author.
40. eBureau requires consumers to provide personal information, a copy of a government-issued ID, and a current utility, phone, or credit card bill.

41. Natasha Singer, "Consumer Data, but Not for Consumers," *New York Times* (July 21, 2012), www.nytimes.com/2012/07/22/business/acxiom-consumer-data-often-unavailable-to-consumers.html?pagewanted=all.
42. eBureau, "About Us," www.ebureau.com/about (last visited Jan. 14, 2014).
43. ID Analytics, Adding a New Dimension to Determining Credit Risk, www.idanalytics.com/solutions/enterprise/credit-risk/ (last visited Jan. 14, 2014).
44. eBureau, Income Estimator: Data Sheet, http://www.ebureau.com/sites/default/files/file/datasheets/ebureau_income_estimator_datasheet.pdf.
45. Reg. B, 12 C.F.R. pt. 1002 [pt. 202].
46. 15 U.S.C. §§ 1691/-/1691f.
47. Reg. B, 12 C.F.R. § 1002.2(p)(1) [§ 202.2(p)(1)]. See also Official Staff Commentary to Regulation B § 1002.2(p)-1 [§ 202.2(p)-1].
48. Reg. B, 12 C.F.R. § 1002.2(p)(1)(i) [§ 202.2(p)(1)(i)].
49. *Id*. § 1002.2(p)(1)(ii) [§ 202.2(p)(1)(ii)].
50. *Id*. § 1002.2(p)(1)(iii) [§ 202.2(p)(1)(iii)].
51. *Id*. § 1002.2(p)(1)(iv) [§ 202.2(p)(1)(iv)]. No definition of "periodically" is given in the regulation. See also Official Staff Commentary to Regulation B § 1002.2(p)-2 [§ 202.2(p)-2] (providing guidance on revalidation procedures).
52. See Official Staff Commentary to Regulation B § 1002.2(p)-1 [§ 202.2(p)-1] (describing the difference as relating only to how age is used as a predictive factor).
53. For example, the Office of the Comptroller of Currency (OCC) has required the national banks that it regulates to ensure that the bank's scoring models meet the validation requirements of Regulation B's definition of credit scoring. OCC, *Credit Scoring Models*, OCC Bull. 97-24 (May 20, 1997); OCC, Fair Lending Handbook, Appendix B, Jan. 2010, available at www.occ.gov/publications/publications-by-type/comptrollers-handbook/Fair%20Lending%20Handbook.pdf. The National Credit Union Administration has similarly required that the credit scoring systems used by the credit unions that NCUA regulates meet the criteria of Regulation B's definition of credit scoring. National Credit Union Administration, NCUA Letter to Credit Unions No. 174 (Aug. 1995). See also Fed. Fin. Insts. Examination Council, *Interagency Fair Lending Examination Procedures Guide*, Appendix II—Considering Automated Underwriting and Credit Scoring (Aug. 2009), available at www.ffiec.gov/pdf/fairappx.pdf (inter-agency examination procedures for reviewing a financial institution's credit scoring models).
54. 15 U.S.C. § 1681a(f).
55. 15 U.S.C. § 1681a(d)(1).
56. 15 U.S.C. 1681g(f)(2)(A)(ii). See National Consumer Law Center, *Fair Credit Reporting* § 14.2.3.1 (8th ed. 2013).
57. See Robinson v. TSYS Total Debt Mgmt., Inc., 447 F. Supp. 2d 502 (D. Md. 2006) ("likelihood of collection score" is a consumer report).
58. FTC Staff Summary § 603(d)(1) item 6A. See National Consumer Law Center, *Fair Credit Reporting* § 2.3 (8th ed. 2013).
59. National Consumer Law Center, *Fair Credit Reporting* § 2.3.4 (8th ed. 2013).
60. 15 U.S.C. § 1681a(d)(1).
61. 15 U.S.C. § 1681a(d)(1) (emphasis added).
62. 15 U.S.C. § 1681a(c). See National Consumer Law Center, *Fair Credit Reporting* § 2.3.3 (8th ed. 2013).
63. National Consumer Law Center, *Fair Credit Reporting* § 2.3.3 (8th ed. 2013).

64. FTC, 40 Years Staff Report Accompanying FTC Staff Summary § V(D); FTC Staff Summary § 603(d)(1) item 5A.
65. 15 U.S.C. § 1681e(b).
66. See generally National Consumer Law Center, *Fair Credit Reporting* § 4.2.3 (8th ed. 2013).
67. See generally National Consumer Law Center, *Fair Credit Reporting* § 4.2.3 (8th ed. 2013).
68. Persis S. Yu & Sharon M. Dietrich, Nat'l Consumer Law Ctr, *Broken Records: How Errors by Criminal Background Checking Companies Harm Workers and Businesses* (April 2012), available at www.nclc.org/issues/broken-records.html.
69. 15 U.S.C. § 1681g.
70. 15 U.S.C. § 1681m(a). See National Consumer Law Center, *Fair Credit Reporting* § 3.3.6 (8th ed. 2013).
71. Persis S. Yu & Sharon M. Dietrich, Nat'l Consumer Law Ctr, *Broken Records: How Errors by Criminal Background Checking Companies Harm Workers and Businesses* (April 2012), available at www.nclc.org/issues/broken-records.html.
72. See National Consumer Law Center, *Fair Credit Reporting* § 3.5 (8th ed. 2013).
73. 15 U.S.C. § 1681i(a). See National Consumer Law Center, *Fair Credit Reporting* § 4.5 (8th ed. 2013).
74. See e.g., www.lexisnexis.com/risk/solutions/contact-locate/skip-tracing-accurint-collections.aspx (last visited Dec. 6, 2013). Adams v. LexisNexis Risk & Info. Analytics Group, Inc. 2010 WL 1931135 (D.N.J. May 12, 2010).
75. Consent Decree, United States v. Spokeo, Inc., No. CV12-05001 MMM (C.D. Cal. June 7, 2012) (prohibiting Spokeo from violating the FCRA). See also Robins v. Spokeo, Inc., 2011 WL 1793334 (C.D. Cal. May 11, 2011) (allegations that defendant regularly accepts money in exchange for reports that "contain data and evaluations regarding consumers' economic wealth and creditworthiness" were sufficient to support a plausible inference that the conduct fell within the scope of the FCRA).
76. Spokeo, Terms of Use, http://www.spokeo.com/terms-of-use (last visited Nov. 15, 2013).
77. See National Consumer Law Center, *Fair Credit Reporting* § 2.5.3.2 (8th ed. 2013) .
78. In re Filiquarian Publishing, L.L.C., Choice Level, L.L.C. and Joshua Linsk, Docket No. C-4401 (FTC Apr. 30, 2013) (final decision and order). See National Consumer Law Center, *Fair Credit Reporting* § 2.3.5.1 (8th ed. 2013).
79. Letters from Maneesha Mithal, Fed. Trade Comm'n, to Everify, Inc., InfoPay, Inc., and Intelligator, Inc. (Feb. 7, 2012), available at www.ftc.gov/opa/2012/02/mobileapps.shtm.
80. See National Consumer Law Center, *Fair Credit Reporting* § 2.3.5.1 (8th ed. 2013).
81. Jessica J. González, National Hispanic Media Coalition, "Communications Technology Use in Communities of Color," (March 18, 2013), Slide 6.
82. *Id.*, Slide 11.
83. Susannah Fox, Pew Internet & American Life Project, "51% of U.S. Adults Bank Online" (Aug. 7, 2013), www.pewinternet.org/Reports/2013/Online-banking/Findings.aspx; Michael P. Tremoglie, "Mobile Banking Is More Common Among Blacks and Hispanics," *MainStreet*, Aug. 13, 2013, www.mainstreet.com/article/smart-spending/technology/mobile-banking-more-common-among-blacks-and-hispanics?page=1; Board of Governors of the Federal Reserve Board, Consumers and Mobile Financial Services (March 2012), www.federalreserve.gov/econresdata/mobile-devices/2012-current-use-mobile-banking-payments.htm
84. Susannah Fox, Pew Internet & American Life Project, "51% of U.S. Adults Bank Online" (Aug. 7, 2013), www.pewinternet.org/Reports/2013/Online-banking/Findings.aspx.

85. Chris Hoenig, "Survey: Mobile Banking Most Common Among Blacks, Latinos," *DiversityInc*, Aug. 2013, www.diversityinc.com/news/survey-mobile-banking-most-common-among-blacks-latinos.
86. See e.g., MaxMind, GeoIP/IP Address Location Database, www.maxmind.com/en/geolocation_landing (last visited Dec. 4, 2013).
87. "A Tale of Two Prices," *Wall Street Journal*, Dec. 24, 2012, http://online.wsj.com/news/articles/SB10001424127887323476304578198373572451206 (last visited Dec. 3, 2013).
88. Lori Andrews, "Facebook Is Using You," *New York Times*, Feb. 4, 2012, www.nytimes.com/2012/02/05/opinion/sunday/facebook-is-using-you.html?pagewanted=all.
89. 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.
90. See National Consumer Law Center, *Credit Discrimination*, § 4.3.1 (6th ed. 2013).
91. *Id.* at § 4.3.2.5.
92. Official Staff Commentary to Regulation B, 12 C.F.R. § 202.6(a)-2.
93. A disparate treatment claim could be brought under the Civil Rights Acts of 1866. 42 U.S.C. §§ 1981 and 1982.
94. Consumer Federation of America, "How Payday Loans Work," www.paydayloaninfo.org/facts#1 (last visited Jan. 14, 2014).
95. See e.g., Susanna Montezemolo, Center for Responsible Lending, *Payday Lending Abuses And Predatory Practices* (Sept. 2013), available at www.responsiblelending.org/state-of-lending/reports/10-Payday-Loans.pdf; National Consumer Law Center, *Stopping the Payday Loan Trap* (June 2010), available at http://www.nclc.org/images/pdf/high_cost_small_loans/payday_loans/report-stopping-payday-trap.pdf.
96. National Consumer Law Center, *Stopping the Payday Loan Trap* (June 2010), available at http://www.nclc.org/images/pdf/high_cost_small_loans/payday_loans/report-stopping-payday-trap.pdf.
97. Presta, the only rent-to-own product evaluated, resembles the cost of traditional lease-to-own products. We treat the Presta product as a loan product because it serves the same function as a long for the price of the item being acquired. As a form of credit, an item bought through Presta carries an APR of 202.07% for a product valued at $300.
98. 15 U.S.C. § 1693k; Reg. E, 12 C.F.R. § 1005.10(e)(1) [§ 205.10(e) (1)]; Official Interpretations of Reg. E, 12 C.F.R. pt. 1005 [pt. 205], Supp. I, § 1005.10(e) [§ 205.10(e)].
99. 16 C.F.R. Part 444.

NCLC®

NATIONAL CONSUMER LAW CENTER®

Advancing Fairness in the Marketplace for All

Boston Headquarters:
7 Winthrop Square
Boston, MA 02110-1245
Phone: 617/542-8010
Fax: 617/542-8028
www.nclc.org

Washington Office:
1001 Connecticut Ave, NW
Suite 510
Washington, DC, 20036
Phone: 202/452-6252
Fax: 202/463-9462