BIG DATA

A BIG DISAPPOINTMENT FOR SCORING CONSUMER CREDIT RISK

"The problem with quotes on the Internet is that it is hard to verify their authenticity."
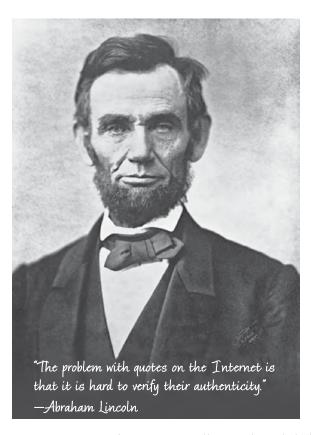—*Abraham Lincoln*

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.