FinRegLab

The Use of Cash-Flow Data in Underwriting Credit

Empirical Research Findings

JULY 2019

About FinRegLab

FinRegLab is a non-profit research organization that was founded on the premise that independent, rigorous research is a primary ingredient in helping develop market norms and policy solutions that enable responsible innovation in financial services.

Acknowledgments

FinRegLab was established with generous support from Flourish. Flourish also supported FinRegLab's evaluation of cash-flow data in consumer and small business credit underwriting and this summary report, *The Use of Cash-Flow Data in Underwriting Credit: Empirical Research Findings*.

Flourish is an evergreen, early-stage venture fund investing globally in entrepreneurs whose innovations help people achieve financial health and prosperity. Spun out of Omidyar Network in 2019 with an existing portfolio of $200 million, Flourish received a new commitment of $300 million from Pam and Pierre Omidyar, the founder of eBay. Flourish invests in a number of themes in fintech, insurtech, regtech and other technologies, as well as supports nonprofit organizations, that empower people and help foster a fair, more inclusive economy. Flourish is managed by a global team with offices in Silicon Valley, Washington DC, London, and India.

For more information visit www.flourishventures.com or join our community through Twitter, LinkedIn, Instagram, and Medium.

Additional support for this report was provided by the Milken Institute, which enabled FinRegLab to evaluate cash-flow data in small business credit underwriting.

The Milken Institute is a nonprofit, nonpartisan think tank that helps people build meaningful lives, in which they can experience health and well-being, pursue effective education and gainful employment, and access the resources required to create ever-expanding opportunities for themselves and their broader communities.

We are particularly grateful to Marsha Courchane and Arthur Baines, Charles River Associates, for their counsel in developing the research design, data analysis, and reporting of the research results. We are also grateful to Thomas P. Brown, Paul Hastings, for legal assistance throughout the course of the research process.

We would also like to acknowledge the FinRegLab team who worked on the research and the report(s). They include: Jennifer Chasseur, Kelly Thompson Cochran, Sarada Dhulipala, Aaron Milner, Stephen Stolzenberg, and Jared Yee.

CONTENTS

1. EXECUTIVE SUMMARY

Access to affordable credit can play a major role in improving the financial health of both consumers and small businesses. From smoothing short-term gaps between inflows and outflows to expanding long-term financial capacity through investments in housing, education, transportation, or business expansion, credit access can be a critical gateway to improved financial stability and well-being.

However, millions of consumers and small businesses struggle to achieve consistent access to affordable credit in today's markets. This is due in part to gaps and weaknesses in traditional credit reporting systems, which many lenders rely upon heavily for information to assess credit applications. For example, an estimated 45 million to 60 million consumers lack sufficient history to generate reliable credit scores that can be used to predict their repayment risk.[1] Concerns about the predictiveness of information available to underwrite small businesses also contributed to many traditional lenders' decisions to reduce their activities in that market in the wake of the 2008 financial crisis.[2]

To fill these gaps, both traditional incumbents and new entrants are experimenting with various sources of "alternative" or "non-traditional" data. One of the most promising of these alternatives is cash-flow data — such as records of transactions in and out of consumers' deposit and card accounts and feeds from small businesses' accounting software — because it provides a relatively detailed and comprehensive picture of how applicants manage their finances on an ongoing basis. Yet while recent technological and market developments are making it easier for lenders to access cash-flow information electronically, the adoption of underwriting models that rely on detailed analyses of such information is uneven in the United States. For instance, while there is increasing interest in such models in small business credit markets, adoption in consumer lending appears to be slower particularly among banks and credit unions, despite the fact that they have direct access to such information for their existing customers.

In light of the potential for cash-flow based underwriting to improve risk prediction and access to credit in both consumer and small business markets, FinRegLab set out to conduct empirical and policy analyses to assess the benefits and risks of using such data in credit underwriting and the

1 Consumer Financial Protection Bureau Office of Research, Data Point: Credit Invisibles 4-6 (2015) (hereinafter CFPB Credit Invisibles); Peter Carroll & Saba Rehmani, Point of View: Alternative Data and the Unbanked 5, Oliver Wyman (2017)

2 Karen G. Mills, Fintech, Small Business and the American Dream: How Technology Is Transforming Lending and Shaping a New Era of Small Business Opportunity Chapters 4, 6 [eBook] (2019); Peter Carroll & Ben Hoffman, Financing Small Businesses: How 'New-Form Lending' Will Reshape Banks' Small Business Strategies 3, Oliver Wyman (2013)

hurdles to its wider adoption. FinRegLab is a non-profit research organization that was founded in 2018 based on the premise that independent, rigorous research is a primary ingredient in helping develop market norms and policy solutions that enable responsible innovation in financial services. This report, along with two companion documents, is our first effort to provide such research and begin a conversation on themes that we expect to recur in our subsequent work.

In particular, this Empirical Research Findings report provides a detailed summary of our applied research based on data from six non-bank financial services providers—Accion, Brigit, Kabbage, LendUp, Oportun, and Petal—that have begun using cash-flow variables and scores in an effort to increase the provision of credit to consumers and small businesses who may have difficulty obtaining loans from traditional sources. FinRegLab retained Charles River Associates to help us design and conduct an independent analysis of the predictiveness of the participants' cash-flow variables and scores based on actual loan performance. We also compared the predictiveness of the cash-flow metrics to traditional scores and variables, as well as to combined models using both types of information. Where data permitted, we also analyzed the extent to which the research participants are providing credit to traditionally underserved populations and whether the use of the cash-flow variables and scores introduces fair lending risk for credit eligibility determinations.

As discussed in more detail below, our analysis validates that varying types of cash-flow data are being used to underwrite credit for a range of unsecured consumer and small business credit products across a broad set of U.S. geographies. More specifically:

- **Predictiveness:** For the participants for which loan-level data was available, we found compelling evidence that indicates that the cash-flow variables and scores tested were predictive of credit risk and loan performance across the heterogenous set of providers, populations, and products studied. Standing alone, the cash-flow metrics generally performed as well as traditional credit scores, which suggests that cash-flow variables and scores can provide meaningful predictive power among populations and products similar to those studied where traditional credit history is not available or reliable. Moreover, our analysis indicates that the cash-flow data and traditional credit data provided different insights into credit risk, such that the cash-flow data frequently improved the ability to predict credit risk among borrowers that are scored by traditional systems as presenting similar risks of default. These results occurred across traditional credit score bands.
- **Inclusion:** We found evidence that the study participants are serving borrowers who may have historically faced constraints on their ability to access credit, although data limitations did not permit a consistent quantitative analysis to be applied across all participants. We used a variety of benchmarks depending on data availability, including the percentage of borrowers with low or no traditional credit scores, borrower income levels, and residence in zip codes in which racial minorities exceed 50 percent or 80 percent of the total population.
- **Fair lending effects:** Finally, where data was available for analysis, we found that the degree to which the cash-flow data was predictive of credit risk appeared to be relatively consistent across borrowers who likely belong to different demographic groups. Rather than acting as proxies for race and ethnicity or gender, the cash-flow variables and scores appeared to provide independent predictive value across all groups. Moreover, when compared to traditional credit scores and attributes, the cash-flow based metrics appear to predict creditworthiness within the subpopulations at least as well as the traditional metrics, and better in selected cases. These results suggest that cash-flow variables and scores do not create a disparate impact among protected populations.

This report is the only publicly available independent evaluation of cash-flow data of which we are aware. Although some of the sample sizes were relatively modest, the fact that we obtained relatively consistent, statistically significant results across a range of participants, products, and borrower populations is notable. Given that cash-flow data is increasingly available in electronic form to both bank and nonbank lenders, this suggests that further attention is warranted.

The companion reports, which will be released later in summer 2019, provide broader market context and policy analysis for these research results. The Small Business Spotlight report provides a broader picture of cash-flow based underwriting in the small business market and an overview of policy issues that may be particularly important in determining the pace of adoption going forward. The Market Context and Policy Analysis report provides deeper policy analyses of the current state of cash-flow based underwriting in the United States across both consumer and small business markets, challenges and risks in the emerging markets, and options for developing and extending beneficial practices. It focuses on market, legal, and policy issues both in credit underwriting and in the underlying transfers of cash-flow data between companies. Both of these reports build off three working groups that FinRegLab convened to solicit insight and opinion from more than 80 representatives of fintech companies, banks, data aggregators, advocacy organizations, and research institutions, as well as individual stakeholder interviews.

Collectively across the three reports, we conclude that using cash-flow data in credit underwriting holds substantial promise for improving credit risk prediction, expanding access to credit, and spurring market innovation and competition. While the scope of our research and data do not permit us to answer all relevant questions, the reports suggest that stakeholders should invest more resources into reducing the technological, competitive, and compliance challenges that are slowing adoption of beneficial practices and mitigation of risks in today's markets. With thoughtful development, cash-flow based underwriting has the potential to become a win-win for borrowers and financial services providers alike.

2. BACKGROUND

In recent decades, underwriting processes have become increasingly automated across both consumer and small business credit markets. Automated systems can potentially cut costs, increase the consistency of treatment, and improve the prediction of credit risk across different populations. However, they increase lenders' dependence on standardized data and can create fair lending concerns if not carefully structured.

2.1 Credit underwriting and risk prediction

Underwriting credit is a complex process that typically includes consideration of a wide variety of factors that are designed to assess both whether a particular applicant has the *financial capacity* to repay the loan and the *willingness* to do so. These concepts are often described as ability and propensity to repay. Historically, such assessments were made by individual loan officers and underwriters based on both objective information and subjective assessments of the applicants' financial situation, habits, and character. Such underwriting systems are often called judgmental or manual systems. But over the last several decades, lenders have increasingly adopted automated underwriting models that use statistical analyses of financial data to evaluate both applicants' ability and propensity to repay for purposes of determining whether to offer credit and on what terms (e.g., interest rates, loan amounts, etc.).[3]

Automated underwriting models have traditionally relied in large part on data that is provided in credit reports on individual consumer or small business applicants. In the consumer market, the most widely used of such reports are produced by three companies—Equifax, Experian, and TransUnion—which are called "nationwide consumer reporting agencies" (NCRAs) because of their size and scope.[4] The NCRAs' reports are made up largely of information about how individuals are repaying or have repaid previous loans and other major obligations, as well as information from

3 Board of Governors of the Federal Reserve System, Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit O-4, 3, 10-11 (2007) (hereinafter, FRB Credit Scoring Report). Small business lenders' transition to automated underwriting has been slower and more variable than in consumer underwriting, in part because of challenges in obtaining standardized information that is sufficiently predictive of credit risk across a broad range of small business types. Mills, Chapters 4, 6.

4 Consumer Financial Protection Bureau, Key Dimensions and Processes in the U.S. Credit Reporting System: A Review of How the Nation's Largest Credit Bureaus Manage Consumer Data 3 (2012) (noting that the three NCRAs maintain files on more than 200 million U.S. adults concerning more than 1.3 billion consumer credit accounts or other "trade lines") (hereinafter, CFPB Key Dimensions); FRB Credit Scoring Report at 13-16.

BOX 2.1.1 CREDIT SCORES IN BUSINESS LENDING

Specialized credit scoring systems started to develop substantially later in the small business lending market than in consumer lending. At the time of the financial crisis in 2008, many traditional lenders were still relying primarily on the personal scores of business owners rather than commercial credit scores. Concerns about the predictiveness of available data and heavy losses prompted some large lenders to restrict their activities after the crisis, for instance by increasing their minimum loan amounts.

In the decade since the crisis, the commercial credit reporting industry has moved toward more standardized records of payments to vendors, equipment purchases, and creditors, similar to the kind of "trade line" information that is common in consumer reports. Nevertheless, the algorithms for business credit scores are not as standardized as for consumer scores, and business scores are more frequently available for established small businesses than for startups.

A 2018 survey of nearly 5000 small businesses with at least one employee found that 86 percent of companies relied in whole or in part on their owners' personal scores to obtain financing, with only 14 percent relying solely on business scores. Reliance on personal scores is even stronger among sole proprietorships and other firms without employees.

Sources: Allen N. Berger et al., The Surprising Use of Credit Scoring in Small Business Lending by Community Banks and the Attendant Effects on Credit Availability and Risk, Federal Reserve Bank of Atlanta Working Paper 2009-9, at 1-4 (March 2009); Mills, Chapters 4, 6; Carroll & Hoffman at 3; Claire Tsosie & Steve Nicastro, Business Credit Score 101, nerdwallet (Oct. 6, 2017); Federal Reserve Banks, 2018 Small Business Credit Survey Report on Employer Firms 9 (2019); Federal Reserve Banks, 2017 Small Business Credit Survey Report on Nonemployer Firms 9 (2018); Federal Reserve Bank of New York, 2016 Small Business Credit Survey: Report on Startup Firms 8 (2017).

public records sources about bankruptcies and liens.[5] So-called specialty CRAs produce reports that may focus on repayment of specific types of expenses, such as rent or very short-term loans that are not typically reported to the NCRAs.[6] The commercial credit reporting market also includes a mix of companies, including Dun & Bradstreet, Equifax, and Experian, as well as various niche bureaus.[7]

Credit report information can be helpful to assess both ability and propensity to repay, since it may show both current obligations and past repayment history. A number of companies have also developed "credit scoring" models that use historical data from credit reports or other sources to group applicants into bands reflecting their predicted likelihood of default. Over the last several decades, so-called generic or third-party credit scores relying solely on data from NCRAs have become widely used in consumer lending; the most well-known of these scores are provided by the Fair Isaac Corporation (FICO) and a joint venture of the NCRAs called VantageScore.[8] Small business underwriting often relies on the personal scores of business owners in addition to commercial scores for the businesses, where available.[9]

Lenders may factor third-party scores into their own underwriting processes in a variety of ways, for instance by establishing minimum score thresholds under which credit will not be extended

5 CFPB Key Dimensions at 8-10.

6 Consumer Financial Protection Bureau, List of Consumer Reporting Companies (2019) (publishing annual list of consumer reporting agencies, including ten categories of specialty companies). In addition to various uses in credit markets, credit reports are also used frequently in eligibility determinations for employment, certain decisions relating to insurance, rental housing, and (along with deposit account history) checking accounts. CFPB Key Dimensions at 5.

7 Gail Gardner, What Are the Credit Reporting Agencies for Businesses?, Small Business Trends (Jan. 4, 2019), available at smallbiztrends.com/2019/01/business-credit-reporting-agencies.html.

8 FRB Credit Scoring Report at O-4, 8-9, 22-24; CFPB Key Dimensions at 10. Such models generally group consumers based on estimates of the likelihood that they will become seriously delinquent on any of their credit accounts in the near future (typically 18 to 24 months). One method of developing generic models is to take snapshots of the credit records for a representative sample of consumers at two points in time separated by about 18 to 24 months. The predictive characteristics are calculated from the earlier sample, and compared to the records in the second snapshot that reflect which borrowers have become seriously delinquent on any credit accounts in the intervening period. Model developers then perform statistical analyses to determine which characteristics are most predictive of delinquency and to assign weights to reflect their relative importance. CFPB Key Dimensions at 10; FRB Credit Scoring Report at 8-9, 22-27.

9 Federal Reserve Banks, 2018 Small Business Credit Survey Report on Employer Firms 9 (2019); Federal Reserve Banks, 2017 Small Business Credit Survey Report on Nonemployer Firms 9 (2018).

BOX 2.1.2 TRADITIONAL REPORTING SYSTEMS GAPS

In the consumer reporting system, gaps in coverage can occur for several reasons. First, because reporting is voluntary, variations in industry reporting patterns and individual companies' decisions about whether and what types of information to report can affect the ability of their customers to be scored and assessed by other lenders and credit report users downstream. For example, mortgage payments are far more likely to be reported to the NCRAs than rental payments.

Challenges in merging monthly updates from thousands of disparate information sources have also produced substantial concerns about accuracy and consistency across individual consumers' credit files. A 2012 study by the Federal Trade Commission reported that 26 percent of participating consumers found potentially material errors in their NCRA credit files, 13 percent obtained corrections that improved their credit scores, and 5 percent obtained corrections that were so large that they changed credit risk tiers.

Overall, an estimated 45 to 60 million American adults cannot be scored using traditional generic models because they either have no credit files with NCRAs or their files are too limited to produce reliable scores. These "no file" and "thin file" consumers face an irresolvable conflict: they often need a score to qualify for loans and obtain better pricing on credit, and yet to generate a score they need to have borrowed before. Consumers who have stabilized their finances after a previous adverse event often face similar problems in that it is difficult to access credit without already having a positive credit history.

Small businesses owners are often vulnerable to reporting gaps, inaccuracies, and other weaknesses in both commercial and consumer credit information systems, given that lenders frequently use both types of reports and scores to make underwriting decisions. See Box 2.1.1.

Sources: CFPB Key Dimensions 3, 11-12, 21-26; Federal Trade Commission, Report to Congress under Section 319 of the Fair and Accurate Credit Transactions Act of 2003 at i to ii, 35-57 (2012); CFPB Credit Invisibles 4-6; Carroll & Rehmani at 5.

and/or by using them as a variable in more tailored proprietary underwriting algorithms. Because third-party scores facilitate consistent comparisons, they may also be used to monitor portfolios, expedite securitization, and provide investment benchmarks even when they are not used in the underwriting process itself, as well as to facilitate certain types of credit marketing.[10]

Yet while reliance on traditional credit report data and scoring models has been credited particularly in consumer credit markets with improving the consistency of credit evaluation, reducing both up-front underwriting costs and back-end losses, and increasing access to credit,[11] these sources cannot provide a complete assessment of applicants' finances. Traditional credit reports only reflect applicants' payment history on certain obligations—not their incomes, balance sheets, or even a complete picture of all recurring expenses. And because of various coverage gaps and accuracy problems with the data, millions of consumers and small businesses do not have sufficient credit history to generate reliable traditional scores relative to the general population. For these reasons, lenders have historically collected information from other sources, for instance by inquiring into applicants' income and computing metrics such as debt-to-income ratios. But gathering, verifying, and analyzing a detailed picture of applicants' full financial situations can take substantial time and labor, and lenders must balance these costs against competitive pressures to process and approve credit applications quickly.

Thus, recent market and technological advances that make it easier for lenders to gain electronic access to transaction account records and other sources of detailed cash-flow data are potentially

10 FRB Credit Scoring Report at 3, 8-9, 29-32.

11 Credit Scoring Report at O-4 to O-6, 12-13, 39-49; Allen N. Berger & W. Scott Frame, Small Business Credit Scoring and Credit Availability, 47 J. of Small Business Management 5 (2007).

BOX 2.1.3 NEW ENTRANTS AND ALTERNATIVE/NON-TRADITIONAL DATA

Over the past decade, a large number of technology-based firms have entered various markets for financial services and products as both competitors and service providers to banks and other traditional incumbents. These new entrants rely heavily on data and financial technology (fintech) to develop new products and services and to create new methods of customer acquisition, internal operations, and service delivery.

In the credit space, many of these fintech companies started are called marketplace lenders or platforms. They tend to operate almost entirely online, to rely on heavily automated underwriting models, and to sell loans individually or in pieces directly to investors rather than pooling entire portfolios of loans together for securitization. Some fintech companies originate and hold loans directly, while others operate as servicers or partners to banks and other traditional lenders.

Companies that provide payment processing services, e-commerce platforms, and accounting software to small businesses have also begun providing credit options to their customers.

Both new entrants and incumbents are exploring so-called alternative or non-traditional data for purposes of credit underwriting. Those terms do not have precise definitions, but are often use to refer to types of electronic data that are not typically reflected in traditional credit reports or collected in lender applications (such as annual income). For example, those terms are sometimes used to describe cash-flow data, payment history information from landlords and utility companies that have historically not reported extensively to NCRAs, on-line footprint and e-commerce information, and items such as a person's education or employment.

NCRAs and traditional third-party scoring companies are also focusing on alternative data generally and cash-flow data in particular. Experian has launched a product called ExperianBoost that augments consumers' traditional credit files where consumers give permission to access their transaction account data to obtain payments history for utility and telecommunications information. FICO, Experian, and data aggregator Finicity have also announced a joint venture called "UltraFICO" that will create adjusted credit scores where consumers authorize accessing their account data to analyze factors such as the length of time that accounts have been open, recency and frequency of bank transactions, evidence of consistent cash on hand, and history of positive account balances.

Sources: Congressional Research Service, Marketplace Lending: Fintech in Consumer and Small-Business Lending 3-4 (2018); U.S. Department of the Treasury, A Financial System that Creates Economic Opportunities: Nonbank Financials, Fintech, and Innovation 4-6 (2018); Peter Rudegeair, A $150,000 Small Business Loan—From an App, (Dec. 28, 2018); Experian, Alternative Credit Data 5 (2018); TransUnion, The State of Alternative Data 3 (2015); Susan Henson, Blog, Introducing Experian Boost, a New Way to Instantly Improve Your Credit Score (Dec. 18, 2018, updated April 8, 2019); AnnaMaria Andriotis, Why Your FICO Score Could Get a Boost in 2019, Wall St. J. (Oct. 21, 2018).

transformational for both lenders and applicants alike.[12] Because cash-flow data includes both inflows and outflows, it can provide more detailed and holistic information about how consumers and small businesses manage their finances on an ongoing basis than can be obtained from traditional credit reports. Such data also can provide greater sensitivity and timeliness in detecting changes in an applicant's financial position, particularly for small businesses. And because more U.S. households maintain transaction accounts with banks or prepaid providers than have credit products that are likely to be reflected in reports from national credit bureaus,[13] cash-flow data may provide an important source of information for underwriting applicants who fall into gaps in the traditional credit reporting systems.

12 Over the last twenty years, electronic transaction account data has become much more widely available as banks and other account providers have implemented electronic platforms that permit customers to download their statements and conduct transactions online. In addition, technology intermediaries called "data aggregators" have emerged to facilitate transfers of such data between financial services providers at the direction of the consumers or businesses who own the accounts. See generally U.S. Department of the Treasury, A Financial System that Creates Economic Opportunities: Nonbank Financials, Fintech, and Innovation 22-38 (2018).

13 Federal Deposit Insurance Corporation, 2017 FDIC National Survey of Unbanked and Underbanked Households 1, 7, 9-11, 12, table ES.5, 34-38, 48-58 (2018). The FDIC's most recent survey indicates that 94 percent of American households have at least one checking or savings account, and an additional 2 percent have one or more prepaid accounts. In contrast, about 80 percent of households have one or more credit products from what the FDIC describes as "mainstream" bank and non-bank lenders that are likely to report to credit bureaus. *Id.*

A group of fintech companies and other non-bank lenders has begun experimenting with cash-flow based underwriting for unsecured credit products in both consumer and small business credit markets. Some NCRAs and traditional credit scoring companies are also augmenting traditional reports and developing specialized consumer credit scores using cash-flow sources. Some traditional banks have also formed partnerships with fintechs or launched internal initiatives to increase use of electronic cash-flow data in small business underwriting, but appear to be moving more slowly with regard to its use in consumer credit markets. To date, little independent research has been made publicly available in either consumer or small business markets to assess the extent to which such data sources can efficiently and effectively model credit risk or expand access to populations whose information is not fully and accurately reflected in traditional credit reporting systems.[14] In the absence of such information, it remains unclear whether and how quickly more U.S. lenders will adopt cash-flow based underwriting, particularly in the consumer context.

2.2 Fair lending analysis

Beyond basic predictiveness, a second critical consideration in developing or modifying credit underwriting models is potential fair lending risk. The Equal Credit Opportunity Act (ECOA) prohibits discrimination in "any aspect of a credit transaction" for both consumer and commercial credit on the basis of race, color, national origin, religion, sex, marital status, age, or certain other protected characteristics.[15] ECOA has two principal theories of liability. The first is "disparate treatment," in which creditors treat applicants differently based on protected characteristics. The second is "disparate impact," in which use of facially neutral practices has a disproportionately negative effect on members of a protected class, unless those practices are meeting a legitimate business need that cannot reasonably be achieved by less impactful means.[16]

Many observers note that one of the advantages of the shift from manual and judgmental underwriting toward automated systems is that such methodologies tend to decrease the risk of disparate treatment. Such underwriting models are generally prohibited from factoring in protected characteristics,[17] and because they apply algorithms to standardized credit information, a given set of inputs produces the same outputs each time. This promotes consistent treatment even when dealing with a large number of variables that may have complex relationships with each other. Thus,

14 Some research has focused on the general question of whether fintech companies are in fact increasing access to credit and/or lowering prices for underserved populations, but has not focused specifically on the use of specific types of data to predict credit risk. See, e.g., Marco Di Maggio & Vincent W. Yao, Fintech Borrowers: Lax-Screening or Cream Skimming (updated February 2019); Julapa Jagtiani & Catharine Lemieux, The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the LendingClub Consumer Platform, Federal Reserve Bank of Philadelphia Working Paper 18-15 (updated January 2019); Julapa Jagtiani & Catharine Lemieux, Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks? Journal of Economics & Business (November-December 2018). In addition, Experian and FICO have released some statistics based on early analysis of the impact of their consumer cash-flow based initiatives based on sample populations. See Henson; Andriotis.

15 15 U.S.C. § 1691(a). Additional protected characteristics include receipt of public assistance and exercise of certain legal rights under federal consumer financial laws. *Id*. The Fair Housing Act also prohibits discrimination with regard to credit transactions relating to housing on the grounds of race, color, national origin, religion, sex, family status, and handicap. 42 U.S.C § 3605.

16 12 C.F.R. §§ 1002.4(a), 1002.6(a), 1002.6(b)(1); *id*. Supp. I, cmt. 4(a)-1, 6(a)-2. The Consumer Financial Protection Bureau stated in May 2018 that it was reexamining ECOA requirements concerning the disparate impact doctrine in light of recent Supreme Court case law and Congressional disapproval of a prior Bureau bulletin concerning indirect auto lender compliance with ECOA and its implementing regulations. Consumer Financial Protection Bureau, Statement of the Bureau of Consumer Financial Protection on Enactment of S.J. Res. 57 (May 21, 2018).

17 12 C.F.R § 1002.6(b)(1). Age may be considered in certain narrow circumstances. *Id*. § 1002.6(b)(2).

BOX 2.2.1 RESEARCH ON TRADITIONAL CREDIT SCORING

The most comprehensive publicly available fair lending analysis of traditional credit scoring was published in 2007 by economists at the Federal Reserve Board acting pursuant to a mandate from Congress, with further analysis published in 2012. Because credit scoring models are proprietary, the economists had to construct their own model using criteria that are reflected on traditional consumer reports by the NCRAs and using general industry practice to the extent possible. They then applied that model to a nationally representative sample database of 300,000 consumer records that incorporated demographic information from the Social Security Administration.

The report started by looking at differences among demographic groups with regard to average credit scores that were available from TransUnion as well as generated by the Board's model, and at differences in performance outcomes for different demographic groups relative to what the various scores predicted. The report found substantial differences in the median scores of African-Americans and Hispanics relative to whites and Asians. Many of these differences were reduced to the extent that the study authors were able to factor in a census-tract-based estimate of income, but they lacked the data to account fully for differences in such factors as wealth, employment, and education.

Turning to a more sophisticated multivariate analysis of the Board's own model, the study found that it was predictive of credit risk for the population as a whole and for all major demographic groups. When demographic status was controlled for, the model maintained predictiveness but showed some shifts with regard to factors focusing on the length of credit history. Demographically neutral models caused the scores of younger individuals and recent immigrants to increase slightly and the scores of older individuals to decrease slightly. The study concluded that the traditional characteristics used do not serve as proxies for race, ethnicity, or gender, and that their impacts with regard to length of credit history were outweighed by the substantial independent predictive power of those variables.

Sources: FRB Credit Scoring Report; Robert B. Avery, Kenneth P. Brevoort, & Glenn Canner, Does Credit Scoring Produce a Disparate Impact? 40 Real Estate Economics 965 (2012).

automated underwriting generally decreases the risk of inconsistency and personal bias that are inherent in subjective assessments by individuals.[18]

However, automated systems can still pose concerns about fair lending—and fairness in a broader sense—in a number of different ways. For example, if algorithms are developed based on a database that is made up primarily of one type of borrower, they may not work well in predicting the default risk for other types of borrowers. Monitoring model performance over time is also important, since changes in borrower behavior, economic conditions, or lender policies can cause models to lose predictiveness with regard to particular groups or overall.[19]

More broadly, there is a concern that relying on databases that reflect the past results of discrimination to develop predictive models may tend to perpetuate its effects. In the credit context, for example, particularly in light of historical discrimination in employment, education, housing, and lending, advocates have raised concerns about the risk that use of traditional reports and scoring systems may perpetuate previous inequities. Studies frequently find large differences in traditional

18 FRB Credit Scoring Report at O-5, 11, 36-37, 52; CFPB Key Dimensions at 11. For studies finding disparities in treatment between testers from different demographic groups posing as loan applicants, see, e.g., Sterling A. Bone et al., Shaping Small Business Lending Policy Through Matched-Pair Mystery Shopping, 38 J. of Public Policy & Marketing 391 (2019); U.S. Department of Housing & Urban Development Office of Policy Development & Research, All Other Things Being Equal: A Paired Testing Study of Mortgage Lending Institutions (2002). One recent study suggests that unexplained pricing differentials between demographic groups in the mortgage market have dropped substantially from 2009 to 2015, during a period of increasing reliance on automated underwriting models and heightened on-line competition. In addition, the study finds that mortgage lenders that rely heavily on online applications and automated underwriting do not have unexplained differentials in accept/reject decisions and have smaller unexplained differentials in pricing among demographic groups relative to lenders who are more reliant on face-to-face channels and may use less automated underwriting systems. Robert Bartlett et al., Consumer-Lending Discrimination in the FinTech Era, National Bureau of Economic Research Working Paper No. 25943, at 1-2, 15-16, 32 (updated June 2019).

19 Carol A. Evans, Keeping Fintech Fair: Thinking About Fair Lending and UDAP Risks, Consumer Compliance Outlook 4-9 (2nd Issue 2017); Federal Trade Commission, Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues 27-32 (2016); Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Cal. L. Rev. 671 (2016).

BOX 2.2.2 DISPARATE IMPACT ANALYSIS

Litigation and enforcement actions involving disparate impact claims against lenders generally follow a three-step process that has been developed in the employment discrimination context:

- At the first step, a plaintiff must make an initial showing that the particular practice causes a disproportionate adverse effect on protected groups.
- If that showing is made, the burden shifts to the creditor to show that the practice furthers a legitimate business need.
- In the third stage, the burden shifts to the plaintiff to demonstrate whether the legitimate business need can reasonably be achieved by using an alternative practice that would have less adverse impact on protected classes.

Practitioners are still debating the impact of a 2015 Supreme Court decision applying disparate impact analysis under the Fair Housing Act with regard to what showings must be made at each stage. Case law and regulatory guidance do not provide a precise definition of what constitutes a "legitimate business need," for example, although in the credit underwriting context the analysis often focuses on whether there is a "demonstrable relationship" between variables or models and predicting individuals' credit risk. For example, some banking agency guidance on credit scoring models focuses on whether the variable is statistically related to loan performance and has an understandable relationship to creditworthiness.

Statistical tests can be important at each stage of litigation or enforcement, and more generally when lenders set out to evaluate their degree of fair lending compliance risk with regard to adopting or changing their underwriting models. This evaluation process often starts with basic descriptive tests to determine whether there are correlations between demographic status and particular outcomes, variables, or scores. Where particular variables are correlated both with credit performance and with demographic characteristics, analysts may use various techniques to control for the influence of demographic characteristics in order to evaluate the extent to which the variables lose predictive power. For example, they may calculate the predictiveness of a credit model as applied to each demographic group separately to determine whether there are differences that would negatively impact particular protected groups.

Where adverse effects are detected, statistical analyses may also be used to compare the extent of the negative effect to the extent to which particular variables have independent predictive value. Statistical analyses may also be used to determine whether alternative variables or models would have less adverse impact without materially degrading predictive value.

Sources: 12 C.F.R. 1002, supp. I, § 1002.6(a)-2; *Texas Dep't of Housing & Community Affairs v. Inclusive Communities Project, Inc.*, 135 S. Ct. 2507 (2015); OCC, Examination Guidance on Credit Scoring Models, Office of the Comptroller of the Currency Bull. 97-24, app. at 11 (May 20, 1997); David Skanderson & Dubravka Ritter, Fair Lending Analysis of Credit Cards, Federal Reserve Bank of Philadelphia Payment Cards Center Discussion Paper 34-40 (August 2014).

credit scores between different demographic groups, but due to data limitations they generally cannot control fully for the fact that income, assets, and wealth also tend to vary between the study populations.[20] Concerns have also been raised that racial minorities' payment histories may be negatively affected to the extent that they may lack geographic access to banks and are targeted by lenders who offer credit products with higher prices and riskier structures. Differentials in traditional consumer credit scores have remained a continuing concern for advocates even after the Federal Reserve Board performed a large national study of the issue as directed by Congress in 2007.[21]

In light of this context, model validation and governance protocols generally and disparate impact analysis in particular can be an important check on the fairness of credit scoring and other

20 FRB Credit Scoring Report at S-4 to S-6, O-12 to O-24. For recent studies analyzing rare data sources with both income and credit score information, but not racial demographics, see Rachael Beer et al., Are Income and Credit Scores Highly Correlated?, FEDS Notes (Aug. 13, 2018); Stephania Albanesi et al., Credit Growth and the Financial Crisis: A New Narrative, National Bureau of Economic Research Working Paper No. 23740 (August 2017).

21 FRB Credit Scoring Report at S-4 to S-6, O-12 to O-24 (summarizing analysis as of 2007); Robert B. Avery, Kenneth P. Brevoort & Glenn B. Canner, Does Credit Scoring Produce a Disparate Impact? 40 Real Estate Economics S65 (2012) (further analysis); Lisa Rice & Deidre Swesnik, Discriminatory Effects of Credit Scoring on Communities of Color, 46 Suffolk L. Rev. 935 (2013); National Consumer Law Center, Past Imperfect: How Credit Scores and Other Analytics "Bake In" and Perpetuate Past Discrimination (2016).

underwriting algorithms.[22] The legal inquiry for disparate impact is structured as a multi-stage analysis, which generally involves several types of statistical tests as well as consideration of the broader facts and circumstances to assess such questions as the extent to which an underwriting model creates differential effects among demographic groups, the extent to which models or individual variables provide independent value in predicting credit risk, and the availability of less burdensome alternatives. One further complication is that federal law generally prohibits lenders from collecting demographic information on applicants and borrowers for most types of credit.[23] As a result, disparate impact analyses often can be conducted only by first applying proxy methodologies to estimate the likelihood that a particular borrower belongs to a particular demographic group based on one or more factors such as name and geography.[24] This further adds to the complexity and uncertainty of the analysis.

2.3 FinRegLab's research

This background informed FinRegLab's decision to focus its first major research and policy analysis project on the use of cash-flow data in credit underwriting. We organized two initiatives to support the broader project. The first was to conduct independent empirical research on the predictiveness of cash-flow attributes and scores, both in isolation and relative to traditional credit history information. The second was to convene a broad range of stakeholders to develop a more fulsome picture of the challenges that are shaping both the adoption of cash-flow based underwriting and the transfer of cash-flow data between companies for use in credit and other financial services.

Our goal across both workstreams was to use cash-flow based underwriting as a stepping stone to broader questions about how customer-directed data sharing can be structured to promote customer data sovereignty and protect privacy, while preserving space for firms to use that data to create financial products and services that better serve the public. Particularly given that electronic transaction account data is becoming widely available to both banks and nonbanks and is more directly reflective of applicants' finances than other forms of alternative or non-traditional data, we wanted to assess the extent to which it could make underwriting of underserved populations more cost-effective and inclusive.

The forthcoming Small Business Spotlight report provides a more focused discussion of the state of cash-flow based underwriting in the small business market, including a distillation of the empirical analysis presented here, a broader survey of recent developments in that market, and a discussion of policy issues that are of particular interest to small business applicants and credit providers.

The forthcoming Market Context and Policy report puts the results of this Empirical Research Findings report in a broader market and analytical context by building on the insights generated by the stakeholder convenings, which involved more than 80 representatives of fintech companies, banks, data aggregators, advocacy organizations, and research institutions. Representatives from the federal banking regulators and Consumer Financial Protection Bureau participated as observers. The stakeholders met over more than eight weeks in working groups to address three broad topics relating to cash-flow based underwriting: fair and inclusive access to credit, consumer understanding

22 For background on model governance expectations for federal banks, see Office of the Comptroller of the Currency, Examination Guidance on Credit Scoring Models, OCC Bulletin 97-24 (May 20, 1997); Board of Governors of the Federal Reserve & Office of the Comptroller of the Currency, Supervisory Guidance on Model Risk Management, SR 11-7 & OCC Bulletin 2011-12 (April 4, 2011).

23 12 C.F.R. § 1002.5(b). The major exception to this rule is in mortgage lending, where collection of demographic information is required under the Home Mortgage Disclosure Act. 12 U.S.C. § 2803. A 2010 amendment to ECOA that has not yet been implemented requires collection of similar information for business loans. 15 U.S.C. § 1691c-2.

24 See Subsection 4.1.3 for more discussion.

and consent issues in connection with both cash-flow based underwriting and related data transfers, and other policy concerns raised by the emergence of a new type of information ecosystem to facilitate consumer-directed transfers of transaction account data for both credit and other uses.

The balance of this report is organized into four sections, focusing on the research design and participants, methodology, key findings and implications, and conclusion. Charles River Associates' report to FinRegLab is attached as an appendix and provides more detailed summaries of the methodology and the results of the analyses performed on each individual participant's loan data.

3. RESEARCH DESIGN & PARTICIPANTS

FinRegLab's purpose in undertaking this empirical research was to conduct an independent, quantitative analysis of cash-flow scores and variables that are being used in the market today to underwrite consumers and small businesses. The participants that contributed data to the study are all focused on increasing access to underserved populations but vary widely as to business models, product structures, and underwriting processes.

3.1 Research questions

The focus of our applied research was to evaluate the cash-flow data variables and scores for their ability to predict credit risk, potential for expanding access to credit, and potential fair lending effects. With assistance from CRA, we defined three specific research questions for consideration:

» Are cash-flow variables and scores useful in predicting credit risk in the underwriting process, as compared with traditional credit scores and/or credit bureau attributes?

» Do cash-flow variables and scores expand the availability of credit, particularly with respect to consumers and small business owners who may have experienced constrained access to credit under more traditional underwriting criteria?

» What, if any, risks of creating a disparate impact among different demographic groups appear to arise from the use of cash-flow variables and scores in highly automated underwriting processes?

We structured this research to focus on evaluating the predictiveness of the particular cash-flow scores and metrics supplied by the study participants. The participants did not provide us with the underlying bank account or other records or the algorithms by which they generate cash-flow scores and metrics, make credit eligibility determinations, or determine prices. They commonly use additional information and attributes in their automated underwriting processes beyond the cash-flow metrics that were the focus of our analysis, and they did not provide the weights assigned by their algorithms to each cash-flow attribute. Thus, the participants' cash-flow metrics permitted CRA and FinRegLab to evaluate the predictiveness and fair lending effects of the variables and scores in general, but our analysis does not evaluate their particular proprietary models.

3.2 Study participants

Six non-bank financial services providers—Petal, Oportun, LendUp, Brigit, Kabbage, and Accion—participated in the research by providing data concerning their use of cash-flow variables and/or scores in underwriting unsecured, relatively short-term loans and cash advance products.[25] FinRegLab engaged CRA to conduct an independent analysis of the three research questions using these participants' data. Given each participant's interest in protecting proprietary information, we agreed to anonymize the firms in the findings and present the research results in a way that does not identify individual participants or individual cash-flow variables. In addition, the results for participants who are focused on small business markets are not separately identified from those who focus on consumer populations. Finally, discussion of certain aspects of the participants' lending processes is provided only at a group level.

The research participants are heterogeneous with respect to a wide range of factors, including business models, geographic footprint, operational structure, product offerings, application channels, tenure in specific markets, and overall lending volumes. They also take different approaches to acquiring and using cash-flow data. Four focus on consumer lending, while two serve small businesses. The participants include five for-profit firms and two Community Development Financial Institutions (CDFIs).[26] All of the credit products are unsecured, but the products vary as to closed-end and open-end structures and as to whether they are issued by the participants or by partner banks. Other terms also vary significantly. For example, repayment periods vary from the borrower's next account deposit to 46 months. Fee and rate structures also vary depending on the product type and in some cases the amount borrowed and other factors relating to borrowers' credit characteristics. Several of the participants are nationally based, while others are highly concentrated in selected geographies.

The following provides a brief overview of each of the participants' target markets, product types, and distribution channels:

» **Petal:** Petal partners with Web Bank, an FDIC-insured industrial bank chartered in Utah, to provide an unsecured credit card to consumers in amounts that range from $500 to $10,000. Marketing is aimed at consumers who have a limited credit record. Applications are accepted online.

» **Oportun:** As a certified CDFI, Oportun provides unsecured installment loans to low- and moderate-income consumers. Loans range in size from $300 to $9,000 and in length from 6 to 46 months. The maximum loan amount varies by state, and loans above $6,000 are available specifically to qualified returning customers. Consumers can apply for the loans via retail locations in some states, online, or by phone.[27]

» **LendUp:** LendUp offers installment loans and a single payment loan that is marketed as a payday loan alternative. The company uses a point system based on consumers' repayment history and completion of free on-line education courses; consumers who reach certain point levels can qualify for installment loans with larger loan amounts and lower rates, and opt to have those loans reported to build credit history. Loans meeting certain size

25 Some of the participating companies provide access to credit by partnering with or acting as service providers to financial institutions that extend loans or other credit products, but do not consider themselves to be lenders and do not themselves extend credit.

26 CDFIs are certified by the Community Development Financial Institutions Fund within the U.S. Department of Treasury based on a mission of serving low income communities, and are eligible for various types of CDFI Fund assistance and programs. CDFI Fund, CDFI Certification: Your Gateway to the CDFI Community (2016), available at www.cdfifund.gov/Documents/CDFI_CERTIFICATION_updatedJAN2016.pdf.

27 Oportun loans are available in twelve states with retail locations in the following nine states: Arizona, California, Florida, Illinois, New Jersey, New Mexico, Nevada, Texas, and Utah. Loans for residents of Idaho, Missouri, and Wisconsin are online only.

and pricing thresholds are automatically reported to consumer reporting agencies. Data on the unsecured single payment loan was evaluated in this research. That product ranges in amount from $100 to $500, with repayment due in two to four weeks. Applications are accepted online.

- **Brigit:** Brigit provides cash advances and financial monitoring tools to consumers who have an active bank account. The company uses a flat monthly subscription fee. Brigit monitors consumers' account balances to identify when a customer's balance is likely to become negative. The company will deposit an amount up to $250 to prevent an overdraft. Consumers are also permitted to request advances manually but can only request one at a time. Payment is due after the next account deposit. The Brigit product can be applied for online.
- **Kabbage:** Kabbage provides small businesses with access to unsecured lines of credit between $2,000 and $250,000 through its technology service provider relationship with Celtic Bank, an FDIC-insured industrial bank chartered in Utah. Celtic Bank requires one year of operating history and, on average, revenues of $50,000 annually or $4,200 monthly for the last three months to qualify. Average credit lines are $25,000 and average draws are $6,000; draws are treated as installment loans with terms of 6, 12 or 18 months. All business loans available through Kabbage are issued by Celtic Bank. Applications are accepted online.
- **Accion in the U.S.:** Accion in the U.S. (Accion) is a non-profit small business lender that provides installment loans of $300 or more to underserved entrepreneurs. Repayment periods are typically 24 months. Accion provides small businesses loans nationwide through four independent, regional CDFIs and a national office that coordinates technology and knowledge sharing to benefit the network. Data from one location was evaluated in this research. Accion accepts applications online.

3.3 Participants' underwriting practices

All participants use highly automated underwriting systems. From available cash-flow sources, they distill financial variables reflecting applicants' income, expenses, balances, and activity levels. For example, the cash-flow variables used by the participants may enable them to evaluate income-to-expense ratios, differences in flows of fixed and variable income, minimum balances, and/or the frequency of negative balance events as a measure of recent financial instability.

In the small business context, for example, the participants use cash-flow data to assess the business's historical and projected performance. The data includes incoming revenue, receivables, expenditures, and business obligations. The firm's financial performance may also be evaluated based on such metrics as average monthly revenue and transaction volume.

To assess consumer applicants' eligibility and creditworthiness, participants evaluate cash-flow data extending back by as much as 12 months. The small business participants sometimes consider longer periods depending on the data source and availability. Some participants pull data over time, for instance to monitor whether adjustments in the terms for open-end credit products are warranted.[28]

Across one or more participants, sources of cash-flow data included transaction account data from banks, business accounting software, payments processors, and e-commerce platforms, as well as copies of pay stubs, invoices, bill statements, and similar materials provided by applicants. The

28 We did not have access to any information regarding data that was pulled after the participants' origination decisions in connection with later monitoring or decisionmaking.

latter is part of a broader underwriting process that may allow some participants to extend credit to customers who may lack access to bank accounts and thus do not have digital cash-flow data. The participants generally use one or more data aggregators to access bank account data.

All of the participants use the cash-flow data to create proprietary assessments of repayment risk, but they vary as to the stage at which they use that information, the weight that they assign it in evaluating ability and/or propensity to repay, and the extent to which they rely upon traditional scores or attributes in sequence or in combination with cash-flow variables. For example, in at least one case, the company uses the cash-flow data to assess applicants who do not pass an initial screen using more traditional criteria. In such "second look" models, the cash-flow variables may enable credit to be extended to consumers who otherwise would have been denied credit using only the "first look" attributes.

The participants also vary as to their use of traditional credit bureau attributes and scores. Most participants will grant credit to applicants who do not have traditional credit scores, though they may factor traditional scores and attributes into their underwriting processes where available. However, the consumer participants take different approaches on whether and how they use traditional FICO scores, Vantage scores, and/or information from specialty consumer reporting agencies. Similarly, the small business participants differ as to how they approach use of business credit scores and/or the personal scores of business owners.

4. METHODOLOGY

Our core analysis used individual participants' loan-level performance data to evaluate the effectiveness of cash-flow variables and metrics in predicting credit risk both across sample populations and for specific subgroups. Where data permitted we also compared the predictiveness of cash-flow data relative to traditional credit scores and credit bureau attributes, as well as the predictiveness of models that combined both cash-flow and traditional sources.

4.1 Data and methodology

Working with CRA, FinRegLab requested from each of the participants application- and loan-level data on cash-flow variables and scores, traditional credit scores or other attributes from traditional credit reports, amounts and durations for originated loans, loan performance (e.g., default or delinquency information), and certain demographic proxies to facilitate the fair lending analysis. Five of the six study participants provided data for more than 90,000 originated loans overall, though the scope of the information provided varied somewhat from company to company. The sixth company produced the results of an internal analysis of more than 20,000 loans. CRA and FinRegLab have evaluated those results for insights about the three research questions but cannot independently verify the company's underlying quantitative analysis.

Populations for which data were provided were not subjectively selected by the participants; rather, the participants generally provided data for all applications evaluated and/or loans originated within specified time periods, which were defined to increase the likelihood that the loans were sufficiently seasoned to enable measurement of performance. CRA worked with each participant to refine the data request based on the products, underwriting procedures, and data maintained by each institution.

The diversity of the participants and data prevented combining the data to perform a consolidated analysis. As noted above, the participants differ with regard to the products that they offer, the populations that they serve (consumer vs. small business), the types of cash-flow and traditional data that they rely upon, the ways in which they use such data, and the metrics that they focus on in defining and tracking default, delinquency, or other poor loan performance.

CRA helped FinRegLab to define the analytical approach specific to each of the research questions described above to the data from each of the five participants, as described separately below. CRA performed all of the data analysis as described further in their report, which is provided in the Appendix.[29]

4.1.1 Predictiveness

CRA's analysis uses the loan-level performance data to assess the extent to which cash-flow variables and/or scores can facilitate the evaluation of credit risk. The objective in modeling the default risk is to determine the extent to which the factors are predictive of which customers were, in fact, more likely to repay and which customers were, in fact, less likely to repay. Such models can be used to rank order customers from highest to lowest default risk. Where the participants also made available credit scores or other information from traditional credit reports, the analysis also evaluates the predictiveness of the traditional attributes and of a combination of traditional attributes and cash-flow attributes relative to actual loan performance.

The analyses proceeded in two phases for each set of variables (cash-flow only, traditional only, and combined), as described in more detail in the Appendix. First, CRA used difference of means tests to examine correlations between each individual variable or score and default status. Second, CRA calculated multivariate logit models to ascertain the relationship between all attributes and default or delinquency. In connection with those multivariate models, CRA then computed the "receiver operating characteristic" (ROC) and the "area under the ROC curve" (AUC), which are standard measures of model fit or performance used by developers of credit risk models.

Because the AUC statistics provide an overall performance measure for the various combinations of variables in separating customers who defaulted or were delinquent from those who were not, the summary of results below focuses primarily on this metric. A model that performed no better than random chance would have an AUC of 0.5, while a model that performs perfectly in predicting default would have an AUC of 1.0.

4.1.2 Inclusiveness

In addition to analyzing general predictiveness, we set out to assess the extent to which the study participants are serving consumers and small business owners who may have experienced constrained access to credit under more traditional underwriting criteria. However, both definitional challenges and data limitations made it impracticable to perform consistent quantitative analyses across all participants. For example, we could not determine the precise number of borrowers who lack traditional credit files or have such thin files that their credit scores may not be reliable.[30] Accordingly, CRA looked at a number of different metrics depending on data availability to obtain additional insights about the extent to which cash-flow data may be increasing access to these populations.

For instance, for the participants that provided loan-level data, the analysis reviews a range of factors including income, residence in zip codes in which minorities make up at least 50 percent

29 CRA also defined and provided the logistical support necessary to manage the data transfers, encryption, information technology security, and similar matters.

30 The Consumer Financial Protection Bureau has estimated based on a 1-in-48 representative sample from one of the NCRAs using 2010 data that there are 26 million consumers with no credit files at NCRAs, and another 19 million consumers who have such limited files that they are treated as unscoreable by a commercially available credit scoring model to which researchers had access. However, consumers may have differing amounts of information in their credit files at different NCRAs, and different third-party models may have different criteria for scoreability. CFPB Credit Invisibles at 4-6; FRB Consumer Scoring Report at 16-17.

BOX 4.1.1 AUC METRICS

AUC metrics are often used by developers of credit risk or other predictive models. A model that performed no better than random chance would have an AUC of 0.5, while a model that performs perfectly would have an AUC of 1.0. However, there are no objective benchmarks for AUCs between .5 and 1.0 because their values depend on the usage context. In certain areas of medical research, AUCs of 0.95 or higher may be obtained, but in research on financial services much lower numbers are often reported. Some financial services sources suggest that an AUC of .6 is generally considered desirable in information-scarce environments, while AUCs of .7 or greater are the goal in information-rich environments.

Comparisons across studies are difficult because of these factors, particularly where different researchers are analyzing different products, populations, and underwriting methodologies. However, other studies analyzing underwriting models for unsecured products and populations similar to the ones analyzed in this report have reported AUC values for traditional credit scores in the .6 range, which is similar to the results here.

Sources: Rajkamal Iyer et al., Screening Peers Softly: Inferring the Quality of Small Borrowers, 62 Management Science 1554, 1562 (2016); Tobias Berg et al., On the Rise of the FinTechs—Credit Scoring Using Digital Footprints, FDIC Working Paper 2018-04, at 4 (September 2018); Bowen Baker, Consumer Credit Risk Modeling, MIT Departments of Physics and EECS (December 2015).

of the population, and no or low traditional credit scores to the extent that such information was available. FinRegLab and CRA also reviewed the internal analysis by the other participant which included an evaluation of borrower income levels as discussed below.

4.1.3 Fair lending effects

As described above, fair lending law has two principal theories of liability, disparate treatment and disparate impact. CRA designed the analysis to evaluate potential disparate impact risks in using the cash-flow variables and scores in underwriting algorithms. Four participants provided sufficient data to permit this type of analysis for race/ethnicity, and three participants were also able to provide data to analyze potential impact based on gender.

Because collection of data on protected characteristics is prohibited under fair lending laws for the credit products covered by this research,[31] the analysis was conducted after applying a proxy methodology to assess customers' likely demographic group. These same kinds of techniques are commonly used and accepted by federal regulators in evaluating compliance with fair lending laws.[32] For example, CRA validated that the race/ethnicity probabilities were computed in a manner not materially different from the assumptions reflected in computer code that has been publicly released by the Consumer Financial Protection Bureau.[33]

Due to sample population considerations and data limitations, CRA could not perform certain types of statistical tests that are frequently used in both the first and third phases of disparate impact analyses. For example, it was not possible to calculate the degree to which average approval rates by demographic group were impacted by the cash-flow metrics or scores, as they were only components of the participants' overall underwriting processes and CRA did not have access to the other attributes or the weights assigned to the attributes by the participants' algorithms. It also was not possible to estimate average approval rates using alternative cash-flow metrics, as we did not have access to the underlying data with which to construct alternative variables.

31 12 C.F.R. § 1002.5(b). See Subsection 2.2 for more discussion.

32 *See, e.g.*, Consumer Financial Protection Bureau, Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity: A Methodology and Assessment (2014).

33 Available at github.com/cfpb/proxy-methodology. Multiple commercial software packages are available to create gender proxies.

However, the data did permit calculation of one of the principal tests that the Federal Reserve Board used to assess fair lending considerations with regard to traditional credit scores by evaluating the degree to which particular variables are predictive across different demographic groups.[34] Specifically, the test requires the sample populations to be subdivided by demographic group, such that the predictiveness of the cash-flow metrics can be measured within each group and the results compared across groups. The test is useful for evaluating potential disparate impact risk because if the cash-flow score or metric fails to be predictive or is substantially less predictive of credit risk among a particular demographic group relative to its predictiveness for a relevant comparison group (for example: non-Hispanic white customers), such a result may suggest a heightened risk that the particular variables or scores are acting as a proxy for protected class status rather than providing independent predictive value.

To apply the test, among each participant's sample population, the proxy methodologies were used to identify customers with high probabilities of belonging to each race, ethnicity, and gender group.[35] Similar to the analysis of general predictiveness, for each of the resulting subgroups CRA proceeded first by applying difference in means tests for the individual cash-flow variables and scores and then by calculating AUCs for the multivariate models. Where possible, CRA also calculated AUCs for each demographic group for the models that relied solely on traditional scores and attributes, and on the combined models that used both traditional and cash-flow information.

BOX 4.1.3 PROXY METHODOLOGIES

In conducting fair lending examinations and internal compliance analyses, federal regulators and industry often use a method called Bayesian Improved Surname Geocoding to assess the likely race/ethnicity of borrowers. The technique uses surnames and geography of residence to calculate the likelihood of belonging to particular subpopulations based on a comparison to U.S. Census data. Proxy methodologies for gender often focus primarily on first names as reported by the Social Security Administration.

While such methods are commonly used and accepted by federal financial regulators, by their nature they are somewhat inexact. Academic research indicates that proxy methodologies can produce measurement errors in certain circumstances as both overinclusive (by assigning a high probability of belonging to the wrong group) and underinclusive (by assigning a low probability of belonging to the correct group).

Sources: Consumer Financial Protection Bureau, Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity: A Methodology and Assessment (2014); Patrice Ficklin, Preventing Illegal Discrimination in Auto Lending, www.consumerfinance.gov/about-us/blog/preventing-illegal-discrimination-in-auto-lending/ (Nov. 4, 2013); Yan Zhang, Assessing Fair Lending Risks Using Race/Ethnicity Proxies, 64 Management Science 178 (2018).

34 FRB Scoring Report at 109-116; see also Consumer Financial Protection Bureau, Examination Procedures: ECOA Baseline Review, Module 5(f) (April 2019) (focusing on whether entities evaluate the validity or performance of their models by prohibited basis group). For other discussions of similar techniques, see David Skanderson & Dubravka Ritter, Fair Lending Analysis of Credit Cards, Federal Reserve Bank of Philadelphia Payment Cards Center Discussion/Working Paper 14-02, at 34-40 (August 2014); Elaine Fortowsky & Michael LaCour-Little, Credit Scoring and Disparate Impact, Wells Fargo Home Mortgage Working Paper 20-21 (2001); Stephen L. Ross & John Yinger, The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement (2002).

35 CRA used a probability threshold of 75 percent to define which loans were assigned to which demographic groups. For more discussion of this approach, see the Appendix.

4.2 Implications

Before discussing the specific results of the various analyses, it is helpful to note two important implications with regard to the research approach, data, and methodology.

4.2.1 Heterogeneity

First, the fact that the participants are so heterogeneous along the dimensions described above has both advantages and disadvantages with regard to the structure of the analysis. The strength of this approach is that each participating company represents an independent case study on the use of cash-flow data. Each institution has already invested significant resources to identify and test various relationships among cash-flow data and other factors that impact credit risk. The participants provided a description of their extensive model development efforts, which yielded underwriting models that they believe to be robust and predictive as used in their day-to-day operations. Because this analysis uses their loan-level performance data, variables or scores, and definitions of default or delinquency, it tests the potential predictiveness of the variables and scores using actual performance data over time for products of varying durations used by both consumer and small business populations, rather than theorizing about a potential set of relationships that may exist.

At the same time, there are also some disadvantages. As noted above, FinRegLab and CRA concluded that it was not practicable to aggregate the data across the participants. While most of the providers had substantial loan volumes, allowing us to undertake statistical testing, it was not practicable to draw conclusions about individual cash-flow attributes because not all participants used the same cash-flow attributes or in some cases, even similar ones. In addition, the applicant and loan populations, while sizeable, do not appear to be representative of the overall U.S. population. And while the ability to track actual loan performance for specific products over time is a strength, there is no way to assess the predictiveness of the variables and scores with regard to applicants who were rejected. The analysis is thus different from the way that scoring model developers often assess the predictiveness of potential generic scoring models using large populations to measure the relationship between particular criteria and negative loan performance on any reported credit products over a particular period of time.[36] Finally, due to limitations in the time periods covered, we were not able to assess the actual performance of these models in more adverse economic conditions.

4.2.2 Comparability

Direct comparisons of one participant's results to another's should be discouraged. As noted above, the participants provided individual cash-flow variables and scores that they rely upon as components in their overall underwriting processes rather than their full underwriting models. Thus, these results should not be interpreted as any participant's overall ability or approach to modelling credit risk. Comparisons are also inapposite because the participants are serving different populations with different credit products and tracking different measures of delinquency, default, or other poor loan performance. Further, automated underwriting processes that use cash-flow data for a second-look analysis would be expected to have different results than algorithms that use such variables to evaluate all applications from the outset.

Interpretation of the comparisons of cash-flow variables and scores to traditional scores or other attributes also requires some caution. The participants provided traditional credit report information

36 See *supra* n.8.

because it provides insight into which borrowers may historically have faced constraints on their ability to access credit. But much as with the cash-flow variables and scores, the traditional scores and attributes provided and the ways in which they are used varied from company to company. Moreover, with regard to traditional scores that are generated by national consumer reporting agencies or other third parties, as noted above many of these are generic scores that may not be generally very predictive for the particular populations or products that are the focus of the participants. In addition, the traditional scores may have been developed using a different definition of default than the ones used by some participants.

5. KEY FINDINGS & IMPLICATIONS

As described below and further detailed in the Appendix, this analysis confirms that varying types of cash-flow data are being used to underwrite credit for a range of unsecured consumer and small business credit products across a broad set of U.S. geographies. In particular, for the participants for which loan-level data was available, we find compelling evidence indicating that the cash-flow variables and scores were predictive of credit risk and loan performance across the heterogenous set of providers, populations, and products studied.

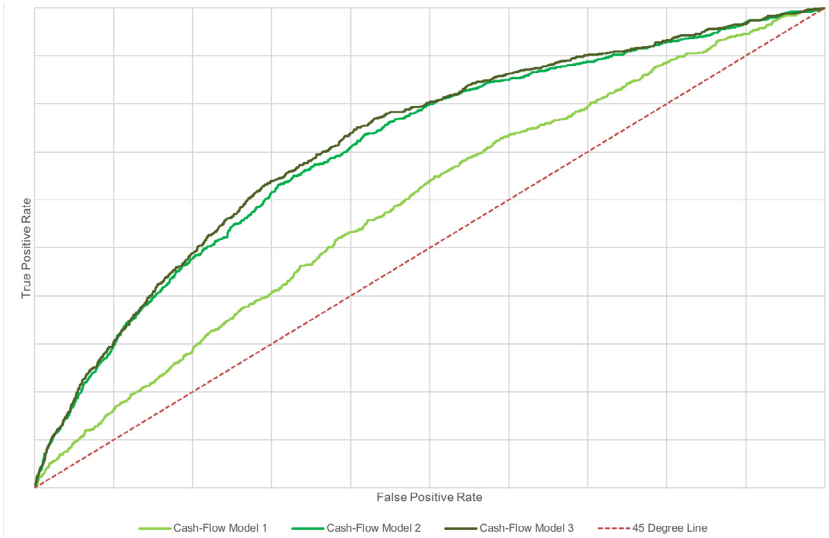
5.1 Predictiveness

We found compelling evidence that indicates that among the sample populations and products, cash-flow variables and scores are predictive of credit risk and loan performance across the highly heterogeneous set of research participants. In separate analyses of the five participants that provided loan-level data, the results appear to be robust across both consumer and small business populations as well as across the credit spectrum, including among borrowers with no or very low traditional credit scores. The cash-flow metrics were both predictive in their own right and also frequently improved the ability to predict credit risk in combination with traditional credit scores or other metrics.

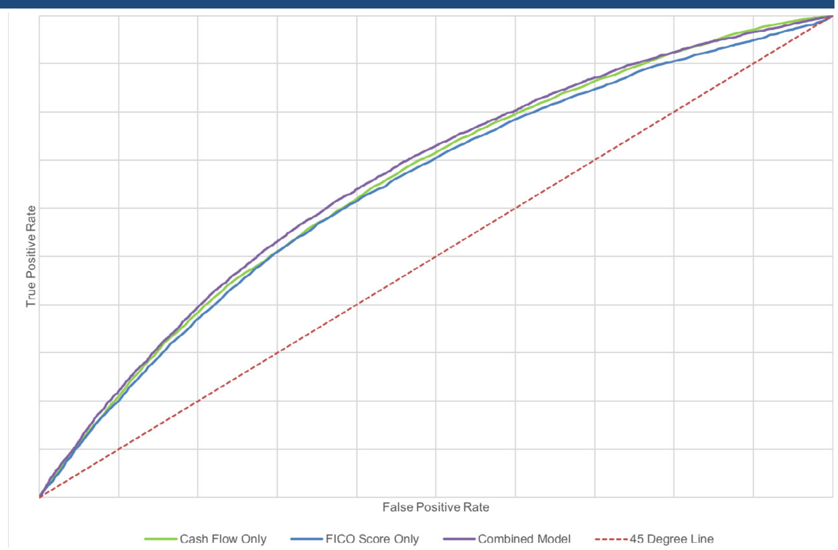
More specifically, for four of the five participants for which loan-level analyses were conducted, we found compelling evidence that indicates that the cash-flow variables and/or scores are correlated with the likelihood of default. The AUCs for various cash-flow only models ranged from .592 to .725. As illustrated by the attached graphs, these results meaningfully diverge from .5 (which is depicted as a 45-degree line and indicates no predictive power), and are at levels that in the experience of CRA suggest a relatively robust ability to predict likelihood of default within the test samples, independent of any use of traditional metrics. The fifth company's AUC was .572. Although consistent with the broader finding, these results permit a less conclusive interpretation because of a relatively small number of delinquent loans in the time period studied.

For four participants, we were also able to calculate AUCs for traditional scores or attributes from traditional consumer reports, as well as for combined metrics that used both traditional and cash-flow data. As illustrated in the graphs, for three participants, the AUCs for the cash-flow only metrics were at least as high as for the traditional-only metrics standing alone. In the fourth case,

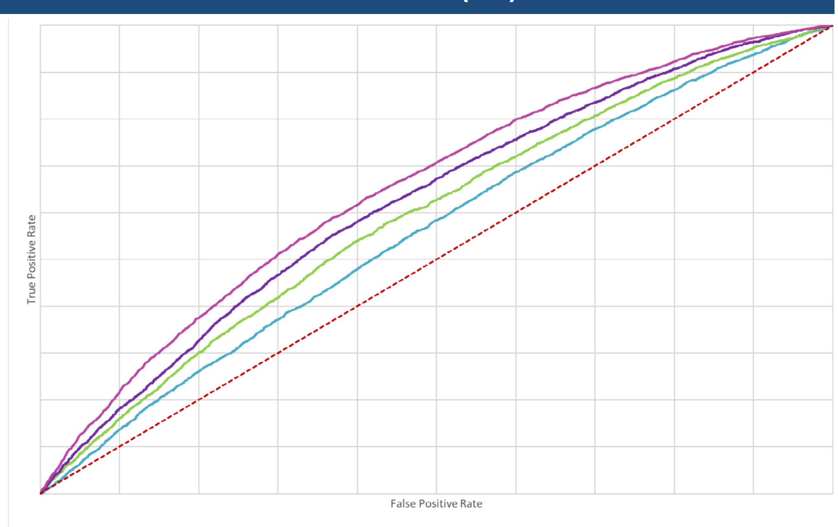
PARTICIPANT #1 RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES FOR MODELS 1-3



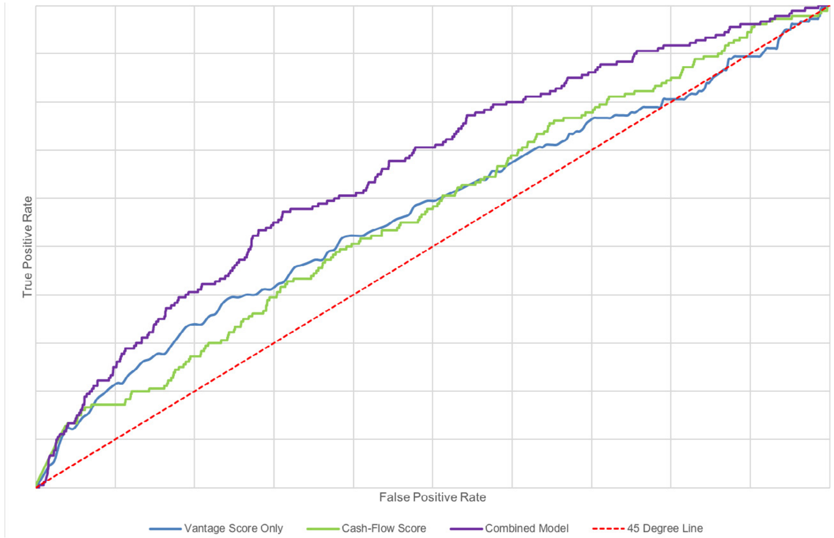
PARTICIPANT #2 RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES FOR MODELS 1-3



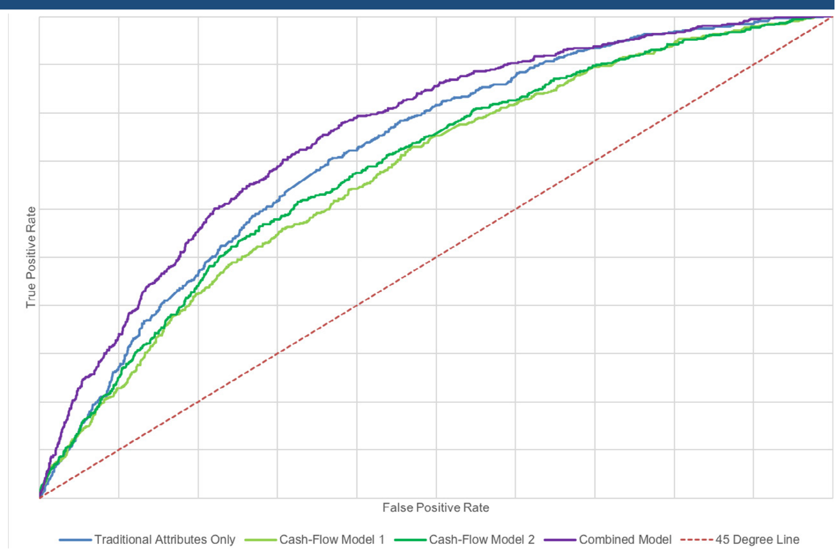
PARTICIPANT #4 RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES FOR MODELS 1-4



PARTICIPANT #5 RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES FOR MODELS 1-3



PARTICIPANT #6 RECEIVER OPERATING CHARACTERISTIC (ROC) CURVES FOR MODELS 1-4



different combinations of cash-flow metrics generated relatively robust AUCs of .675 and .688; although those were lower than the AUC for traditional FICO plus multiple traditional attributes at .720, a combined model incorporating both sets of data generated an AUC of .758. Combined models for the other three lenders also showed improvements in AUCs compared to using only cash-flow or traditional data alone, although the magnitude of the improvements varied.

The participant that performed the internal analysis focused on a cash-flow metric score that it uses in assessing ability to repay and feeds into a more complex model evaluating propensity to repay. Specifically, the analysis benchmarked the cash-flow metric against a traditional debt-to-income (DTI) ratio both for its new borrowers as a whole and separately for borrowers that had valid FICO scores and those that did not.[37] When the entire population was divided into deciles based on their cash-flow metric scores and DTIs, there was a linear relationship between average

37 DTI was calculated using a subset of the factors utilized in the cash-flow metric.

risk of serious delinquency by decile and the cash-flow metric scores except in the highest group. However, the AUC for the entire population was .532 for the cash-flow metric score, compared with .513 for DTI. For the two separate groups, the AUC for the group without a valid FICO score was .537, compared to .523 for borrowers who did have a score. Thus, while the various AUCs were statistically significant, their closeness to .5 does not suggest that the cash-flow metric had a robust ability to predict delinquency across the entire sample population.

Overall, these results have two important implications. First, the overall strength of these results and the nature of the participating companies' underwriting practices suggest that cash-flow variables and scores can provide meaningful predictive power among populations and products similar to those studied where traditional credit history is not available or reliable.

Second, the fact that cash-flow attributes and scores frequently improved predictiveness in combination with traditional credit history is noteworthy. The improvement in AUCs for combined models and our other analyses suggest that cash-flow information separates risk in somewhat different ways than traditional metrics. Overall, the results suggest that with regard to populations and products similar to those studied, cash-flow data can provide meaningful insights by differentiating predictions of credit risk among borrowers that are scored by traditional systems as presenting similar risks of default.

In particular, the following "heat maps" for the four participants provide a visualization of cash-flow metrics' ability to differentiate risk levels relative to traditional metrics. The maps divide each participants' borrower population into percentiles based on their relative traditional credit scores or metrics and their cash-flow scores or metrics. Each row of the charts represents a group of borrowers who are classified as having similar default risk based on traditional scores or metrics, while the columns further separate out those borrowers into bands based on the risk levels predicted by their cash-flow metrics or scores. Cells with more than five observations report the percentage of borrowers in each group that became delinquent or defaulted. Color codes were then assigned to those cells based on the extent to which the group's delinquency/default rates differ from the median delinquency/default rate for the participant's borrower population as a whole, with dark green for the lowest default frequency, yellow for delinquency rates close to the median, and red representing the highest default frequency.[38]

In viewing the maps, relatively consistent patterns emerge for three of the participants; the patterns in the fourth participant's chart are somewhat less clear due in part to a relatively small number of delinquent loans in the time period studied. The first pattern that may strike viewers is that cells in the top left corner tend to be red and the bottom right corner tend to be green, which is not surprising since in those cases both the traditional metrics and the cash-flow metrics tend to predict relatively high probabilities and low probabilities of default, respectively. The second pattern, however, emerges if the rows are viewed one at a time: Customers in the left-hand columns (who are predicted to have more credit risk based on cash-flow metrics) in fact tend to have relatively high delinquency rates relative to the customers in the right-hand columns (who are predicted to be less risky based on cash-flow metrics), notwithstanding the fact that all customers within the particular rows are predicted to have similar default risk based upon traditional credit scores or other metrics. This generally is true even in the bands for middle and high traditional scores or metrics. Particularly when combined with the overall AUC scores, this suggests that cash-flow variables tend to improve the sorting of risk relative to relying on traditional sources alone.

38 Because median default rates vary among the individual participants, a particular default rate might be color coded differently on the heat maps for different participants.

PARTICIPANT #2: DELINQUENCY FREQUENCY BY CASH-FLOW SCORE PERCENTILE AND FICO SCORE PERCENTILE

	CASH-FLOW SCORE																			
FICO SCORE	0 - 5TH	5 - 10TH	10 - 15TH	15 - 20TH	20 - 25TH	25 - 30TH	30 - 35TH	35 - 40TH	40 - 45TH	45 - 50TH	50 - 55TH	55 - 60TH	60 - 65TH	65 - 70TH	70 - 75TH	75 - 80TH	80 - 85TH	85 - 90TH	90 - 95TH	95 - 100TH
0 - 5th	35.5%	26.4%	31.4%	31.5%	27.8%	25.0%	22.9%	9.7%	13.6%	25.0%	18.2%	20.0%	.	60.0%	.	.	.	.	.	.
5 - 10th	33.0%	31.0%	29.7%	28.8%	25.4%	20.3%	29.4%	29.3%	25.0%	40.0%	4.8%	0.0%	11.1%	0.0%	.	.	.	.	.	.
10 - 15th	32.7%	35.6%	27.1%	27.4%	32.7%	23.2%	20.6%	19.3%	14.3%	24.6%	20.9%	8.5%	34.3%	12.0%	11.8%	12.5%	16.7%	.	.	.
15 - 20th	37.7%	25.1%	28.4%	26.2%	28.2%	26.0%	25.4%	21.2%	27.6%	20.6%	21.1%	25.0%	17.9%	14.0%	17.4%	9.1%	18.2%	0.0%	.	.
20 - 25th	30.3%	34.4%	30.5%	28.9%	23.8%	26.8%	23.2%	21.9%	20.7%	20.5%	12.5%	21.3%	12.2%	25.0%	11.5%	11.4%	12.0%	8.3%	20.0%	0.0%
25 - 30th	33.8%	34.1%	29.0%	22.8%	34.6%	23.5%	16.3%	25.2%	24.3%	20.5%	13.9%	19.8%	17.9%	22.4%	10.3%	13.4%	11.5%	5.3%	0.0%	12.5%
30 - 35th	27.4%	30.2%	27.9%	30.9%	24.0%	26.0%	23.5%	19.0%	16.8%	16.1%	16.4%	20.0%	12.2%	11.4%	16.5%	18.0%	19.2%	13.0%	12.5%	9.1%
35 - 40th	24.0%	22.6%	33.3%	25.0%	25.2%	21.4%	19.8%	19.7%	16.2%	17.0%	15.3%	17.2%	15.0%	13.3%	13.4%	13.9%	19.7%	6.4%	2.2%	13.8%
40 - 45th	18.9%	27.5%	33.8%	27.5%	17.1%	19.4%	24.2%	10.1%	21.4%	19.6%	14.4%	10.8%	12.8%	10.5%	12.5%	16.9%	10.8%	14.3%	13.0%	6.5%
45 - 50th	20.7%	7.1%	17.2%	18.8%	22.6%	11.7%	18.1%	24.2%	19.1%	19.0%	20.3%	20.4%	12.7%	10.8%	12.5%	13.5%	10.9%	10.7%	12.0%	3.0%
50 - 55th	32.0%	10.3%	23.9%	16.1%	19.5%	20.0%	15.1%	14.1%	15.4%	17.0%	14.8%	17.8%	12.9%	11.9%	16.2%	12.3%	10.7%	14.5%	8.9%	10.6%
55 - 60th	30.0%	15.2%	14.6%	15.4%	21.5%	22.9%	14.8%	17.3%	15.1%	15.7%	11.4%	16.8%	10.5%	15.4%	9.7%	10.9%	7.3%	8.5%	7.3%	8.5%
60 - 65th	33.3%	20.7%	24.5%	12.5%	20.4%	13.2%	21.0%	15.8%	25.7%	13.7%	12.6%	10.3%	10.4%	16.0%	12.2%	9.5%	9.5%	8.6%	10.3%	9.1%
65 - 70th	30.0%	15.4%	13.6%	20.0%	16.0%	18.4%	7.1%	19.8%	18.8%	13.1%	17.0%	11.6%	8.1%	7.7%	10.6%	11.2%	12.3%	10.3%	3.5%	6.2%
70 - 75th	12.5%	18.8%	19.4%	15.4%	12.3%	9.5%	11.3%	10.6%	14.1%	15.7%	11.8%	11.0%	12.0%	12.6%	14.8%	11.4%	6.7%	9.1%	4.3%	4.4%
75 - 80th	19.0%	10.5%	22.2%	14.3%	17.8%	15.3%	12.7%	12.5%	16.2%	11.2%	17.1%	9.3%	10.0%	11.6%	9.4%	13.4%	9.5%	9.7%	10.4%	2.7%
80 - 85th	18.8%	31.3%	11.8%	12.5%	0.0%	6.6%	14.9%	12.0%	14.8%	10.9%	12.0%	6.4%	9.0%	7.6%	5.3%	7.5%	4.0%	8.1%	7.5%	3.9%
85 - 90th	15.4%	42.1%	33.3%	20.0%	13.0%	8.8%	6.3%	10.7%	14.1%	16.2%	10.7%	6.4%	9.2%	8.1%	8.9%	5.6%	6.1%	4.1%	5.9%	3.0%
90 - 95th	28.6%	36.4%	29.2%	23.3%	6.7%	7.5%	12.5%	6.0%	14.0%	10.1%	20.2%	10.2%	9.9%	6.5%	5.2%	8.3%	6.9%	9.1%	6.0%	2.8%
95 - 100th	18.2%	21.4%	55.6%	38.9%	17.2%	12.5%	10.9%	12.1%	15.9%	11.7%	16.5%	13.3%	13.7%	8.3%	10.9%	11.0%	6.6%	11.2%	4.2%	1.7%

NOTES: 1) Cells are shaded based on values. Green indicates values close to the lowest default frequency, yellow indicates values close to the median default frequency, and red indicates values close to the highest default frequency. 2) Cells with fewer than 5 loans are excluded from this heat map. 3) Percentiles are based on the population of originated loans. 4) 381 originated loans with a missing FICO score were excluded from the frequency table.

PARTICIPANT #4: DEFAULT FREQUENCY BY CFPD PERCENTILE AND TPD PERCENTILE

	CASH-FLOW BASED PROBABILITY OF DEFAULT																			
TRADITIONAL PROBABILITY OF DEFAULT	100TH	95TH	90TH	85TH	80TH	75TH	70TH	65TH	60TH	55TH	50TH	45TH	40TH	35TH	30TH	25TH	20TH	15TH	10TH	0 - 5TH
95 - 100th	33.3%	23.1%	36.4%	20.0%	35.4%	27.0%	19.0%	27.9%	32.5%	13.5%	16.2%	10.6%	23.2%	21.1%	12.8%	16.3%	21.3%	14.3%	10.1%	12.8%
90 - 95th	28.6%	40.0%	36.6%	39.6%	28.3%	32.5%	20.5%	22.2%	18.8%	18.3%	24.7%	22.8%	23.9%	23.5%	11.1%	22.8%	18.4%	17.0%	4.9%	20.5%
85 - 90th	20.0%	.	22.9%	24.1%	30.0%	18.0%	28.6%	33.9%	16.4%	25.3%	25.6%	20.4%	14.3%	18.1%	21.1%	22.2%	13.9%	13.4%	11.8%	12.7%
80 - 85th	.	.	22.5%	27.3%	21.8%	23.3%	18.3%	17.3%	28.1%	22.7%	11.5%	28.0%	16.9%	11.6%	11.3%	7.5%	11.7%	12.5%	16.9%	15.9%
75 - 80th	.	.	23.4%	20.0%	25.4%	23.7%	20.7%	17.6%	14.8%	20.5%	16.2%	13.9%	8.8%	13.8%	16.2%	10.7%	13.2%	16.7%	11.3%	6.1%
70 - 75th	0.0%	17.4%	26.8%	29.7%	16.5%	15.4%	23.1%	19.3%	31.9%	6.9%	15.6%	18.6%	25.0%	22.4%	15.4%	21.7%	10.0%	9.5%	9.7%	12.5%
65 - 70th	.	9.1%	25.8%	21.9%	17.2%	22.9%	20.5%	20.8%	11.4%	18.7%	9.5%	10.9%	9.4%	16.0%	16.7%	20.0%	21.6%	8.7%	8.3%	2.0%
60 - 65th	21.1%	25.8%	16.5%	36.1%	25.4%	20.8%	25.0%	26.3%	15.7%	4.8%	10.3%	20.0%	9.4%	17.7%	18.5%	10.5%	7.7%	10.2%	13.7%	7.6%
55 - 60th	25.9%	27.8%	21.5%	20.8%	27.1%	12.3%	20.9%	28.2%	19.7%	14.7%	16.4%	18.1%	14.3%	14.5%	11.3%	10.2%	8.9%	7.3%	15.4%	12.3%
50 - 55th	25.5%	20.5%	25.4%	21.6%	15.2%	20.7%	17.8%	17.5%	6.5%	20.7%	9.4%	16.7%	10.0%	20.0%	8.1%	11.7%	14.8%	6.3%	4.2%	5.1%
45 - 50th	20.3%	32.3%	24.7%	18.2%	20.8%	18.6%	22.7%	22.2%	12.3%	9.1%	10.8%	9.1%	16.4%	10.8%	19.7%	14.0%	10.2%	3.6%	7.0%	3.3%
40 - 45th	34.6%	20.0%	28.2%	30.0%	19.7%	15.8%	20.6%	17.3%	10.5%	8.8%	5.1%	21.3%	15.6%	8.2%	13.5%	9.3%	10.6%	7.7%	8.0%	11.4%
35 - 40th	23.4%	16.8%	20.4%	22.4%	15.5%	12.5%	15.2%	18.6%	15.7%	12.1%	14.8%	17.5%	5.9%	14.5%	10.7%	13.0%	8.5%	8.8%	12.8%	8.8%
30 - 35th	30.5%	24.8%	18.2%	21.5%	19.3%	13.8%	15.3%	14.5%	11.5%	8.9%	17.9%	14.1%	10.5%	7.4%	14.3%	10.3%	12.2%	7.7%	3.4%	3.3%
25 - 30th	19.0%	24.2%	15.1%	15.9%	19.1%	19.0%	7.4%	16.7%	15.6%	1.7%	10.0%	17.2%	16.0%	11.5%	13.9%	12.5%	7.7%	7.1%	1.8%	13.7%
20 - 25th	21.3%	21.6%	15.4%	9.7%	12.1%	13.6%	17.6%	16.9%	16.9%	8.0%	11.8%	13.2%	17.2%	11.5%	13.4%	9.3%	5.8%	7.4%	3.9%	6.1%
15 - 20th	27.7%	25.6%	19.4%	5.9%	15.5%	16.9%	8.3%	7.3%	11.1%	11.1%	18.4%	14.3%	5.3%	7.5%	7.8%	10.0%	3.7%	8.6%	6.0%	1.5%
10 - 15th	20.6%	21.1%	16.7%	14.0%	4.4%	10.7%	13.6%	13.0%	6.3%	7.2%	11.7%	19.6%	7.8%	5.5%	10.5%	9.1%	9.4%	13.1%	4.8%	7.1%
5 - 10th	21.1%	16.0%	8.3%	22.2%	17.6%	13.6%	6.0%	13.2%	4.4%	14.8%	9.3%	8.1%	12.7%	7.1%	6.7%	12.3%	5.7%	4.3%	3.2%	5.8%
0 - 5th	19.6%	18.8%	16.7%	13.3%	7.5%	6.9%	11.3%	12.0%	12.9%	7.0%	4.0%	12.5%	6.0%	2.5%	7.7%	4.9%	3.4%	4.3%	2.9%	3.9%

NOTES: 1) Cells are shaded based on values. Green indicates values close to the lowest default frequency, yellow indicates values close to the median default frequency, and red indicates values close to the highest default frequency. 2) Cells with fewer than 5 loans are excluded from this heat map. 3) Percentiles are based on the population of originated loans with a known empirical default status.

PARTICIPANT #5: PAST DUE FREQUENCY BY CASH-FLOW AND VANTAGE SCORE PERCENTILE

	CASH-FLOW SCORE									
VANTAGE SCORE	10TH	20TH	30TH	40TH	50TH	60TH	70TH	80TH	90TH	100TH
0 - 10th	3.5%	6.6%	5.0%	2.6%	8.2%	1.1%	5.0%	3.8%	5.1%	0.0%
10 - 20th	0.0%	1.3%	2.5%	4.6%	0.0%	3.2%	2.1%	3.7%	3.4%	10.7%
20 - 30th	1.6%	1.4%	2.8%	0.0%	1.3%	2.2%	0.0%	2.2%	3.0%	0.0%
30 - 40th	1.3%	1.2%	2.2%	0.9%	3.0%	0.9%	1.3%	4.0%	3.2%	0.0%
40 - 50th	2.1%	1.4%	1.1%	5.8%	0.0%	1.2%	2.2%	0.0%	1.0%	3.6%
50 - 60th	0.0%	2.2%	1.0%	0.0%	2.1%	1.0%	4.4%	4.0%	0.0%	0.0%
60 - 70th	1.4%	1.2%	0.0%	3.2%	0.0%	1.5%	1.7%	3.9%	0.0%	3.6%
70 - 80th	2.4%	0.0%	0.0%	1.1%	1.5%	2.6%	0.0%	1.7%	0.0%	0.8%
80 - 90th	2.2%	0.0%	0.0%	1.8%	1.8%	1.8%	0.0%	0.0%	0.0%	1.3%
90 - 100th	0.0%	1.2%	0.0%	0.0%	2.2%	0.0%	0.0%	0.0%	3.0%	2.2%

NOTES: 1) Cells are shaded based on values. Green indicates values close to the lowest default frequency, yellow indicates values close to the median default frequency, and red indicates values close to the highest default frequency. 2) Cells with fewer than 5 loans are excluded from this heat map. 3) Percentiles are based on the population of originated loans. 4) 304 originated loans with a missing Pre-Qual. Vantage score and 335 originated loans with a missing Cash-Flow Score were excluded from the frequency table.

PARTICIPANT #6: DELINQUENCY FREQUENCY BY FICO SCORE PERCENTILE AND MODEL 2's PREDICTED PROBABILITY OF DELINQUENCY PERCENTILE

	MODEL 2's PREDICTED PROBABILITY OF DELINQUENCY																			
FICO SCORE	95 - 100TH	90 - 95TH	85 - 90TH	80 - 85TH	75 - 80TH	70 - 75TH	65 - 70TH	60 - 65TH	55 - 60TH	50 - 55TH	45 - 50TH	40 - 45TH	35 - 40TH	30 - 35TH	25 - 30TH	20 - 25TH	15 - 20TH	10 - 15TH	5 - 10TH	0 - 5TH
0 - 5th	41.7%	22.7%	33.3%	38.5%	37.5%	33.3%	44.4%	.	30.0%	27.3%	50.0%	20.0%	.	.	25.0%	0.0%	0.0%	.	.	.
5 - 10th	25.0%	52.9%	77.8%	27.3%	12.5%	11.1%	22.2%	25.0%	45.5%	33.3%	33.3%	16.7%	22.2%	0.0%	0.0%	33.3%	.	12.5%	20.0%	.
10 - 15th	36.4%	22.2%	55.6%	30.8%	31.3%	23.1%	0.0%	45.5%	33.3%	27.3%	25.0%	0.0%	40.0%	25.0%	16.7%	20.0%	.	.	.	0.0%
15 - 20th	44.4%	28.6%	33.3%	36.4%	27.3%	30.0%	27.3%	16.7%	27.3%	37.5%	10.0%	8.3%	14.3%	12.5%	20.0%	20.0%	33.3%	.	.	0.0%
20 - 25th	35.7%	16.7%	63.6%	42.9%	23.1%	9.1%	0.0%	20.0%	22.2%	33.3%	30.0%	37.5%	20.0%	11.1%	0.0%	0.0%	11.1%	.	14.3%	.
25 - 30th	50.0%	.	8.3%	12.5%	40.0%	23.1%	15.4%	8.3%	0.0%	28.6%	20.0%	16.7%	11.1%	0.0%	0.0%	20.0%	0.0%	0.0%	11.1%	20.0%
30 - 35th	13.3%	15.4%	25.0%	30.0%	7.1%	0.0%	27.3%	9.1%	20.0%	.	0.0%	11.1%	9.1%	20.0%	.	9.1%	20.0%	0.0%	.	.
35 - 40th	42.9%	36.4%	42.9%	25.0%	40.0%	40.0%	0.0%	9.1%	14.3%	11.1%	16.7%	0.0%	22.2%	25.0%	16.7%	0.0%	0.0%	.	.	0.0%
40 - 45th	.	20.0%	33.3%	21.4%	37.5%	66.7%	33.3%	27.3%	8.3%	18.2%	0.0%	0.0%	18.2%	0.0%	16.7%	0.0%	0.0%	0.0%	.	14.3%
45 - 50th	14.3%	.	0.0%	18.2%	25.0%	14.3%	20.0%	0.0%	10.0%	25.0%	0.0%	16.7%	12.5%	14.3%	20.0%	16.7%	12.5%	0.0%	0.0%	0.0%
50 - 55th	25.0%	0.0%	0.0%	33.3%	10.0%	8.3%	25.0%	0.0%	14.3%	.	10.0%	0.0%	0.0%	28.6%	12.5%	0.0%	12.5%	0.0%	10.0%	0.0%
55 - 60th	25.0%	20.0%	0.0%	14.3%	8.3%	0.0%	0.0%	.	9.1%	20.0%	9.1%	0.0%	0.0%	0.0%	16.7%	0.0%	16.7%	13.3%	0.0%	.
60 - 65th	.	.	0.0%	33.3%	16.7%	0.0%	7.7%	27.3%	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%	20.0%	.	0.0%	0.0%	0.0%	0.0%
65 - 70th	20.0%	40.0%	.	33.3%	42.9%	0.0%	0.0%	25.0%	0.0%	0.0%	0.0%	18.2%	12.5%	15.4%	0.0%	0.0%	0.0%	0.0%	0.0%	12.5%
70 - 75th	0.0%	0.0%	.	0.0%	0.0%	22.2%	0.0%	0.0%	12.5%	11.1%	20.0%	0.0%	0.0%	0.0%	7.1%	0.0%	0.0%	0.0%	0.0%	7.1%
75 - 80th	.	28.6%	0.0%	.	.	16.7%	.	0.0%	12.5%	9.1%	0.0%	0.0%	0.0%	10.0%	12.5%	0.0%	8.3%	0.0%	0.0%	0.0%
80 - 85th	.	14.3%	12.5%	14.3%	0.0%	14.3%	.	10.0%	0.0%	7.7%	0.0%	0.0%	0.0%	18.2%	0.0%	15.4%	16.7%	0.0%	8.3%	0.0%
85 - 90th	.	.	.	.	.	11.1%	10.0%	11.1%	0.0%	0.0%	0.0%	30.0%	0.0%	0.0%	5.6%	0.0%	0.0%	18.8%	7.7%	0.0%
90 - 95th	.	0.0%	.	0.0%	0.0%	0.0%	12.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	9.1%	0.0%	0.0%	0.0%	7.1%	0.0%	0.0%
95 - 100th	.	.	0.0%	.	0.0%	.	0.0%	11.1%	0.0%	14.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	6.3%	0.0%	0.0%	5.3%

NOTES: 1) Cells are shaded based on values. Green indicates values close to the lowest delinquent frequency, yellow indicates values close to the median delinquent frequency, and red indicates values close to the highest delinquent frequency. Gray values indicate cells where there were fewer than 5 loans. 2) Percentiles are based on the population of originated loans. 2) 283 originated loans with a missing FICO score were excluded from the frequency table.

5.2 Inclusiveness

In addition to the evidence discussed above about the potential value of cash-flow data to identify creditworthy borrowers among applicants with lower traditional credit scores, we found some additional evidence that the use of cash-flow data in highly automated underwriting systems is expanding access to credit to consumers and small businesses that may have faced historical constraints. While as discussed above we were not able to apply a single consistent quantitative analysis across all participants due to data limitations and other factors, we applied a range of analyses where data permitted.

For three of the participants we were able to study the percentage of borrowers who had no or low traditional credit scores. This group is likely to include relatively high numbers of "no file" and "thin file" borrowers, as well as borrowers may be having some difficulty accessing credit after past periods of financial instability.[39] The percentage of the three participants' borrowers with traditional scores below approximately 650 was roughly 45 percent to 50 percent and the percentage of their borrowers below approximately 600 ranged from 0 to 25 percent. In addition, two participants reported that attempts to pull traditional scores for 3.5 percent and 8 percent of their borrowers were unsuccessful, respectively. They were also able to provide data on the number of open accounts reflected in borrowers' traditional credit reports where available, though that does not define which borrowers would be considered to have a "thin file."[40] For one participant, about 8 percent of borrowers had less than three trade lines; for the other, more than 50 percent had less than two open accounts.

For three participants, we were able to assess some borrower characteristics relative to the zip codes in which they reside. For example, we assessed the percentage of borrowers who live in zip codes in which racial minorities exceed 50 percent or 80 percent of the total population as measured by the 2017 American Community Survey. Such zip codes often tend to be served by fewer financial institutions than other zip codes, so access to affordable credit may be a concern in such areas for all residents. All three participants served substantial populations in such zip codes, with 28 percent to 64 percent of their borrowers residing in "majority minority" zip codes and 8 percent to 29 percent in "predominantly minority" zip codes, respectively.

We were also able to assess borrowers' income relative to the average income for their zip codes for two of the participants that provided loan level data. These analyses evaluated how many individual borrowers' incomes fell below the median household income as reported in the Census for their respective zip codes. We found that 59 percent of borrowers for the one participant and 83 percent of borrowers for the other earned less than the median income for their geographies. However, the results should be interpreted with substantial caution because the income metrics provided by the participants may differ from the Census benchmarks accordingly, they may tend to underestimate borrowers' actual income levels. The participant that provided its internal analysis performed a different type of evaluation that measured the size of the difference between the median incomes of its borrowers relative to the median incomes for the zip codes in which they reside. The company concluded that its borrowers' weighted median incomes were 47 percent of the weighted median household income of their geographies.

39 However, it should not be assumed that all "no score" borrowers lack credit files. Due to differences in populations covered by the three NCRAs some borrowers may lack a credit file with one company but still be scoreable by others. There are also differences in scoring thresholds and coverage among third-party scoring models. See *supra* note 30.

40 Each third-party credit scoring system has its own definitions for what renders a credit file too limited to generate a reliable score. Factors could include trade lines that are too new to contain sufficient payment history or files that are too stale due to no recent reported activity. CFPB Credit Invisibles at 4.

Finally, with regard to the two CDFI participants, it is worth noting that such financial institutions must direct at least 60 percent of their financial activities toward one or more target markets, which are defined to include various types of underserved populations and residents of distressed communities, in order to obtain and maintain certification from the CDFI Fund within the U.S. Department of Treasury. Certified CDFIs are required to report annually with regard to demographic groups served, geographies served, and various other types of community development impacts.[41]

5.3 Fair lending effects

For the four participants that provided data sufficient to perform an analysis based on subpopulations, we found that the degree to which the cash-flow data were predictive of credit risk appeared to be relatively consistent across different demographic groups. Rather than proxying for race and ethnicity or gender, the use of the cash-flow variables and scores appeared to provide independent predictive value across all groups. Moreover, when compared to traditional credit scores, the cash-flow based metrics appeared to predict creditworthiness within the race/ethnicity subpopulations at least as well as the traditional scores, and better in selected cases. These results suggest that use of cash-flow variables and scores does not create a disparate impact among protected populations.

More specifically, for all four participants, we were able to calculate the AUCs for likely white borrowers and compared them to the AUCs for borrowers who likely belong to other demographic groups. We were also able to calculate AUCs for likely male and likely female borrowers for three participants. The AUCs for the different demographic groups all indicated that the cash-flow variables and scores were predictive of credit risk and loan performance. In addition, the cash-flow AUCs for the various demographic groups generally showed relatively small amounts of variance from each other. (For instance, the AUCs for likely African-American borrowers did not vary substantially from the AUCs for likely non-Hispanic white borrowers, or the AUCs for female borrowers vs. male borrowers.) Further, when compared to the AUCs for traditional credit scores, the AUCs for the cash-flow based metrics alone and the combined metrics appeared to predict credit worthiness within the subpopulations at least as well as the traditional scores, and better in selected cases.

This relative consistency suggests that the cash-flow models are not simply proxies for race/ethnicity or gender among the sample populations. Rather, they appear to have independent predictive power and to rank order credit risk to a similar degree within each demographic group, respectively. While we were not able to perform all of the statistical analyses that would typically be conducted for a full compliance evaluation of algorithms for credit scoring, eligibility determinations, or pricing, these results are encouraging in that they suggest that the cash-flow variables are providing similar amounts of predictiveness for each demographic group analyzed.

41 *See, e.g.*, CDFI Fund, CDFI Fund Annual Certification and Data Collection Report Form Instructions (2019).

6. CONCLUSION

Our research suggests that cash-flow data can provide meaningful predictive power among populations and products similar to those studied. While the data may be particularly valuable in situations in which traditional credit history is not available or reliable, the data may also provide insights when used in conjunction with traditional scores and metrics.

The cash-flow metrics generally performed as well as the traditional metrics standing alone, and frequently improved the ability to predict credit risk among borrowers that are scored by traditional systems as presenting similar risks of default. Although data limitations did not permit a consistent quantitative analysis to be applied across all participants, we also found evidence that each of the research participants is serving borrowers who may have historically faced constraints on their ability to access credit as evidenced by their traditional credit scores (or lack thereof) and other metrics.

Finally, we found that the degree to which the cash-flow data are predictive of credit risk appears to be relatively consistent across different demographic groups. Rather than creating a disparate impact by proxying for race/ethnicity or gender, the use of the cash-flow variables and scores appears to be providing independent predictive value across all groups.

One should be cautious in extrapolating these results beyond the parameters discussed above, since we lacked the data to conduct certain additional analyses with regard to the covered products and populations, as well as to study the use of cash-flow metrics in credit pricing, performance in different economic conditions, and predictiveness with regard to underwriting longer-term, larger balance loans. Particularly because new underwriting models using cash-flow data have not yet been tested in economic downturns, additional validation would be helpful.

Nevertheless, we view the results as generally encouraging and as suggesting that investment of additional resources is warranted into research and other efforts to reduce the technological, competitive, and compliance challenges that are slowing adoption of beneficial practices and mitigation of risks in today's market. On balance, the results suggest that cash-flow metrics when used alone or in combination with more traditional credit reports and scoring models hold substantial promise for improving credit risk prediction, expanding access to credit, and spurring market innovation and competition.

BIBLIOGRAPHY

Stephania Albanesi et al., Credit Growth and the Financial Crisis: A New Narrative, National Bureau of Economic Research Working Paper No. 23740 (August 2017)

AnnaMaria Andriotis, Why Your FICO Score Could Get a Boost in 2019, Wall St. J. (Oct. 21, 2018)

Robert B. Avery, Kenneth P. Brevoort, & Glenn B. Canner, Does Credit Scoring Produce a Disparate Impact? 40 Real Estate Economics S65 (2012)

Solon Barocas & Andrew D. Selbst, Big Data's Disparate Impact, 104 Cal. L. Rev. 671 (2016)

Robert Bartlett et al., Consumer-Lending Discrimination in the FinTech Era, National Bureau of Economic Research Working Paper No. 25943 (updated June 2019)

Rachael Beer et al., Are Income and Credit Scores Highly Correlated?, FEDS Notes (Aug. 13, 2018)

Tobias Berg et al., On the Rise of the FinTechs—Credit Scoring Using Digital Footprints, FDIC Working Paper 2018-04 (September 2018)

Allen N. Berger & W. Scott Frame, Small Business Credit Scoring and Credit Availability, 47 J. of Small Business Management 5 (2007)

Allen N. Berger et al., The Surprising Use of Credit Scoring in Small Business Lending by Community Banks and the Attendant Effects on Credit Availability and Risk, Federal Reserve Bank of Atlanta Working Paper 2009-9 (March 2009)

Board of Governors of the Federal Reserve System, Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit (2007)

Board of Governors of the Federal Reserve & Office of the Comptroller of the Currency, Supervisory Guidance on Model Risk Management, SR 11-7 & OCC Bulletin 2011-12 (April 4, 2011)

Sterling A. Bone et al., Shaping Small Business Lending Policy Through Matched-Pair Mystery Shopping, 38 J. of Public Policy & Marketing 391 (2019)

Peter Carroll & Ben Hoffman, Financing Small Businesses: How 'New-Form Lending' Will Reshape Banks' Small Business Strategies, Oliver Wyman (2013)

Peter Carroll & Saba Rehmani, Point of View: Alternative Data and the Unbanked, Oliver Wyman (2017)

CDFI Fund, CDFI Certification: Your Gateway to the CDFI Community (2016), available at www.cdfifund.gov/Documents/CDFI_CERTIFICATION_updatedJAN2016.pdf.

CDFI Fund, CDFI Fund Annual Certification and Data Collection Report Form Instructions (2019)

Congressional Research Service, Marketplace Lending: Fintech in Consumer and Small-Business Lending (2018)

Consumer Finanical Protection Bureau, Examination Procedures: ECOA Baseline Review (2019)

Consumer Financial Protection Bureau, Key Dimensions and Processes in the U.S. Credit Reporting System: A Review of How the Nation's Largest Credit Bureaus Manage Consumer Data (2012)

Consumer Financial Protection Bureau, List of Consumer Reporting Companies (2019)

Consumer Financial Protection Bureau, Statement of the Bureau of Consumer Financial Protection on Enactment of S.J. Res. 57 (May 21, 2018)

Consumer Financial Protection Bureau, Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity: A Methodology and Assessment (2014)

Consumer Financial Protection Bureau Office of Research, Data Point: Credit Invisibles (2015)

Marco Di Maggio & Vincent W. Yao, Fintech Borrowers: Lax-Screening or Cream Skimming (updated February 2019), available at abfer.org/media/abfer-events-2019/annual-conference/real-estate/AC19P6018_Fintech_Borrowers_Lax-Screening_or_Cream-Skimming.pdf

Carol A. Evans, Keeping Fintech Fair: Thinking About Fair Lending and UDAP Risks, Consumer Compliance Outlook 4-9 (2nd Issue 2017)

Experian, Alternative Credit Data (2018)

Federal Deposit Insurance Corporation, 2017 FDIC National Survey of Unbanked and Underbanked Households (2018)

Federal Reserve Bank of New York, 2016 Small Business Credit Survey Report on Startup Firms (2017)

Federal Reserve Banks, 2018 Small Business Credit Survey Report on Employer Firms (2019)

Federal Reserve Banks, 2017 Small Business Credit Survey Report on Nonemployer Firms (2018)

Federal Trade Commission, Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues (2016)

Federal Trade Commission, Report to Congress under Section 319 of the Fair and Accurate Credit Transactions Act of 2003 (2012)

Patrice Ficklin, Blog, Preventing Illegal Discrimination in Auto Lending, Consumer Financial Protection Bureau (Nov. 4, 2013), available at www.consumerfinance.gov/about-us/blog/preventing-illegal-discrimination-in-auto-lending/

Elaine Fortowsky & Michael LaCour-Little, Credit Scoring and Disparate Impact, Wells Fargo Home Mortgage Working Paper (December 2001)

Gail Gardner, What Are the Credit Reporting Agencies for Businesses?, Small Business Trends (Jan. 4, 2019), available at smallbiztrends.com/2019/01/business-credit-reporting-agencies.html

Susan Henson, Blog, Introducing Experian Boost, a New Way to Instantly Improve Your Credit Score, Experian (Dec. 18, 2018, updated April 18, 2019), available at www.experian.com/blogs/ask-experian/introducing-experian-boost/

Rajkamal Iyer et al., Screening Peers Softly: Inferring the Quality of Small Borrowers, 62 Management Science 1554 (2016)

Julapa Jagtiani & Catharine Lemieux, Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks?, Journal of Economics & Business (November-December 2018)

Julapa Jagtiani & Catharine Lemieux, The Roles of Alternative Data and Machine Learning In Fintech Lending: Evidence from the LendingClub Consumer Platform, Federal Reserve Bank of Philadelphia Working Paper 18-15 (updated January 2019)

Karen G. Mills, Fintech, Small Business & the American Dream: How Technology Is Transforming Lending and Shaping a New Era of Small Business Opportunity [eBook] (2019)

National Consumer Law Center, Past Imperfect: How Credit Scores and Other Analytics "Bake In" and Perpetuate Past Discrimination (2016)

Office of the Comptroller of the Currency, Examination Guidance on Credit Scoring Models, OCC Bulletin 97-24 (May 20, 1997)

Lisa Rice & Deidre Swesnik, Discriminatory Effects of Credit Scoring on Communities of Color, 46 Suffolk L. Rev. 935 (2013)

Stephen L. Ross & John Yinger, The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement (2002)

Peter Rudegeair, A $150,000 Small Business Loan—From an App, Wall St. J. (Dec. 28, 2018)

David Skanderson & Dubravka Ritter, Fair Lending Analysis of Credit Cards, Federal Reserve Bank of Philadelphia Payment Cards Center Discussion/Working Paper 14-02 (August 2014)

TransUnion, The State of Alternative Data 3 (2015)

Claire Tsosie & Steve Nicastro, Business Credit Score 101, nerdwallet (Oct. 6, 2017), available at www.nerdwallet.com/blog/small-business/business-credit-score-basics/

U.S. Department of Housing & Urban Development Office of Policy Development & Research, All Other Things Being Equal: A Paired Testing Study of Mortgage Lending Institutions (2002)

U.S. Department of the Treasury, A Financial System that Creates Economic Opportunities: Nonbank Financials, Fintech, and Innovation (2018)

Yan Zhang, Assessing Fair Lending Risks Using Race/Ethnicity Proxies, 64 Management Science 178 (2018)

APPENDIX

CRA Charles River Associates

The Use of Cash Flow Data in Automated Credit Underwriting

Report Submitted for:
FinRegLab

Prepared By:

Dr. Marsha J. Courchane
Arthur P. Baines
Vice Presidents and Co-Practice Leaders,
Financial Economics Practice,
Charles River Associates[1]

July 23, 2019

---

[1] The authors may be reached by email at mcourchane@crai.com and abaines@crai.com.

Table of Contents

1. EXECUTIVE SUMMARY

1. For this research study, we analyzed the use of various types of cash flow data in highly automated underwriting systems utilized by six financial services institutions which offer and originate consumer and small business loans across a broad set of geographies in the U.S. The use of the various types of cash flow data, in conjunction with, or in lieu of, more traditional credit bureau derived data has been used to underwrite credit for both consumer loans and small business loans.

2. We find compelling evidence that indicates that among the sample populations and products CRA analyzed, the cash flow data are predictive of credit risk and loan performance across the highly heterogeneous set of participants.[2] In our separate analyses of each participant, the results appear to be robust across both consumer and small business populations as well as across the credit spectrum, including among borrowers with no, or very low, traditional credit scores, some of which may reflect 'no-file' or 'thin-file' borrowers. Among the sample populations and products, the cash flow data and traditional credit data, when analyzed, displayed some degree of asymmetric information, and the cash flow data frequently improved the sorting of risk among borrowers posing similar credit risks, as measured by the traditional credit data.

3. Where data were available, we observe customers to have lower incomes, on average, as compared to the geographies in which they reside, and many customers reside in majority minority or predominantly minority geographies, suggesting a sizeable share of the sample populations may include customers who traditionally have been credit constrained. This limited evidence suggests that the participants' use of cash flow data in highly automated

[2] CRA did not conduct an analysis of Participant 3's sample population as loan level data were not made available.

underwriting systems expanded access to credit for consumers and small businesses that may traditionally have found it difficult to access credit markets.[3]

4. For the subset of participants for whom we have proxied data for race/ethnicity and/or gender, we were able to test whether or not the cash flow data were predictive of credit risk among demographically neutralized populations. We found the degree to which the cash flow data were predictive of credit risk to be relatively consistent across multiple demographic groups within the sample populations. The cash flow data, rather than proxying for demographic attributes, appear to predict credit risk within each group in the sample populations. The use of cash flow data in the highly automated underwriting processes represented by the sample populations and products did not appear to create a disparate impact.

2. SCOPE OF ASSIGNMENT

5. FinRegLab engaged Charles River Associates ("CRA") to conduct analyses of the use of cash flow data by participating financial services institutions in highly automated underwriting models of credit applications and loan originations.[4] FinRegLab's intent is to undertake a quantitative analysis of important questions raised by the increased use of cash flow data in the market for consumer and small business loans.[5] Those research questions include:

    A. Are cash flow data useful in predicting credit risk in the underwriting process, as compared with traditional credit scores and/or credit bureau attributes?

[3] The evidence is limited due to data constraints.

[4] We use the term financial services institutions to indicate that the participants offer credit products to consumers and small businesses. The term does not suggest they are depository institutions, and not all of the participants are chartered financial institutions.

[5] We generally use the terms customer, applicant and borrower to include both consumers and small businesses in this context.

B. Does the use of cash flow data expand the availability of credit, particularly with respect to consumers and small businesses that may have experienced constrained access to credit under more traditional underwriting criteria?

C. What, if any, fair lending risks appear to arise from the use of cash flow data in such highly automated underwriting processes?

6. To analyze these questions, FinRegLab identified financial services institutions which had built and implemented highly automated underwriting systems that utilized cash flow data in some measure to assess credit risks and to decision credit applications and solicited their participation in the research study.

7. CRA designed the quantitative research. This included the determination of the types of analyses that would be undertaken and the data that would be required from each participant. CRA also defined and provided the logistical support to enable the complex data transfers, encryption and IT security necessary to ensure customer privacy was maintained throughout the process.

3. METHODOLOGY

3.1. Financial Institution Participants

8. FinRegLab recruited six financial services institutions to participate in this research. The institutions are highly heterogeneous with respect to products offered, geography, types of cash flow data utilized, how such data are used, and the sources of the cash flow data. Further, the participants have different lengths of market participation time, and different volumes of applications underwritten and loans originated. Two of the institutions focus on small business lending, while four focus primarily on direct consumer lending. Two of the institutions are certified Community Development Financial Institutions (“CDFIs”). The participants include five for-profit firms and one non-profit. Several of the participants are nationally based, while others are highly concentrated in selected geographies. All

participants share a mission focus on increasing access to markets they view as traditionally underserved.

9. Each institution has developed proprietary algorithms that utilize cash flow data as a component in their assessment of applicant credit risk. The institutions did not provide their algorithms to CRA, but rather provided individual cash flow metrics and, in some cases, the credit scores created by their proprietary algorithms utilizing cash flow metrics. The nature and sources of the cash flow data differ across institutions. Our ability to provide detailed descriptions of each cash flow attribute is limited by their proprietary nature. Some of the institutions utilize the cash flow data in conjunction with various traditional credit bureau attributes and/or scores, while others do not. Most of the institutions utilize the cash flow data as a component of their primary assessment of credit risk; however at least one institution uses the cash flow data as a component of a 'second-chance' underwriting evaluation. Each of the institutions has deployed their proprietary algorithms to originate loans in the marketplace.

10. Each institution takes a unique approach to the use of cash flow data. Each institution has invested significant resources to identify and test various relationships among cash flow data and other factors that impact credit risk. Each participant has provided to CRA a description of extensive model development efforts meant to establish relationships they believe to be robust and predictive. Thus, we have the advantage of testing relationships the participants believe to exist, rather than simply theorizing about a potential set of relationships that may exist. We are able to test cash flow based scores, derived from a number of underlying cash flow metrics, as well as individual cash flow metrics. The relative breadth of lending products offered by the participants allows us to analyze the use of cash flow data on products with varying durations across a diversity of customer-types. The participants have, for the most part, utilized their models in the marketplace for some time, and most have relatively robust information regarding the actual performance of loans originated using the cash flow data in their models to assess credit risk.

11. The heterogeneous nature of the participants does introduce limitations to the potential analyses. For example, the diversity of products and approaches means that the aggregation of data across the institutions is not feasible. The analyses were conducted separately for each institution and those individual analyses are reported in the appendices.[6] While most of the participating institutions have substantial loan volumes, allowing us to undertake statistical testing, our ability to draw conclusions about individual cash flow attributes is more circumscribed, as not all participants utilize the same (or, in some cases, even similar) cash flow attributes. Our ability to utilize the denied applications in our analysis was also limited by the research design, in that there is no performance data for applications that did not result in an originated loan, including approved applicants that chose not to proceed with the loan.[7] These applicant and loan populations, while sizeable, appear not representative of the overall US population. Further, most of the participants began using cash flow attributes to model risk in a period of general economic expansion following the end of the Great Recession.[8] As such, we have limited ability to observe the actual performance of these models in time periods with relatively more adverse economic conditions.

3.2. Data

12. The data requested from each participant included the following:

- Application-level data including credit score measures derived from cash flow data, credit scores derived from traditional bureau attributes, individual cash flow attributes, traditional credit attributes, application status (e.g. approved, declined, etc.), application date, and geography

[6] At the direction of FinRegLab, CRA will not attribute the results of the analyses to specific participants.
[7] It is common in lending markets that some share of approved applications do not result in an originated loan.
[8] Commonly understood to be June 2009; available at: https://www.minneapolisfed.org/publications/special-studies/recession-in-perspective.

- For originated loans, loan-level data on performance of the loan (including default and/or delinquency information)

- Gender proxies and Bayesian Improved Surname Geocoding ("BISG") race and ethnicity probabilities based on the applicant/borrower's surname and geography. BISG probabilities were calculated using assumptions closely mirroring those in the CFPB's publicly-available computer code for calculating BISG probabilities.[9]

13. Most participants provided data on all three dimensions. CRA worked with each participant to refine the data request based upon the specific policies and procedures of each institution. This included identifying those attributes which each institution defined to be cash flow metrics, and those they believed important in their underwriting process. As such, there is an inherently broad definition of the metrics considered to be "cash flow." We worked with each participant to identify performance metrics that were objective and not subject to discretion or judgment. As a result we may be testing performance metrics that differ from a participant's internal performance metrics and those upon which their proprietary algorithms were tested and developed. It is important to understand that not every requested data element was used by each participant in their own underwriting process. Zip code data, for example, were provided by participants that do not use that data as part of their automated underwriting process. Care was taken to assess the validity and completeness of the provided data. Populations generally were defined by time period and were not subjectively selected by the financial institution. Basic diagnostics are reported in the respective Appendices for each participant. Finally, CRA validated that the BISG probabilities were constructed in a manner not materially different from the assumptions reflected in the CFPB's publicly-available computer code for creating BISG probabilities.

[9] Generally, BISG probabilities were calculated by the financial institution, negating the need to provide personally identifiable information ("PII") associated with the applicants and borrowers to CRA.

3.3. Analytical Approaches

14. CRA defined an analytical approach specific to each of the research questions described above.

15. First, we assessed the degree to which the evaluation of credit risk was facilitated through the use of cash flow data. This is commonly referred to as "lift" with respect to default risk modeling. This lift is not one-sided. It does not imply only increasing credit scores. Rather, lift implies movement in both directions: increasing the credit scores of those who are, in fact, more likely to repay, and decreasing the scores of those who are, in fact, less likely to repay. For this reason, it may be more intuitive to conceptualize this exercise as rank ordering risk from highest risk borrowers to lowest risk borrowers. Of particular interest is whether cash flow data can be used to accurately evaluate credit risk for customers for whom a traditional credit score does not exist or for whom the credit score is based on relatively little market experience, such as for those with a 'thin' credit file.

16. For this purpose, we utilized the loan-level performance data. First we assessed the degree to which correlation(s) were observed between the known set of defaulted and non-defaulted accounts, the individual cash flow attributes, and the institution's proprietary credit scores which were derived from the cash flow attributes.[10] Next, we developed a series of multivariate logit models to ascertain the relationship between the cash flow attributes and scores and the probability of default. Finally, we computed the receiver operating characteristics ("ROC") and the area under the ROC curve ("AUC"). These metrics

[10] Throughout this report the term statistically significant should be understood to be based upon a 95% confidence level, unless otherwise stated.

are commonly used to understand the degree to which an attribute 'predicts' likelihood of default.[11]

17. The second research question is somewhat more subjective. While each participant expressly indicated a focus on meeting the needs of traditionally underserved or access-constrained customers, we reviewed the available data for empirical evidence to suggest whether the customers of these participants have attributes that may correlate with consumers or markets that are commonly viewed as underserved or access constrained. Where possible, we utilized credit scores derived from traditional credit bureau attributes as a proxy for the degree to which access may previously have been constrained. Additionally, we have used various publicly available metrics for the geographies associated with the customer-level application and loan data to describe the customers receiving the products. These metrics include median income and majority minority geography status.

18. While these questions allow for an analysis of the potential benefits of cash flow data for the evaluation of credit risk, the final question focuses on an important risk inherent in every underwriting process – fair lending risk. The highly automated processes by which the cash flow attributes and associated credit scores are derived dictates a focus on disparate impact ("DI") risk, rather than disparate treatment risk.[12] Under disparate impact theory, an objective policy or factor, applied uniformly and without judgment or discretion, may create disparate outcomes (e.g. differences in average credit scores, average denial rates or average prices) on a prohibited basis. The most common prohibited bases evaluated by fair lending examiners include race, ethnicity, age, or gender. Where

[11] See, for example, Bowen Baker, "Consumer Credit Risk Modeling," MIT Departments of Physics and EECS, 70 Amherst Street, Cambridge, MA 02142, December 17, 2015. The ROC plots the true positive rate (TPR) relative to the false positive rate (FPR) for a given probability cutoff such that a completely random predictor will produce a straight line from (0, 0) to (1, 1) with an AUC of 0.5. A perfect predictor will produce a square ROC with an AUC of 1.

[12] It was beyond the scope of this project to validate that the cash flow metrics and associated credit scores would be viewed by federal financial regulatory agencies as empirically derived and statistically sound ('EDSS') under Regulation B and prudential guidance.

disparate outcomes are caused by the objective policy or factor, the empirical analysis focuses on the business necessity (or justification) for the use of such a factor. For this research study, the business necessity includes the accurate prediction of credit risk default probabilities. We have undertaken analyses that attempt to discern whether the cash flow attributes or derived scores predict credit risk or may be serving as a proxy for one or more of the prohibited basis groups. The techniques for analyzing this question were developed over the past two decades and have been tested on attributes sourced from traditional credit bureau data on populations where race, ethnicity, age and gender were known.[13]

19. We have employed similar analytical techniques here, which require dividing the sample populations into demographic groups, but with the important caveat that we had to proxy for race, ethnicity and gender because they are unknown for the populations in this analysis.[14] Using proxies, we isolated sub-populations with a relatively high likelihood of belonging to a given race, ethnicity or gender group.[15] Within each group, we then applied similar analytical techniques to those used to answer the credit evaluation question.[16] By restricting the tests to analyses within prohibited basis groups, we are measuring the degree to which these attributes can be used to evaluate credit risk among a group of customers belonging to the same race, ethnicity or gender.

---

[13] See Avery, Brevoort, Canner "Does Credit Scoring Produce a Disparate Impact?" *Real Estate Economics*, Vol. 40, Issue S1, December 2012, S65 – S114.

[14] Financial Institutions are generally prohibited from collecting demographic information on prohibited status with the notable exception of mortgage activity reportable under the Home Mortgage Disclosure Act ("HMDA").

[15] We adopt the threshold approach using the BISG probabilities. If a consumer has an estimated BISG probability of 75% or more, we consider them likely to belong to a particular subgroup. While CFPB research has asserted that the continuous approach, which assigns to each individual a vector of probabilities for each race/ethnicity, may be more accurate in determining the total count of each demographic subgroup across a nationally representative population, for the analysis performed in this report we believed the threshold approach was more useful because it provides greater confidence that the borrowers designated as likely belonging to a given subgroup are, in fact, members of the subgroup. One could use other thresholds than 75%, but we considered that a higher threshold might further decrease population size and lower thresholds might blur the DI measures.

[16] Avery, Brevoort and Canner refer to the within group tests as estimating the model in demographically neutral environments.

20. See appendix G for a glossary of technical terms.

3.4. Use of Proxies

21. This analysis utilizes BISG to develop race and ethnicity proxies. Gender proxies were generally provided by the participants, and the underlying approaches utilized the applicant's or borrower's first name in combination with data from the Census Bureau.[17] We believe these proxies to be useful for this type of testing, and we observe these approaches to be commonly used and accepted by federal financial regulatory agencies, including, for example, the CFPB.[18] The use of such proxies, however, is not without limitations and necessitates cautious interpretation of the results. A relatively small but growing body of academic research finds that the use of the proxies can be accompanied by sizeable measurement errors.[19] In certain circumstances, the proxies are subject to substantial Type 1 and Type 2 errors. Specifically, the proxies fail to identify properly actual members of each group (or assign a very low probability of belonging to a group, when the person belongs to the group), and incorrectly assign individuals to the wrong group (or assign a high probability of belonging to the wrong group).

4. FINDINGS

22. Below we report the findings for each participant. Due to the proprietary nature of the algorithms developed by the participants and the resulting cash flow metrics, we describe the cash flow metrics in broadly generic categories. It is important to understand that we are not evaluating the predictiveness of each participant's overall underwriting process. All of the participants' respective automated underwriting processes utilize additional information and attributes beyond the cash flow data. We have isolated the cash flow

17 There are numerous commercial software packages available to create gender proxies.

18 The CFPB has made public the computer code it uses to calculate BISG probabilities, and it is available at: https://github.com/cfpb/proxy-methodology.

19 Zhang, "Assessing Fair Lending Risks Using Race/Ethnicity Proxies," *Management Science*, Vol 64, Issue 1, Jan. 2018. https://doi.org/10.1287/mnsc.2016.2579, Published Online, November 17, 2016.

metrics and/or scores from that overall process. As such, these results should not be interpreted as reflecting each participant's overall ability to model credit risk.

23. Care should be taken in making comparisons of the results across the participants. The heterogeneous nature of the participants, as discussed above, makes such comparisons potentially misleading.

4.1. Participant #1

24. Participant 1's automated underwriting process uses a series of cash flow metrics derived from the applicant's transactional history via proprietary algorithms. The algorithms are applied to several recent months of account transactions and used to calculate cash flow metrics related to income, expenses, balances and activity levels. Participant 1 provided to CRA a loan-level data file containing twenty-four cash flow metrics for each of 10,957 originated loans, as well as the source of the transaction data from which the applicant's transaction history was analyzed, the geography of the borrower, and a delinquency indicator. See Appendix A, Table 1 for basic diagnostics on the data provided.

25. We found compelling evidence that the cash flow metrics are correlated with the likelihood of default in the sample population. We separated the borrowers into delinquent and non-delinquent populations and performed a difference in means test between the two groups on each of the cash flow metrics. Sixteen of the 24 provided cash flow metrics were observed to have statistically significant differences among the delinquent as compared to non-delinquent borrowers. See Appendix A, Table 2 for the test results.

26. Next, we estimated several logit models of the likelihood of delinquency and calculated AUCs based on each. In the first model, we included as predictors the five cash flow variables identified by Participant 1 as among the most important in their underwriting process. In the second model, we included as predictors the cash flow metrics found to have statistically significant differences in means among delinquent borrowers as compared to non-delinquent borrowers. In the third model, we included all of the cash flow metrics as predictors.

27. The AUCs obtained were .597, .713, and .725 for models 1 through 3, respectively. See Appendix A, Tables 3 and 4 and Chart 1 for complete model results. These AUCs meaningfully diverge from .5 (which would indicate no predictive power) and are at levels which, in our experience, suggest a relatively robust ability to predict likelihood of default within the test sample.

28. Our ability to evaluate Participant 1 with respect to the question of the possible expansion of credit access was constrained by the available data. We were not able to examine traditional score ranges, number of trade lines, length of time on bureau or other attributes frequently used to identify consumers or markets with potentially less access to credit.

29. The data included zip code and a proxy for income, which allowed us to make some potential inferences as to the demographics of customers obtaining credit from Participant 1. Approximately 64% of the loans in the sample population were made to customers residing in a majority minority zip code, based upon data from the 2017 American Community Survey ("ACS") (see Appendix A, Table 5).[20] Approximately 29% of the loans were made to customers residing in predominantly minority zip codes, based upon data from the 2017 ACS (see Appendix A, Table 6).[21] Such metrics are difficult to put into context. Nonetheless, these shares suggest a relatively high level of minority customers seeking and gaining access to the product offered by Participant 1. We also report (see Appendix A, Tables 5 and 6) the shares of delinquent and non-delinquent customers by majority minority zip code and by predominantly minority zip code. While we do not observe a difference in delinquency rates among customers residing in majority minority zip code as compared to those not residing in such zip codes, a slightly higher delinquency rate is observed among customers residing in predominantly minority zip codes as compared to

---

[20] Majority minority zip codes are those in which the minority population exceeds the non-Hispanic white population. That is, less than 50% of the residents in the zip code are reported as non-Hispanic white, based upon the 2017 ACS.

[21] Predominantly minority zip codes are those in which the minority population exceeds 80% of the total population of the zip code, based upon the 2017 ACS.

those not residing in such zip codes. These are raw delinquency rates, uncontrolled for any differences in customers' creditworthiness.

30. We compared the income proxy available for each customer to the median household income of the zip code in which each customer resides. The income proxy is based upon Participant 1's proprietary algorithm and is calculated without the application of judgment; however it is not directly comparable to the zip code level household income reported by the US Census bureau.[22] This may lead to a downward bias in the income proxy, and it likely underestimates, on average, customers' actual income levels. Thus, the observation that approximately 83% of the customers have incomes at or below the median income of the zip code in which they reside should be interpreted with caution (see Appendix A, Table 7).

31. We could not evaluate disparate impact risk for Participant 1 as demographic attributes were unavailable.

4.2. Participant #2

32. Participant 2's automated underwriting process uses a cash flow score ("CFS") derived from the applicant's transactional history via proprietary algorithms. Participant 2 provided to CRA a transaction-level data file containing 212,949 applications, which resulted in 40,911 originated loans. Where available, they provided their proprietary CFS, a traditional credit score, as well as a delinquency indicator. See Appendix B, Tables 1 and 3 for basic diagnostics on the data provided.

33. We found compelling evidence that the CFS is correlated with likelihood of delinquency in the sample population. We separated the borrowers into delinquent and non-delinquent populations and performed a difference in means test between the two groups on the CFS.

[22] A detailed explanation of the method utilized to calculate the income proxy is not possible without unmasking the participant.

The cash flow score was statistically significantly lower for those loans that went delinquent. See Appendix B, Table 4 for the test results. To further understand the relationship between the loan performance, the CFS, and the traditional credit score, we divided the loans into twenty groups from lowest to highest CFS and FICO scores, and calculated the delinquency rate within each group. The resulting 'heat map' is reported in Appendix B, Table 6. As expected, the observed delinquency rates were higher among those areas of the heat map which represent relatively higher credit risk. Each row of the heat map provides a visualization of the CFS's ability to separate risk among a group of customers with a similar level of credit risk based upon the traditional credit score. A clear pattern is observed in the rows whereby the customers on the left most columns have relatively high delinquency frequency relative to the customers in the right hand columns, notwithstanding that all customers in the row have a similar credit risk as measured by the traditional score. Each column shows the traditional credit score's ability to separate risk among a group of customers with a similar level of credit risk based upon the CFS.

34. Next, we estimated three logit models of the likelihood of delinquency and calculated AUCs based on each. In the first model, we included a control for the traditional credit bureau score only. In the second model, we included only a control for the CFS, and in the third model we included controls for both the traditional credit score and the CFS. The AUCs obtained were .640, .652, and .660 for models 1 through 3, respectively. See Appendix B, Table 5 and Chart 1 for complete model results. These AUCs meaningfully diverge from .5 (which would indicate no predictive power) and are at levels which, in our experience, suggest a relatively robust ability to predict the likelihood of delinquency within the sample population. The cash flow score and traditional score have similar AUCs. The results suggest that among the sample populations, the CFS adds incremental ability to sort credit risk, beyond that contained in the traditional credit score.

35. The average credit score for Participant 2's customers was 660, with 44% having a score below 650, and 16% having a score under 600. This suggests that Participant 2 lends to borrowers who might struggle to qualify for loans using a traditional score.

36. The data included zip code which allowed us to make some potential inferences as to the demographics of customers obtaining credit from Participant 2. Approximately 28% of the loans in the sample population were made to customers residing in a majority minority zip code. This zip code level demographic information is based upon data from the 2017 American Community Survey ("ACS") (see Appendix B, Table 7).[23] Approximately 8% of the loans were made to customers residing in predominantly minority zip codes. (See Appendix B, Table 8).[24] We also report (see Appendix B, Tables 7 and 8) the shares of delinquent and non-delinquent customers by majority minority zip code and by predominantly minority zip code. A higher delinquency rate was observed among customers residing in predominantly minority or majority minority zip codes as compared to those not residing in such zip codes. These are raw delinquency rates uncontrolled for any differences in customers' creditworthiness.[25]

37. With regard to fair lending risk, the evidence suggests that the use of the CFS did not create a disparate impact among the sample population. The BISG probabilities were used to identify separate groups of borrowers with a high likelihood of belonging to each race/ethnicity group. Gender proxies were used to identify separate groups of borrowers with high likelihood of belonging to each gender group. First, we divided the not past due and past due populations into demographically neutralized sub-populations and tested the difference in means within each race/ethnicity group and by gender. The cash flow score demonstrates statistically significant difference between past due and not past due loans among all tested groups in the sample population. The same is true with respect to the traditional credit score. (See Appendix B, Table 10.)

---

23 Majority minority zip codes are those in which the minority population exceeds the non-Hispanic white population. That is, less than 50% of the residents in the zip code are reported as non-Hispanic white, based upon the 2017 ACS.

24 Predominantly minority zip codes are those in which the minority population exceeds 80% of the total population of the zip code, based upon the 2017 ACS.

25 The subsequent analyses control for credit within demographically neutralized groups.

38. Next, we tested the ability of the three models to rank order risk in the demographically neutralized sample populations. We obtained an AUC of .651 when testing the CFS's ability to rank order credit risk among a group of highly likely non-Hispanic white borrowers (from Model 3). This compares to AUCs of .638, .640, and .633 for likely African American, Hispanic and Asian borrowers, respectively. See Appendix B, Table 11. We repeated this process with respect to gender and obtained AUCs of .657 and .644 for male and female borrowers, respectively. The consistency of the AUCs across these demographically neutralized samples is encouraging, and indicates that it is unlikely that the three cash flow models were simply proxies for race/ethnicity or gender. Rather, they rank ordered risk within demographic groups with relatively equal effectiveness within the sample population. See Appendix B, Tables 12-17 for the full model output for each logistic regression.

4.3. Participant #3

39. Participant 3's automated underwriting process uses several cash flow metrics derived from measures of the applicant's income, debt and expenses. Their algorithm estimates a cash flow metric score ("CFMS") to predict delinquency, which does not consider the customer's traditional credit history. Thus, it is our understanding that two applicants with the same cash flow metrics would have the same CFMS regardless of differences in previous access to credit, delinquencies or defaults and homeownership status. Participant 3 provided to CRA a summary-level analysis of a sample population in excess of 20,000 loans.[26]

40. In this section, we report findings from Participant 3's internally generated summary analysis, which Participant 3 attests to be accurate. We note that CRA did not have the ability to verify the analyses, as loan level data were not made available to us.

[26] The loan count is the minimum loan count implied by the summary analysis provided by the Participant. It, as well as the other loan counts reported by the participant, should not be interpreted as a rounded version of the total loan count.

41. With respect to the evaluation of credit risk, Participant 3 divided the sample population into decile groups from lowest to highest score and reported the average rate at which loans went into delinquency in each group.[27] A linear relationship was observed across the first nine deciles, as the delinquency rate declines at a relatively consistent rate as the score deciles increase, with higher ability to repay. The relationship inverts in the last decile (highest ability to repay) and the delinquency rate is observed to be higher among this group as compared to the 9th decile. Notwithstanding the linear relationship observed across the average delinquency rates by decile, Participant 3 reported an AUC of .532 when assessing the CFMS's correlation with delinquency within the sample population. (See Appendix C, Table 1.)[28] Participant 3 reported that the AUC differs from .5 with statistical significance at the 95% confidence level. It remains difficult to conclude that these AUCs meaningfully diverged from .5 and that CFMS had a robust ability to predict delinquency within the sample population.

42. This process was repeated using debt to income ("DTI").[29] An AUC of .513 was reported for DTI's ability to rank order credit risk, and Participant 3 reports that it differs from .5 with statistical significance at the 95% confidence level. (See Appendix C, Chart 1.) The reported statistical tests confirm that the CFMS displayed a stronger correlation with delinquency as compared with DTI alone among the sample population.

43. Participant 3 divided the sample population into two groups: FICO valid customers[30] and FICO invalid customers, and both groups are reported to contain more than 10,000 observations.[31] The analyses described above were replicated on both the FICO valid and

---

27 Delinquency is defined by Participant 3 to be 60+ days delinquent or when the loan is charged off, rewritten, or where the borrower has filed bankruptcy in first 12 months subsequent to loan origination.

28 All of the Tables in Appendix C were created by Participant 3, and CRA was unable to validate the content.

29 DTI was calculated using a subset of the factors utilized in the CFMS.

30 FICO-valid customers are those with FICO scores between 300 and 850.

31 FICO-invalid customers are those with FICO scores <300, >850, or missing.

invalid groups. AUCs of .523 and .537 were reported for the FICO valid and invalid groups, respectively, based upon the CFMS. Participant 3 reports these to differ from .5 with statistical significance at the 95% confidence level. Similarly, Participant 3 reported that AUCs of .508 and .507 for the FICO valid and invalid groups, respectively, based upon the ability of the DTI measure to rank order credit risk, differed from .5 with statistical significance at the 95% confidence level. See Appendix C, Chart 2 and 3. The reported statistical tests confirm that the CFMS displayed a stronger correlation with delinquency as compared with DTI alone for both subgroups in the sample population. We note that the statistical tests suggest the CFMS had a slightly stronger correlation with delinquency among the FICO invalid group as compared to the FICO valid group. Regardless of the statistical significance asserted, it is difficult to conclude that these AUCs meaningfully diverged from .5 and that CFMS had a robust ability to predict delinquency within either sub group in the sample population.

44. With respect to credit expansion, Participant 3's summary analysis is useful in demonstrating that they were able to extend credit to large numbers of customers with either no traditional credit score or very low credit scores. Additionally, Participant 3 reported the weighted median income of their customers to be 47% of the weighted median household income of the zip codes in which they reside.[32] (See Appendix C, Chart 4.) While more customer attributes would be helpful, these FICO scores and income comparisons are consistent with a population of customers that may be challenged in accessing traditional sources of credit.

45. We could not evaluate disparate impact risk for Participant 3 as demographic attributes and loan-level data were unavailable.

32 Median household income in the zip code was based upon the American Community Survey. While CRA used the same median household income in our analysis of other participants, we formulated our calculation differently. Each approach has its merits, but they are not directly comparable.

4.4. Participant #4

46. Participant 4's underwriting process consists of two highly automated steps. The first utilizes traditional credit attributes to build a traditional probability of default, upon which the initial underwriting decision is based. For those applicants that exceed an established probability of default threshold and would otherwise be declined, the applicant is given the option to provide access to their account information for cash flow based underwriting. In this second step, a cash flow based probability of default ("CFPD") score is calculated using proprietary cash flow metrics calculated from the applicant's recent account transaction history. Cash flow metrics used relate to income, expenses, balances and activity levels.

47. Participant 4 provided to CRA a transaction-level data file containing 86,288 applications, which resulted in 25,953 originated loans. Where available, they provided their CFPD score and seven underlying cash flow metrics, a traditional probability of default ("TPD") score, and actual loan performance data, among other data. See Appendix D, Table 1 and Table 3 for basic diagnostics on the application data provided. To better understand the underwriting outcomes, we separated the applicants into approved and declined groups and performed a difference in means test between the two groups on the CFPD score and the individual cash flow metrics. All test results were statistically significant. See Appendix D, Table 2 for the test results.

48. With regard to the rank ordering of credit risk, we found compelling evidence that the cash flow metrics are correlated with likelihood of default among the sample population. We separated the borrowers into defaulted and non-defaulted groups and performed a difference in means test between the two groups on the CFPD score, individual cash flow metrics, TPD scores and other provided attributes.[33] All of the test results were statistically significant, but for one of the non-cash flow attributes. (See Appendix D, Table 4.) To

[33] There were 1,137 loans without a provided default status. These loans were excluded from all analyses of default.

further understand the relationship between the default rates, the CFPD score and the TPD score, we divided the loans into twenty groups from lowest to highest CFPD and TPD scores, and calculated the default rate within each group. The resulting 'heat map' is reported in Appendix D, Table 6. As with the previous heat map, each row is a visual representation of the CFPD's ability to differentiate credit risk among a group of customers with similar level of credit risk as measured by the TPD. Here again, the rows provide evidence that the CFPD score appears to contain incremental ability to sort credit risk after the TPD has been considered.

49. Next, we estimated several logit models and calculated AUCs based on each. In the first model, we included only the TPD score as a predictor. In the second model we included only the CFPD score as a predictor. In the third model, we included both the TPD and CFPD scores as predictors.

50. The AUCs obtained were .559, .592 and .620 for models 1 through 3, respectively (see Appendix D, Table 5 and Chart 1 for complete model results). These AUC values suggest that the CFPD score has a slightly better ability to rank order credit risk, compared to the TPD score. Further, even after the traditional credit attributes have been considered, the cash flow attributes provide incremental ability to rank order credit risk within the sample population.

51. We also expanded our analysis to include other data fields that appeared to be used to develop the TPD and CFPD scores. See Appendix D, Table 5, which reports the results of a logit model of default that includes both the TPD and CFPD score controls and controls for the other fields present in the data (model 4). The TPD and CFPD scores remain statistically significant. Fraud score and the number of accounts are also statistically significant variables. The remaining controls have quite small estimated coefficients. This is evidence that the CFPD and TPD scores are the dominant predictors of default. These controls are likely highly correlated with the TPD and CFPD controls, thus explaining their small coefficients or lack of significance in the combined model. The AUC for model 4 is .650, compared to .620 for the model including only the TPD and CFPD scores, indicating that the

combined model is only slightly better at predicting default than the model including only the two scores.

52. Our ability to evaluate Participant 4 with respect to the question of the possible expansion of credit access was constrained by the available data.

53. With regard to fair lending risk, the evidence suggests that the use of the CFPD score did not create a disparate impact among the sample population. The BISG probabilities were used to identify separate groups of borrowers with a high likelihood of belonging to each race/ethnicity group. Gender probabilities were applied in a similar fashion to identify a group of likely male borrowers and a separate group of likely female borrowers. First, we examined the average values of the key data fields among loans that defaulted and those that did not within each race, ethnicity, and gender group (see Appendix D, Table 8). For almost all target groups, we found statistically significant differences in the average TPD and CFPD scores between loans that defaulted and those that did not.

54. Next, we tested the ability of the TPD and CFPD scores to rank order risk in the demographically neutralized sample populations. We obtained an AUC of .603 when testing the CFPD's ability to rank order credit risk among a group of highly likely non-Hispanic white borrowers. This compares to AUCs of .584, .602 and .583 for likely African American, Hispanic and Asian borrowers, respectively. (See Appendix D, Table 9.) We repeated this process with respect to gender and obtained AUCs of .606 and .584 for male and female borrowers, respectively. See Appendix D, Tables 10 – 17 for the full model output for each logistic regression. The relative consistency of the AUC across these demographically neutralized samples is encouraging, and suggests that the CFPD was unlikely to simply proxy for race/ethnicity or gender, but was able to rank order risk within demographic groups with relatively equal effectiveness within the sample population. We ran the same tests with respect to the TPD score for each of demographically neutralized sample. It is noteworthy that among these sample populations, the CFPD's ability to rank order credit risk appeared to be superior to the TPD's ability to rank order credit risk in every comparison.

4.5. Participant #5

55. Participant 5's automated underwriting process uses a series of cash flow metrics derived from the applicant's account transactional history via proprietary algorithms. The algorithms are applied to several recent months of account transactions and used to calculate cash flow metrics related to income, expenses, balances and activity levels, as well as a pre-qualification cash flow score ("CFS"). Participant 5 provided to CRA a transaction-level data file containing 229,952 applications, which resulted in 8,751 originated loans. Where available, they provided two individual cash flow metrics, their cash flow based score (a pre-qualification probability of default), traditional credit bureau attributes and scores, and the days each loan was past due. See Appendix E, Table 1 and 3 for basic diagnostics on the data provided. To better understand the underwriting outcomes, we separated the applicants into approved and declined groups and performed a difference in means test between the two groups on the CFS and the traditional credit bureau attributes and scores. These test results were statistically significant. See Appendix E, Table 2 for the test results.

56. Among the population provided, only a small proportion are delinquent (180 out of 8,751), so it is difficult to find evidence that the cash flow metrics are correlated with likelihood of default. Even with the small default population, we found the two cash flow metrics, one traditional metric and the Vantage score, to have statistically significant differences between past due and non-past due loans. (See Appendix E, Table 4.) To further understand the relationship between the past due rates, the CFS, and Vantage score, we divided the loans into ten groups from lowest to highest CFS and Vantage scores, and calculated the past due rate within each group. The resulting 'heat map' is reported in Appendix E, Table 7(b). The rows and columns are interpreted in the same manner as the previous heat maps.

57. Next, we estimated three logit models of delinquency and calculated AUCs based on each. In the first model, we included as controls both the Vantage score itself and a control indicating having a Vantage score. In the second model, we included only the cash flow

metric, and in the third model, we included both the Vantage score and the cash flow metrics.

58. The AUCs obtained were .573, .572, and .659 for models 1 through 3, respectively. See Appendix E, Tables 5, 6, and Chart 1 for complete model results. Given the very small number of delinquent loans it is difficult to conclude if these AUCs meaningfully diverge from .5 (which would indicate no predictive power) or if any of these scores have a robust ability to predict likelihood of default.

59. Participant 5 has a number of customers with limited or no credit experience, as approximately 3.5% of Participant 5's customers did not have a Vantage score and 7.7% of originations have less than three open trade lines. Among customers with a Vantage Score, approximately 50% had a score below 654.

60. With respect to fair lending risk, we found evidence that the use of the cash flow metrics and CFS did not create a disparate impact among the sample population; however the small size of the population means we should interpret this with caution at this time. Using the BISG probabilities to identify separate groups of borrowers with a high likelihood of belonging to each race/ethnicity group,[34] we divided the not past due and past due populations into demographically neutralized sub-populations and tested the difference in means within each race/ethnicity group. The two cash flow metrics demonstrated statistically significant differences between past due and not past due loans among nearly all race/ethnicity groups. The same is not true with respect to the traditional credit metrics. (See Appendix E, Table 9.)

61. We tested the ability of the CFS and Vantage scores to rank order risk in the demographically neutralized sample populations. We obtained an AUC of .55 when testing the CFS's ability to rank order credit risk among a group of highly likely non-Hispanic white

[34] We were not able to test gender.

borrowers. This compares to AUCs of .672, .557 and .649 for likely African American, Hispanic and Asian borrowers, respectively. (See Appendix E, Table 10, Model 2.) The relative larger AUCs across the minority samples likely reflect the relatively larger past-due populations among these groups. Nonetheless, the result is encouraging. We ran the same tests with respect to the Vantage score (Model 1) and CF and Vantage score combined (Model 3). For Model 3 we obtained an AUC of .665 when testing the CFS and Vantage scores' combined ability to rank order credit risk among a group of highly likely non-Hispanic white borrowers. This compares to AUCs of .689, .731 and .693 for likely African American, Hispanic and Asian borrowers, respectively. See Appendix E, Table 10, Model 3. This result may most closely reflect the process utilized by the Participant's highly automated underwriting process, and the results suggest more consistent ability to rank order credit risk within each demographically neutralized population among the sample population. See Appendix E, Tables 11 – 13 for the full model output for each logistic regression.

4.6. Participant #6

62. Participant 6's automated underwriting process uses a series of cash flow metrics, but does not utilize a cash flow based score. Participant 6 provided to CRA a transaction-level data file containing 13,431 applications, which resulted in 3,776 originated loans. Where available, they provided their twenty-five cash flow metrics, as well as traditional credit bureau information and credit scores, and a delinquency indicator. See Appendix F, Tables 1 and 3 for basic diagnostics on the data provided. To better understand the underwriting outcomes, we separated the applicants into approved and declined groups and performed a difference in means test between the two groups on the cash flow metrics and the traditional credit bureau attributes and scores. See Appendix F, Table 2 for the test results.

63. We found compelling evidence that the cash flow metrics are correlated with likelihood of default within the sample population. We separated the borrowers into delinquent and non-delinquent populations and performed a difference in means test between the two groups on each of the cash flow metrics. Twenty-two of the twenty-five provided cash flow

metrics were observed to have statistically significant differences among the delinquent as compared to non-delinquent borrowers. See Appendix F, Table 4 for the test results.

64. Next, we estimated several logit models of delinquency and calculated AUCs based on each. In the first model, we included as predictors the traditional credit score and bureau information. In the second model, we included as predictors the cash flow metrics found to have statistically significant differences in means between delinquent borrowers and non-delinquent borrowers. In the third model, we included all of the cash flow metrics as predictors. In the fourth model, we included all of the cash flow metrics and the traditional credit bureau information and scores as predictors.

65. The AUCs obtained were .720, .675, .688, and .758 for models 1 through 4, respectively. See Appendix F, Table 5 and Chart 1 for complete model results. These AUCs meaningfully diverge from .5 (which would indicate no predictive power) and are at levels which, in our experience, suggest a relatively robust ability to predict likelihood of default within the sample population. While the traditional credit score and bureau information outperforms the cash flow scores on their own, the model is improved by using by both the traditional score and the cash flow information. To further understand the relationship between the default rates, the cash flow metrics, and traditional credit score measures, we used the results of model 2 to estimate the default probability of each loan as predicted by the cash flow metrics. We divided the loans into twenty groups from lowest to highest default probability and traditional credit scores and calculated the default rate within each group. The resulting 'heat map' is reported in Appendix F, Table 6. The rows and columns are interpreted in the same manner as for the previous heat maps.

66. Participant 6 has a number of customers with limited or no credit experience. Eight percent of the approvals did not have a FICO score and 6% had no open accounts. Among Participant 6's customers with a FICO score, more than 50% had a score below 650, and 25% had a score under 597. Participant 6 was able to approve 45% of applications that did not have a FICO score compared with 76% who did have a FICO score. More than 50% of Participant 6's customers have only one open account on their credit bureau. These metrics

suggest Participant 6 was able to lend to borrowers who might struggle to qualify for loans using a traditional score.

67. The data included zip code and a proxy for income which allowed us to make some potential inferences as to the demographics of customers obtaining credit from Participant 6. Approximately 51% of the loans in the sample population were made to customers residing in a majority minority zip code, based upon data from the 2017 American Community Survey ("ACS") (see Appendix F, Table 7).[35] Approximately 29% of the loans were made to customers residing in predominantly minority zip codes, based upon data from the 2017 ACS (see Appendix F, Table 8).[36] While such metrics are difficult to put into context, these shares suggest a relatively high level of minority customers seeking and gaining access to the product offered by Participant 6. We also report (see Appendix F, Tables 7 and 8) the shares of delinquent and non-delinquent customers by majority minority zip code and by predominantly minority zip code. We observe a higher delinquency rate among customers residing in predominantly minority or majority minority zip codes as compared to those not residing in such zip codes. These are raw delinquency rates, uncontrolled for any differences in customers' creditworthiness.

68. Finally, we compared the income proxy available for each customer relative to the median household income of the zip code in which each customer resides. The income proxy is based upon information in the application and measures personal net income. Thus, the observation that approximately 59% of the customers have incomes below the median household income of the zip code in which they reside should be interpreted with caution (see Appendix F, Table 9).

[35] Majority minority zip codes are those in which the minority population exceeds the non-Hispanic white population. That is, less than 50% of the residents in the zip code are reported as non-Hispanic white, based upon the 2017 ACS.

[36] Predominantly minority zip codes are those in which the minority population exceeds 80% of the total population of the zip code, based upon the 2017 ACS.

69. With respect to fair lending risk, we found evidence that use of the cash flow data did not create a disparate impact among the sample population. The BISG probabilities were used to identify separate groups of borrowers with a high likelihood of belonging to each race/ethnicity group. Gender proxies were also available for testing. First, we divided the not past due and past due populations into demographically neutralized sub-populations and tested the difference in means within each race/ethnicity and gender. The majority of cash flow metrics demonstrated statistically significant differences between past due and not past due loans among nearly all groups in the sample population. The same was true with respect to the traditional credit score. See Appendix F, Table 11.

70. We tested the ability of the cash flow metrics (Models 2 and 3) to rank order risk in the demographically neutralized sample populations.[37] We obtained an AUC of .802 when testing the cash flow data's ability to rank order credit risk among a group of highly likely non-Hispanic white borrowers (from model 3). This compares to AUCs of .766, and .759, for likely African American and Hispanic borrowers, respectively (the population of Asian borrowers was too small for reliable estimation and comparison across all models). (See Appendix F, Table 12.) We repeated this process with respect to gender and obtained AUCs of .702 and .711 for male and female borrowers, respectively. The relative consistency of the AUC across these demographically neutralized sample populations is encouraging, and suggests that the cash flow models are likely not simply proxies for race/ethnicity, but are able to rank order risk within demographic groups within the sample population. See Appendix F, Tables 13 – 18 for the full model output for each logistic regression.

[37] We were unable to get Model 4 to converge when run on demographically neutralized sample populations.

APPENDIX A: Participant 1

Appendix A. Participant #1

Appendix A. Participant #1

Table 1. Data Diagnostics: Originated Loans

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Cash Flow Metric #1	Delinquent	748	0	0	$385	$470	$737	$957	$1,269	$1,957	$4,038	$1,065
	Not Delinquent	10,209	0	5	$0	$473	$729	$986	$1,380	$2,283	$9,441	$1,140
	All	10,957	0	5	$0	$472	$729	$983	$1,370	$2,272	$9,441	$1,134
Cash Flow Metric #2	Delinquent	748	0	17	0.0	1.0	2.0	4.0	5.0	6.0	6.0	3.7
	Not Delinquent	10,209	0	136	0.0	1.0	3.0	5.0	5.0	6.0	6.0	4.1
	All	10,957	0	153	0.0	1.0	3.0	5.0	5.0	6.0	6.0	4.1
Cash Flow Metric #3	Delinquent	748	16	0	1.0	1.0	2.0	4.0	11.0	18.0	165.0	8.0
	Not Delinquent	10,209	227	0	1.0	1.0	2.0	4.0	8.0	15.0	178.0	6.5
	All	10,957	243	0	1.0	1.0	2.0	4.0	8.0	15.0	178.0	6.6
Cash Flow Metric #4	Delinquent	748	0	35	0.0	1.0	5.0	11.0	13.0	17.0	32.0	9.6
	Not Delinquent	10,209	5	317	0.0	1.0	5.0	9.0	13.0	15.0	37.0	8.8
	All	10,957	5	352	0.0	1.0	5.0	9.0	13.0	15.0	37.0	8.9
Cash Flow Metric #5	Delinquent	748	0	273	0.0	0.0	0.0	6.0	19.5	42.0	61.0	11.7
	Not Delinquent	10,209	0	3,246	0.0	0.0	0.0	10.0	25.0	46.0	62.0	14.4
	All	10,957	0	3,519	0.0	0.0	0.0	9.0	24.0	46.0	62.0	14.2
Cash Flow Metric #6	Delinquent	748	0	0	$27	$34	$56	$71	$98	$160	$317	$82
	Not Delinquent	10,209	0	5	$0	$37	$60	$79	$109	$181	$1,025	$91
	All	10,957	0	5	$0	$37	$60	$78	$109	$180	$1,025	$90
Cash Flow Metric #7	Delinquent	748	0	33	$0	$28	$53	$70	$96	$153	$282	$77
	Not Delinquent	10,209	0	184	$0	$34	$59	$78	$108	$177	$454	$89
	All	10,957	0	217	$0	$34	$58	$77	$107	$175	$454	$88
Cash Flow Metric #8	Delinquent	748	0	0	$150	$1,083	$1,982	$2,734	$3,993	$6,664	$21,424	$3,209
	Not Delinquent	10,209	0	0	$214	$1,344	$2,227	$3,143	$4,521	$7,799	$62,413	$3,679
	All	10,957	0	0	$150	$1,322	$2,200	$3,119	$4,476	$7,736	$62,413	$3,647
Cash Flow Metric #9	Delinquent	748	0	3	$0	$342	$810	$1,216	$1,768	$3,630	$24,081	$1,541
	Not Delinquent	10,209	0	6	$0	$413	$820	$1,257	$1,907	$3,627	$76,069	$1,579

	All	10,957	0	9	$0	$410	$820	$1,253	$1,898	$3,629	$76,069	$1,577
Cash Flow Metric #10	Delinquent	748	0	0	$10	$1,040	$1,937	$2,699	$3,895	$6,582	$23,121	$3,178
	Not Delinquent	10,209	0	0	$134	$1,317	$2,213	$3,124	$4,465	$7,747	$61,758	$3,654
	All	10,957	0	0	$10	$1,298	$2,192	$3,097	$4,421	$7,700	$61,758	$3,622
Cash Flow Metric #11	Delinquent	748	0	3	$0	$427	$842	$1,208	$1,824	$3,667	$28,428	$1,549
	Not Delinquent	10,209	0	6	$0	$411	$806	$1,227	$1,868	$3,490	$74,914	$1,540
	All	10,957	0	9	$0	$412	$809	$1,225	$1,863	$3,506	$74,914	$1,540
Cash Flow Metric #12	Delinquent	748	0	0	0.61	0.83	0.97	1.02	1.09	1.79	257.28	1.81
	Not Delinquent	10,209	0	0	0.39	0.84	0.97	1.02	1.08	1.56	448.47	1.29
	All	10,957	0	0	0.39	0.84	0.97	1.02	1.08	1.58	448.47	1.33
Cash Flow Metric #13	Delinquent	748	0	5	0.00	0.03	0.11	0.23	0.43	1.84	512.61	1.80
	Not Delinquent	10,209	0	14	0.00	0.05	0.14	0.23	0.42	1.29	632.84	0.79
	All	10,957	0	19	0.00	0.05	0.14	0.23	0.42	1.32	632.84	0.86
Cash Flow Metric #14	Delinquent	748	0	0	9.34%	29.83%	37.53%	43.89%	52.89%	66.47%	87.15%	45.28%
	Not Delinquent	10,209	0	0	0.55%	29.81%	33.70%	38.46%	45.60%	59.02%	92.07%	40.25%
	All	10,957	0	0	0.55%	29.81%	33.88%	38.89%	46.11%	59.89%	92.07%	40.60%
Cash Flow Metric #15	Delinquent	748	114	53	0.00%	0.00%	3.85%	13.94%	30.39%	55.80%	67.96%	19.39%
	Not Delinquent	10,209	770	1,026	0.00%	0.00%	2.22%	8.89%	22.78%	54.14%	70.00%	15.38%
	All	10,957	884	1,079	0.00%	0.00%	2.22%	9.29%	23.20%	54.40%	70.00%	15.63%
Cash Flow Metric #16	Delinquent	748	0	17	-$193	-$35	$2	$36	$175	$783	$4,735	$170
	Not Delinquent	10,209	0	118	-$413	-$7	$12	$90	$305	$1,089	$11,323	$254
	All	10,957	0	135	-$413	-$10	$11	$87	$294	$1,061	$11,323	$249
Cash Flow Metric #17	Delinquent	748	0	0	-$2,071	-$266	$94	$195	$355	$779	$5,807	$250
	Not Delinquent	10,209	0	0	-$196,145	-$156	$161	$282	$472	$1,004	$17,313	$334
	All	10,957	0	0	-$196,145	-$167	$155	$275	$464	$998	$17,313	$328
Cash Flow Metric #18	Delinquent	748	0	0	$57	$145	$245	$349	$508	$1,034	$15,189	$457
	Not Delinquent	10,209	0	0	$34	$176	$279	$403	$599	$1,171	$15,306	$517
	All	10,957	0	0	$34	$173	$276	$399	$593	$1,159	$15,306	$513
Cash Flow Metric #19	Delinquent	748	0	1	-$1,144	$25	$447	$659	$980	$1,740	$6,033	$781
	Not Delinquent	10,209	0	8	-$196,901	$236	$517	$772	$1,144	$2,166	$17,468	$912

	All	10,957	0	9	-$196,901	$213	$512	$765	$1,134	$2,145	$17,468	$903
Cash Flow Metric #20	Delinquent	748	0	1	-$1,929	-$217	$148	$334	$562	$1,115	$5,810	$417
	Not Delinquent	10,209	0	9	-$180,686	-$71	$238	$411	$698	$1,508	$16,770	$513
	All	10,957	0	10	-$180,686	-$87	$231	$405	$690	$1,492	$16,770	$507
Cash Flow Metric #21	Delinquent	748	0	1	$0	$118	$243	$344	$521	$1,055	$15,328	$459
	Not Delinquent	10,209	0	8	$0	$123	$233	$349	$539	$1,158	$15,610	$473
	All	10,957	0	9	$0	$123	$235	$348	$538	$1,150	$15,610	$472
Cash Flow Metric #22	Delinquent	748	0	1	$0	$70	$188	$297	$462	$1,014	$16,489	$409
	Not Delinquent	10,209	0	9	$0	$102	$209	$325	$512	$1,125	$56,925	$453
	All	10,957	0	10	$0	$98	$208	$323	$508	$1,115	$56,925	$450
Cash Flow Metric #23	Delinquent	748	0	0	1.0	1.0	1.0	1.0	1.0	2.0	4.0	1.2
	Not Delinquent	10,209	0	5	0.0	1.0	1.0	1.0	1.0	2.0	5.0	1.2
	All	10,957	0	5	0.0	1.0	1.0	1.0	1.0	2.0	5.0	1.2
Cash Flow Metric #24	Delinquent	748	0	427	0.0	0.0	0.0	0.0	1.0	3.0	7.0	0.7
	Not Delinquent	10,209	0	4,656	0.0	0.0	0.0	1.0	2.0	3.0	9.0	1.0
	All	10,957	0	5,083	0.0	0.0	0.0	1.0	2.0	3.0	9.0	1.0

Appendix A. Participant #1 Table 2. Difference of Means Tests: Originated Loans[38]					
Variable	**Sample**	**#**	**Mean**	**T-Stat**	**P-Value**
Cash Flow Metric #1	Delinquent	748	$1,065	.	.
	Not Delinquent	10,209	$1,140	3.79	0.000
Cash Flow Metric #2	Delinquent	748	3.7	.	.
	Not Delinquent	10,209	4.1	7.23	0.000
Cash Flow Metric #3	Delinquent	732	8.0	.	.
	Not Delinquent	9,982	6.5	-2.69	0.007
Cash Flow Metric #4	Delinquent	748	9.6	.	.
	Not Delinquent	10,204	8.8	-3.46	0.001
Cash Flow Metric #6	Delinquent	748	$82	.	.
	Not Delinquent	10,209	$91	5.98	0.000
Cash Flow Metric #7	Delinquent	748	$77	.	.
	Not Delinquent	10,209	$89	6.98	0.000
Cash Flow Metric #8	Delinquent	748	$3,209	.	.
	Not Delinquent	10,209	$3,679	6.14	0.000
Cash Flow Metric #9	Delinquent	748	$1,541	.	.
	Not Delinquent	10,209	$1,579	0.68	0.494
Cash Flow Metric #10	Delinquent	748	$3,178	.	.
	Not Delinquent	10,209	$3,654	6.16	0.000
Cash Flow Metric #11	Delinquent	748	$1,549	.	.
	Not Delinquent	10,209	$1,540	-0.15	0.880
Cash Flow Metric #12	Delinquent	748	1.81	.	.
	Not Delinquent	10,209	1.29	-1.37	0.170
Cash Flow Metric #13	Delinquent	748	1.80	.	.
	Not Delinquent	10,209	0.79	-1.37	0.172
Cash Flow Metric #14	Delinquent	748	45.28%	.	.
	Not Delinquent	10,209	40.25%	-11.15	0.000
Cash Flow Metric #15	Delinquent	634	19.39%	.	.
	Not Delinquent	9,439	15.38%	-5.43	0.000
Cash Flow Metric #16	Delinquent	748	$170	.	.
	Not Delinquent	10,209	$254	5.78	0.000
Cash Flow Metric #17	Delinquent	748	$250	.	.
	Not Delinquent	10,209	$334	3.04	0.002

38 The significance test tests the difference in means between the delinquent and not delinquent populations using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% level.

Cash Flow Metric #18	Delinquent	748	$457	.	.
	Not Delinquent	10,209	$517	2.44	0.015
Cash Flow Metric #19	Delinquent	748	$781	.	.
	Not Delinquent	10,209	$912	4.11	0.000
Cash Flow Metric #20	Delinquent	748	$417	.	.
	Not Delinquent	10,209	$513	3.32	0.001
Cash Flow Metric #21	Delinquent	748	$459	.	.
	Not Delinquent	10,209	$473	0.54	0.589
Cash Flow Metric #22	Delinquent	748	$409	.	.
	Not Delinquent	10,209	$453	1.63	0.104
Cash Flow Metric #23	Delinquent	748	1.2	.	.
	Not Delinquent	10,209	1.2	1.11	0.267
Cash Flow Metric #24	Delinquent	748	0.7	.	.
	Not Delinquent	10,209	1.0	7.92	0.000

Appendix A. Participant #1

Table 3. Logistic Models for Delinquency Results[39]

Model	AUC
(1) Cash Flow Metrics Important in Underwriting	0.597
(2) Statistically Significant Cash Flow Metrics, Dates and Institution Controls	0.713
(3) All Cash Flow Metrics, Dates and Institution Controls	0.725

[39] The dependent variable is a 0/1 indicator for delinquent, with values of 1 indicating delinquent and 0 indicating not delinquent. Model 1 includes only the five fields that participant 1 identifies as among the most important in their underwriting process. Model 2 includes all cash flow metrics found to have statistically significant differences in means among delinquent borrowers as compared to non-delinquent borrowers as well as statistically significant dates and institution controls. Model 3 includes all cash flow metrics as predictors as well as statistically significant dates and institution controls. The full model output was estimated using a "training" data set. This training data set contains a random sample of 75% of the records from the full data set.

Appendix A. Participant #1

Table 4. Logistic Model for Delinquency Specifications[40]

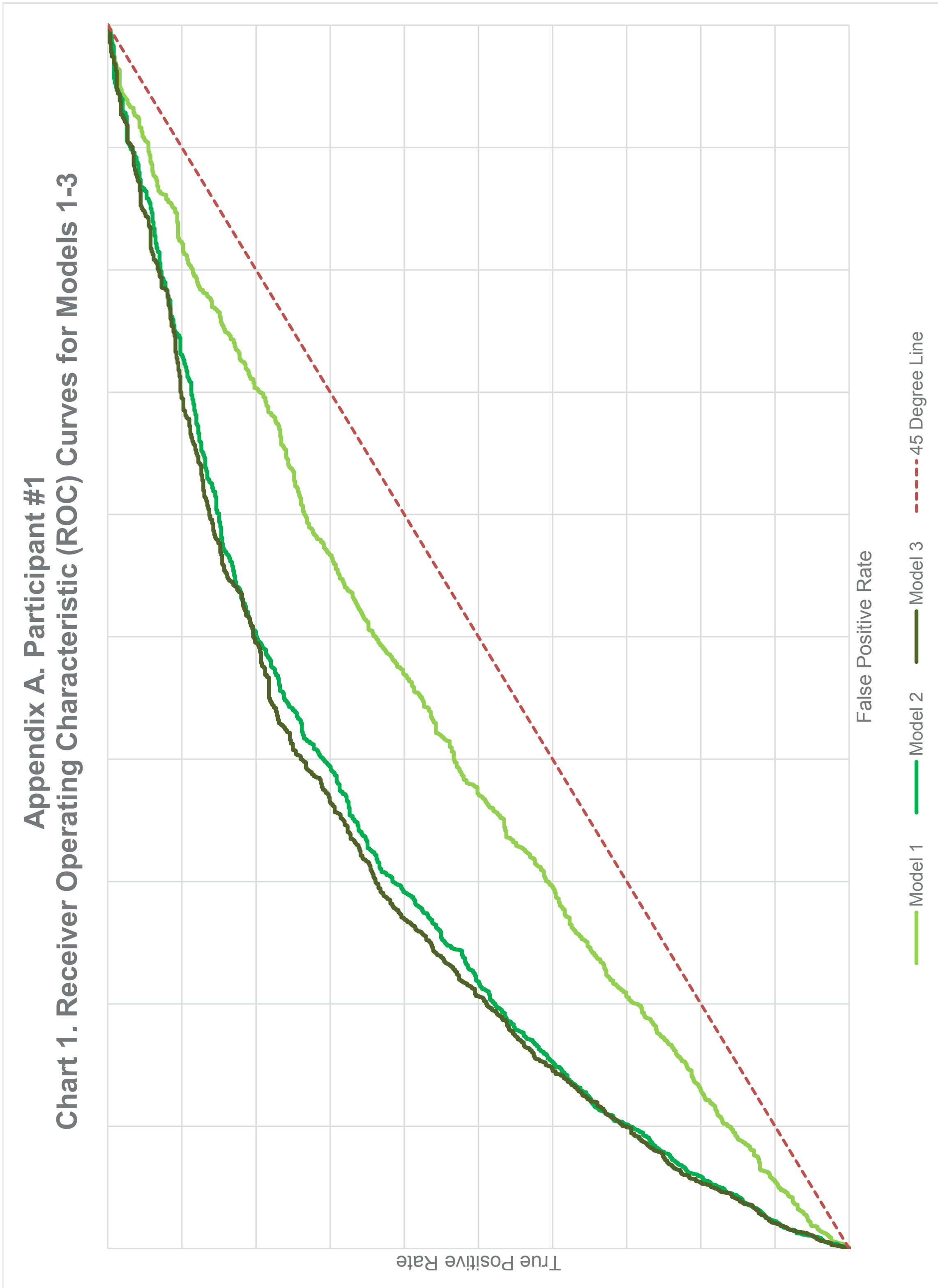
Control Variable	Comparison Group	Model 1		Model 2		Model 3	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Metric #1	--	1.01	0.38	1.01	0.66	1.00	0.82
Cash Flow Metric #7	--	0.54	0.00	0.66	0.06	0.66	0.06
Cash Flow Metric #16	--	0.95	0.01	0.97	0.09	0.97	0.10
Cash Flow Metric #19	--	0.99	0.66	0.94	0.02	0.94	0.01
Cash Flow Metric #20	--	1.01	0.68	0.99	0.77	0.98	0.48
Loan Amount ($100)	--	.	.	.	.	1.57	0.01
Cash Flow Metric #11	--	.	.	.	.	1.03	0.04
Cash Flow Metric #14	--	.	.	27.05	0.00	17.25	0.00
Cash Flow Metric #17	--	.	.	1.07	0.00	1.08	0.00
Cash Flow Metric #18	--	.	.	0.99	0.67	0.94	0.03
Cash Flow Metric #21	--	.	.	.	.	1.05	0.04
Cash Flow Metric #24	--	.	.	0.89	0.12	0.88	0.10
Cash Flow Metric #2	--	.	.	0.91	0.00	0.92	0.01
Cash Flow Metric #3	--	.	.	1.00	0.20	1.00	0.15
Missing Cash Flow Metric #3	Not Missing Cash Flow Metric #3	.	.	0.85	0.61	0.88	0.69
Cash Flow Metric #4	--	.	.	1.02	0.15	1.01	0.24

40 The dependent variable is a 0/1 indicator for delinquent, with values of 1 indicating delinquent and 0 indicating not delinquent. Model 1 includes only the five fields that participant 1 identifies as among the most important in their underwriting process. Model 2 includes all cash flow metrics found to have statistically significant differences in means among delinquent borrowers as compared to non-delinquent borrowers as well as statistically significant dates and institution controls. Model 3 includes all cash flow metrics as predictors as well as statistically significant dates and institution controls. The full model output was estimated using a "training" data set. This training data set contains a random sample of 75% of the records from the full data set. The units of the cash flow variables are in $100's.

Missing Cash Flow Metric #4	Not Missing Cash Flow Metric #4	.	.	.	.	.	.
Cash Flow Metric #5	--	.	.	1.00	0.70	1.00	0.51
Cash Flow Metric #6	--	.	.	1.00	0.98	0.94	0.75
Cash Flow Metric #8	--	.	.	1.01	0.58	1.01	0.79
Cash Flow Metric #9	--	.	.	.	.	0.98	0.14
Cash Flow Metric #10	--	.	.	1.00	0.93	1.00	0.89
Cash Flow Metric #12	--	.	.	.	.	0.95	0.08
Cash Flow Metric #13	--	.	.	.	.	1.04	0.05
Cash Flow Metric #15	--	.	.	1.72	0.05	1.68	0.07
Missing Cash Flow Metric #15	Not Missing Cash Flow Metric #15	.	.	2.83	0.01	2.72	0.01
Cash Flow Metric #22	--	.	.	.	.	.	.
Cash Flow Metric #23	--	.	.	.	.	1.19	0.15
Date #1 Bucket B	Date #1 Bucket C	.	.	.	.	1.39	0.05
Date #1 Bucket A		.	.	.	.	1.37	0.32
Date #2 Bucket B	Date #2 Bucket C	.	.	1.15	0.18	0.92	0.62
Date #2 Bucket A		.	.	0.79	0.60	0.67	0.46
Constant		0.13	0.00	0.03	0.00	0.02	0.00
Pseudo R Squared		0.013		0.075		0.082	
AUC		0.597		0.713		0.725	
Sample Size		8,218		8,155		8,155	

Appendix A. Participant #1

Chart 1. Receiver Operating Characteristic (ROC) Curves for Models 1-3



Appendix A. Participant #1 Table 5. Summary of Whether Applicant's Zip Code Population is at least 50% Minority, by Delinquency Status[41]									
	Delinquent			Not Delinquent			All		
Value	#	Row %	Col %	#	Row %	Col %	#	%	P-val
Missing	64	8.9%	8.6%	658	91.1%	6.4%	722	6.6%	0.032
False	213	6.7%	28.5%	2,986	93.3%	29.2%	3,199	29.2%	0.677
True	471	6.7%	63.0%	6,565	93.3%	64.3%	7,036	64.2%	0.477
All	748	6.8%	100.0%	10,209	93.2%	100.0%	10,957	100.0%	.

Appendix A. Participant #1 Table 6. Summary of Whether Applicant's Zip Code Population is at least 80% Minority, by Delinquency Status									
	Delinquent			Not Delinquent			All		
Value	#	Row %	Col %	#	Row %	Col %	#	%	P-val
Missing	64	8.9%	8.6%	658	91.1%	6.4%	722	6.6%	0.032
False	460	6.5%	61.5%	6,596	93.5%	64.6%	7,056	64.4%	0.089
True	224	7.0%	29.9%	2,955	93.0%	28.9%	3,179	29.0%	0.559
All	748	6.8%	100.0%	10,209	93.2%	100.0%	10,957	100.0%	.

Appendix A. Participant #1 Table 7. Summary of Whether Applicant's Income Exceeds Zip Code's Median Income, by Delinquency Status									
	Delinquent			Not Delinquent			All		
Value	#	Row %	Col %	#	Row %	Col %	#	%	P-val
Missing	66	8.8%	8.8%	680	91.2%	6.7%	746	6.8%	0.029
False	616	6.8%	82.4%	8,498	93.2%	83.2%	9,114	83.2%	0.543
True	66	6.0%	8.8%	1,031	94.0%	10.1%	1,097	10.0%	0.283
All	748	6.8%	100.0%	10,209	93.2%	100.0%	10,957	100.0%	.

41 Missing demographic data is the result of invalid zip codes, zip codes outside of the 50 States, or zip codes that do not have an associated ZCTA (Zip Code Tabulation Area).

APPENDIX B: Participant 2

Appendix B. Participant #2

Appendix B. Participant #2
Table 1. Data Diagnostics: All Applications

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Cash Flow Score	Denied	154,425	154,425	0	.	.	.	.	.	.	.	.
	Approved	58,524	10	0	318	602	659	691	715	735	850	683
	All Applications	212,949	154,435	0	318	602	659	691	715	735	850	683
FICO Score	Denied	154,425	119,915	0	538	546	576	614	661	750	850	626
	Approved	58,524	4,879	0	538	570	625	662	702	771	850	665
	All Applications	212,949	124,794	0	538	553	602	646	690	765	850	650

Appendix B. Participant #2
Table 2. Difference of Means Tests: All Applications[42]

Variable	Sample	#	Mean	T-Stat	P-Value
Cash Flow Score	Denied	0	.	.	.
	Approved	58,514	683	.	.
FICO Score	Denied	34,510	626	.	.
	Approved	53,645	665	-92.66	0.000

42 The significance test tests the difference in means between the approved and denied populations using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% level. Counts in this table are of non-missing values of the indicated variable.

Appendix B. Participant #2
Table 3. Data Diagnostics: Originated Loans[43]

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Cash Flow Score	Not Delinquent	33,984	0	0	356	598	655	687	713	734	850	680
	Delinquent	6,927	0	0	318	579	626	661	691	723	756	657
	Originated Loans	40,911	0	0	318	593	649	683	710	733	850	676
FICO Score	Not Delinquent	33,984	322	0	538	569	624	662	702	770	850	665
	Delinquent	6,927	59	0	538	555	597	631	669	743	850	637
	Originated Loans	40,911	381	0	538	565	619	657	697	767	850	660

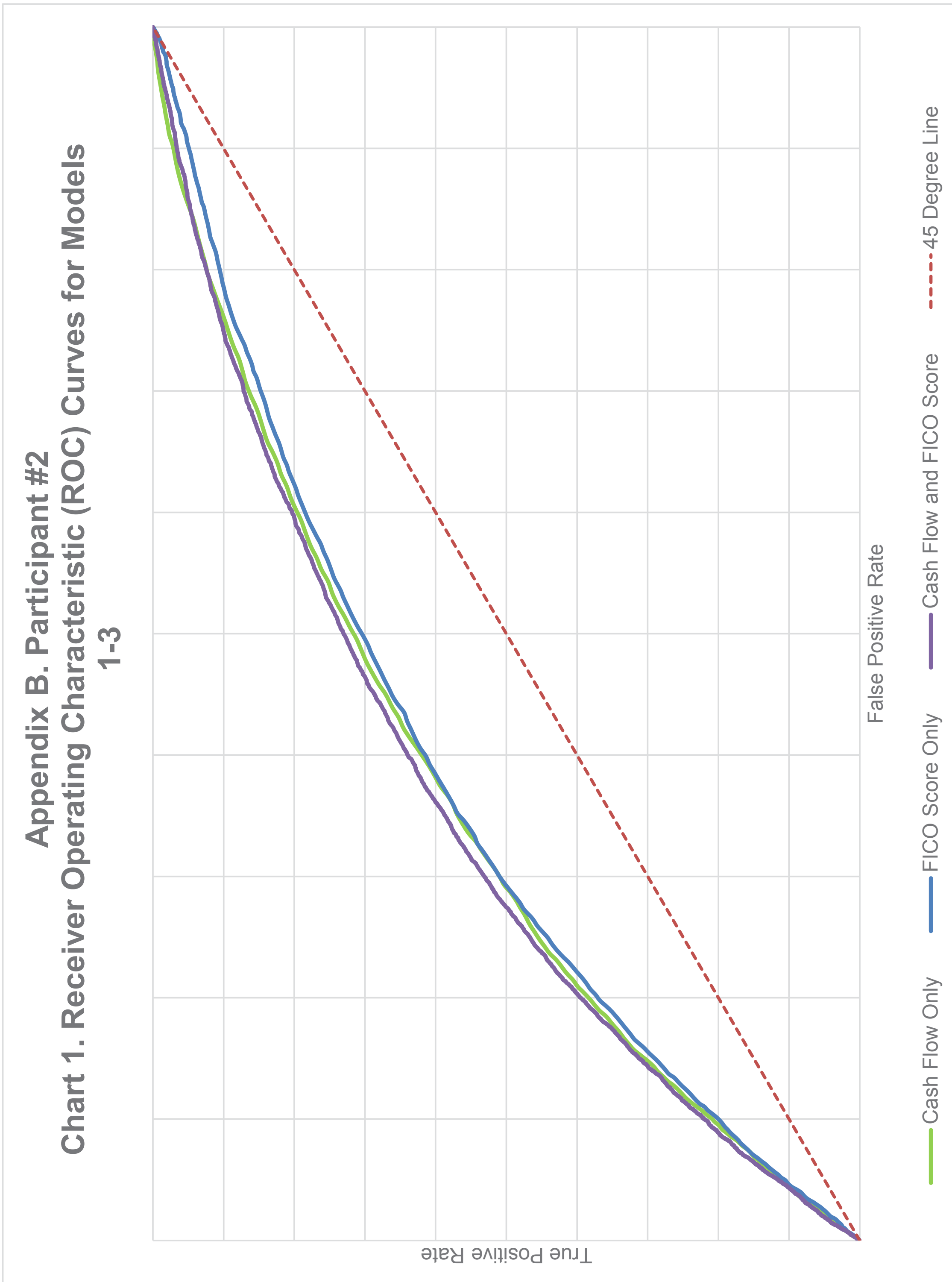
Appendix B. Participant #2
Table 4. Difference of Means Tests: Originated Loans

Variable	Sample	#	Mean	T-Stat	P-Value
Cash Flow Score	Not Delinquent	33,984	680	.	.
	Delinquent	6,927	657	39.26	0.000
FICO Score	Not Delinquent	33,662	665	.	.
	Delinquent	6,868	637	35.94	0.000

43 Delinquent status reflects loans with a positive bad balance.

Appendix B. Participant #2 Table 5. Logistic Model for Delinquency Specifications[44]							
		FICO Score Only		Cash Flow Score Only		Cash Flow Score and FICO Score	
Control Variable	**Comparison Group**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
Cash Flow Score	--	.	.	0.99	0.00	0.99	0.00
FICO Score	--	0.99	0.00	.	.	1.00	0.00
Missing FICO	Not Missing FICO	0.00	0.00	.	.	0.05	0.00
Constant		49.03	0.00	428.33	0.00	812.80	0.00
Pseudo R-Squared		0.034		0.041		0.047	
AUC		0.640		0.652		0.660	
Sample Size		40,911		40,911		40,911	

[44] The dependent variable is a 0/1 indicator for delinquent, with values of 1 indicating delinquent and 0 indicating not delinquent.



Page **43** of **161**

Appendix B. Participant #2

Table 6. Delinquency Frequency by Cash Flow Score Percentile and FICO Score Percentile[45]

	Cash Flow Score																			
FICO Score	**0 - 5th**	**5 - 10th**	**10 - 15th**	**15 - 20th**	**20 - 25th**	**25 - 30th**	**30 - 35th**	**35 - 40th**	**40 - 45th**	**45 - 50th**	**50 - 55th**	**55 - 60th**	**60 - 65th**	**65 - 70th**	**70 - 75th**	**75 - 80th**	**80 - 85th**	**85 - 90th**	**90 - 95th**	**95 - 100th**
0 - 5th	35.5	26.4	31.4	31.5	27.8	25.0	22.9	9.7	13.6	25.0	18.2	20.0	.	60.0	.	.	.	.	.	.
5 - 10th	33.0	31.0	29.7	28.8	25.4	20.3	29.4	29.3	25.0	40.0	4.8	0.0	11.1	0.0	.	.	.	.	.	.
10 - 15th	32.7	35.6	27.1	27.4	32.7	23.2	20.6	19.3	14.3	24.6	20.9	8.5	34.3	12.0	11.8	12.5	16.7	.	.	.
15 - 20th	37.7	25.1	28.4	26.2	28.2	26.0	25.4	21.2	27.6	20.6	21.1	25.0	17.9	14.0	17.4	9.1	18.2	0.0	.	.
20 - 25th	30.3	34.4	30.5	28.9	23.8	26.8	23.2	21.9	20.7	20.5	12.5	21.3	12.2	25.0	11.5	11.4	12.0	8.3	20.0	0.0
25 - 30th	33.8	34.1	29.0	22.8	34.6	23.5	16.3	25.2	24.3	20.5	13.9	19.8	17.9	22.4	10.3	13.4	11.5	5.3	0.0	12.5
30 - 35th	27.4	30.2	27.9	30.9	24.0	26.0	23.5	19.0	16.8	16.1	16.4	20.0	12.2	11.4	16.5	18.0	19.2	13.0	12.5	9.1
35 - 40th	24.0	22.6	33.3	25.0	25.2	21.4	19.8	19.7	16.2	17.0	15.3	17.2	15.0	13.3	13.4	13.9	19.7	6.4	2.2	13.8
40 - 45th	18.9	27.5	33.8	27.5	17.1	19.4	24.2	10.1	21.4	19.6	14.4	10.8	12.8	10.5	12.5	16.9	10.8	14.3	13.0	6.5
45 - 50th	20.7	7.1	17.2	18.8	22.6	11.7	18.1	24.2	19.1	19.0	20.3	20.4	12.7	10.8	12.5	13.5	10.9	10.7	12.0	3.0
50 - 55th	32.0	10.3	23.9	16.1	19.5	20.0	15.1	14.1	15.4	17.0	14.8	17.8	12.9	11.9	16.2	12.3	10.7	14.5	8.9	10.6
55 - 60th	30.0	15.2	14.6	15.4	21.5	22.9	14.8	17.3	15.1	15.7	11.4	16.8	10.5	15.4	9.7	10.9	7.3	8.5	7.3	8.5
60 - 65th	33.3	20.7	24.5	12.5	20.4	13.2	21.0	15.8	25.7	13.7	12.6	10.3	10.4	16.0	12.2	9.5	9.5	8.6	10.3	9.1
65 - 70th	30.0	15.4	13.6	20.0	16.0	18.4	7.1	19.8	18.8	13.1	17.0	11.6	8.1	7.7	10.6	11.2	12.3	10.3	3.5	6.2
70 - 75th	12.5	18.8	19.4	15.4	12.3	9.5	11.3	10.6	14.1	15.7	11.8	11.0	12.0	12.6	14.8	11.4	6.7	9.1	4.3	4.4
75 - 80th	19.0	10.5	22.2	14.3	17.8	15.3	12.7	12.5	16.2	11.2	17.1	9.3	10.0	11.6	9.4	13.4	9.5	9.7	10.4	2.7
80 - 85th	18.8	31.3	11.8	12.5	0.0	6.6	14.9	12.0	14.8	10.9	12.0	6.4	9.0	7.6	5.3	7.5	4.0	8.1	7.5	3.9
85 - 90th	15.4	42.1	33.3	20.0	13.0	8.8	6.3	10.7	14.1	16.2	10.7	6.4	9.2	8.1	8.9	5.6	6.1	4.1	5.9	3.0
90 - 95th	28.6	36.4	29.2	23.3	6.7	7.5	12.5	6.0	14.0	10.1	20.2	10.2	9.9	6.5	5.2	8.3	6.9	9.1	6.0	2.8
95 - 100th	18.2	21.4	55.6	38.9	17.2	12.5	10.9	12.1	15.9	11.7	16.5	13.3	13.7	8.3	10.9	11.0	6.6	11.2	4.2	1.7

45 Cells are shaded based on values. Green indicates values close to the lowest default frequency, yellow indicates values close to the median default frequency, and red indicates values close to the highest default frequency. Cells with fewer than 5 loans are excluded from this heat map. Percentiles are based on the population of originated loans. 381 originated loans with a missing FICO score were excluded from the frequency table.

Appendix B. Participant #2

Table 7. Summary of Whether The Applicant's Zip Code Population is at least 50% Minority, by Delinquency Status

Value	Delinquent #	Delinquent Row %	Delinquent Col %	Not Delinquent #	Not Delinquent Row %	Not Delinquent Col %	All #	All %	P-Val
Missing	35	16.7%	0.5%	175	83.3%	0.5%	210	0.5%	1.000
False	4,557	15.6%	65.8%	24,572	84.4%	72.3%	29,129	71.2%	0.000
True	2,335	20.2%	33.7%	9,237	79.8%	27.2%	11,572	28.3%	0.000
All	6,927	16.9%	100.0%	33,984	83.1%	100.0%	40,911	100.0%	.

Appendix B. Participant #2

Table 8. Summary of Whether The Applicant's Zip Code Population is at least 80% Minority, by Delinquency Status[46]

Value	Delinquent #	Delinquent Row %	Delinquent Col %	Not Delinquent #	Not Delinquent Row %	Not Delinquent Col %	All #	All %	P-Val
Missing	35	16.7%	0.5%	175	83.3%	0.5%	210	0.5%	1.000
False	6,176	16.5%	89.2%	31,175	83.5%	91.7%	37,351	91.3%	0.000
True	716	21.4%	10.3%	2,634	78.6%	7.8%	3,350	8.2%	0.000
All	6,927	16.9%	100.0%	33,984	83.1%	100.0%	40,911	100.0%	.

Appendix B. Participant #2

Table 9. Summary of Actions Taken[47]

	All Applications	Approved Applications		Denied Applications		Originated Loans		Delinquent Loans	
	Count	Count	Percent	Count	Percent	Count	Percent	Count	Percent[1]
All	212,949	58,524	27.48%	154,425	72.52%	40,911	19.21%	6,927	16.93%

46 Missing demographic data is the result of invalid zip codes, zip codes outside of the 50 States, or zip codes that do not have an associated ZCTA (Zip Code Tabulation Area).

47 The percentages in the delinquent loans column are calculated out of originated loans.

Appendix B. Participant #2 Table 10. Difference of Means Tests Within Demographic Group: Originated Loans[48]						
Variable	**Demographic Group**	**Sample**	**Count**	**Mean**	**T-Stat**	**P-Value**
Cash Flow Score	Originated Loans	Not Delinquent	33,984	680	.	.
		Delinquent	6,927	657	.	.
		All	40,911	676	39.261	0.000
	African American 75%	Not Delinquent	1,420	666	.	.
		Delinquent	483	643	9.123	0.000
	Hispanic 75%	Not Delinquent	2,496	675	.	.
		Delinquent	593	654	10.472	0.000
	Asian 75%	Not Delinquent	1,282	687	.	.
		Delinquent	254	670	6.464	0.000
	Non-Hispanic White 75%	Not Delinquent	19,671	682	.	.
		Delinquent	3,538	660	28.136	0.000
	Other or Missing BISG	Not Delinquent	9,115	677	.	.
		Delinquent	2,059	655	20.812	0.000
	Female	Not Delinquent	7,841	675	.	.
		Delinquent	1,752	652	18.599	0.000
	Male	Not Delinquent	22,443	682	.	.
		Delinquent	4,291	659	32.109	0.000
	Gender Unassigned	Not Delinquent	3,700	677	.	.
		Delinquent	884	656	12.235	0.000
FICO Score	Originated Loans	Not Delinquent	33,662	665	.	.
		Delinquent	6,868	637	.	.
		All	40,530	660	35.944	0.000
	African American 75%	Not Delinquent	1,406	645	.	.
		Delinquent	481	622	8.508	0.000
	Hispanic 75%	Not Delinquent	2,483	655	.	.
		Delinquent	591	631	10.214	0.000
	Asian 75%	Not Delinquent	1,258	675	.	.
		Delinquent	251	653	5.438	0.000
	Non-Hispanic White 75%	Not Delinquent	19,495	668	.	.
		Delinquent	3,514	641	25.094	0.000
	Other or Missing BISG	Not Delinquent	9,020	662	.	.
		Delinquent	2,031	635	19.400	0.000
	Female	Not Delinquent	7,775	656	.	.

48 T-tests assume unequal variances and are conducted on the delinquent and non-delinquent populations. Yellow highlighting indicates a difference between the delinquent and non-delinquent groups that is statistically significant at the 95% confidence level (P-value < 0.05). Highlighting is shown regardless of the direction of the difference. Counts displayed are the counts of non-missing values for each variable, by demographic group and status.

		Delinquent	1,740	635	13.242	0.000
	Male	Not Delinquent	22,234	668	.	.
		Delinquent	4,257	639	31.431	0.000
	Gender Unassigned	Not Delinquent	3,653	661	.	.
		Delinquent	871	636	11.631	0.000

Appendix B. Participant #2

Table 11. Logistic Model for Delinquency Results Within Demographic Group[49]

		FICO Score Only	Cash Flow Only	Cash Flow and FICO Score
Demographic Group	**Count**	**AUC**	**AUC**	**AUC**
Originated Loans	40,911	0.640	0.652	0.660
African American 75%	1,903	0.622	0.638	0.644
Hispanic 75%	3,089	0.633	0.640	0.652
Asian 75%	1,536	0.613	0.633	0.638
Non-Hispanic White 75%	23,209	0.641	0.651	0.659
Other or Missing BISG	11,174	0.635	0.649	0.657
Female	9,593	0.614	0.644	0.644
Male	26,734	0.652	0.657	0.670
Gender Unassigned	4,584	0.626	0.635	0.642

49 Models with a FICO Score control include a flag for missing values. The ROC analyses are restricted to the Race/Ethnicity or gender group listed and uses an indicator for "delinquent" as the reference variable and the listed score as the rating. The estimation samples may differ slightly from the displayed count based on missing values and perfect prediction among the set of predictor variables.

Appendix B. Participant #2
Table 12. Logistic Model Specification with FICO Score Within Race/Ethnicity Group

Control Variable	Comparison Group	African American 75%		Hispanic 75%		Asian 75%		Non-Hispanic White 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Score	--	.	.	.	.	.	.	.	.
FICO Score	--	0.99	0.00	0.99	0.00	0.99	0.00	0.99	0.00
Missing FICO	Not Missing FICO	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Constant		64.47	0.00	80.42	0.00	19.88	0.00	36.81	0.00
Pseudo R-Squared		0.031		0.033		0.023		0.032	
AUC		0.622		0.633		0.613		0.641	
Sample Size		1,903		3,089		1,536		23,209	

Appendix B. Participant #2
Table 13. Logistic Model Specification with FICO Score Within Gender Group

Control Variable	Comparison Group	Female		Male	
		Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Score	--	.	.	.	.
FICO Score	--	0.99	0.00	0.99	0.00
Missing FICO	Not Missing FICO	0.01	0.00	0.00	0.00
Constant		15.16	0.00	78.80	0.00
Pseudo R-Squared		0.021		0.040	
AUC		0.614		0.652	
Sample Size		9,593		26,734	

Appendix B. Participant #2 Table 14. Logistic Model Specification with Cash Flow Score Within Race/Ethnicity Group									
		African American 75%		Hispanic 75%		Asian 75%		Non-Hispanic White 75%	
Control Variable	**Comparison Group**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
Cash Flow Score	--	0.99	0.00	0.99	0.00	0.99	0.00	0.99	0.00
FICO Score	--	.	.	.	.	.	.	.	.
Missing FICO	Not Missing FICO	.	.	.	.	.	.	.	.
Constant		243.44	0.00	256.37	0.00	229.92	0.00	452.02	0.00
Pseudo R-Squared		0.038		0.035		0.028		0.040	
AUC		0.638		0.640		0.633		0.651	
Sample Size		1,903		3,089		1,536		23,209	

Appendix B. Participant #2 Table 15. Logistic Model Specification with Cash Flow Score Within Gender Group					
		Female		Male	
Control Variable	**Comparison Group**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
Cash Flow Score	--	0.99	0.00	0.99	0.00
FICO Score	--	.	.	.	.
Missing FICO	Not Missing FICO	.	.	.	.
Constant		283.81	0.00	587.28	0.00
Pseudo R-Squared		0.039		0.042	
AUC		0.644		0.657	
Sample Size		9,593		26,734	

Appendix B. Participant #2

Table 16. Logistic Model Specification with Cash Flow Score and FICO Score Within Race/Ethnicity Group

Control Variable	Comparison Group	African American 75%		Hispanic 75%		Asian 75%		Non-Hispanic White 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Score	--	0.99	0.00	0.99	0.00	0.99	0.00	0.99	0.00
FICO Score	--	1.00	0.00	0.99	0.00	1.00	0.00	1.00	0.00
Missing FICO	Not Missing FICO	0.04	0.00	0.02	0.00	0.04	0.00	0.05	0.00
Constant		487.23	0.00	692.22	0.00	651.36	0.00	757.45	0.00
Pseudo R-Squared		0.042		0.042		0.035		0.046	
AUC		0.644		0.652		0.638		0.659	
Sample Size		1,903		3,089		1,536		23,209	

Appendix B. Participant #2

Table 17. Logistic Model Specification with Cash Flow Score and FICO Score Within Gender Group

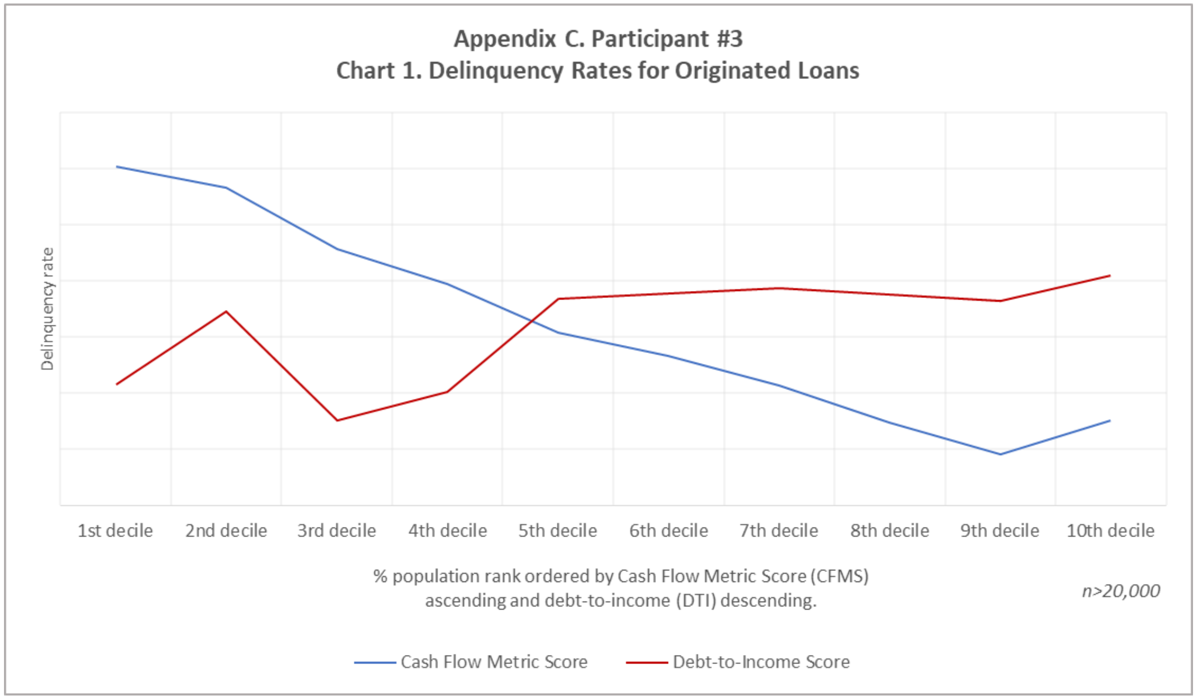
Control Variable	Comparison Group	Female		Male	
		Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Score	--	0.99	0.00	0.99	0.00
FICO Score	--	1.00	0.01	0.99	0.00
Missing FICO	Not Missing FICO	0.26	0.01	0.02	0.00
Constant		357.29	0.00	1,313.62	0.00
Pseudo R-Squared		0.040		0.052	
AUC		0.644		0.670	
Sample Size		9,593		26,734	

APPENDIX C: Participant 3[50]

Appendix C. Participant #3

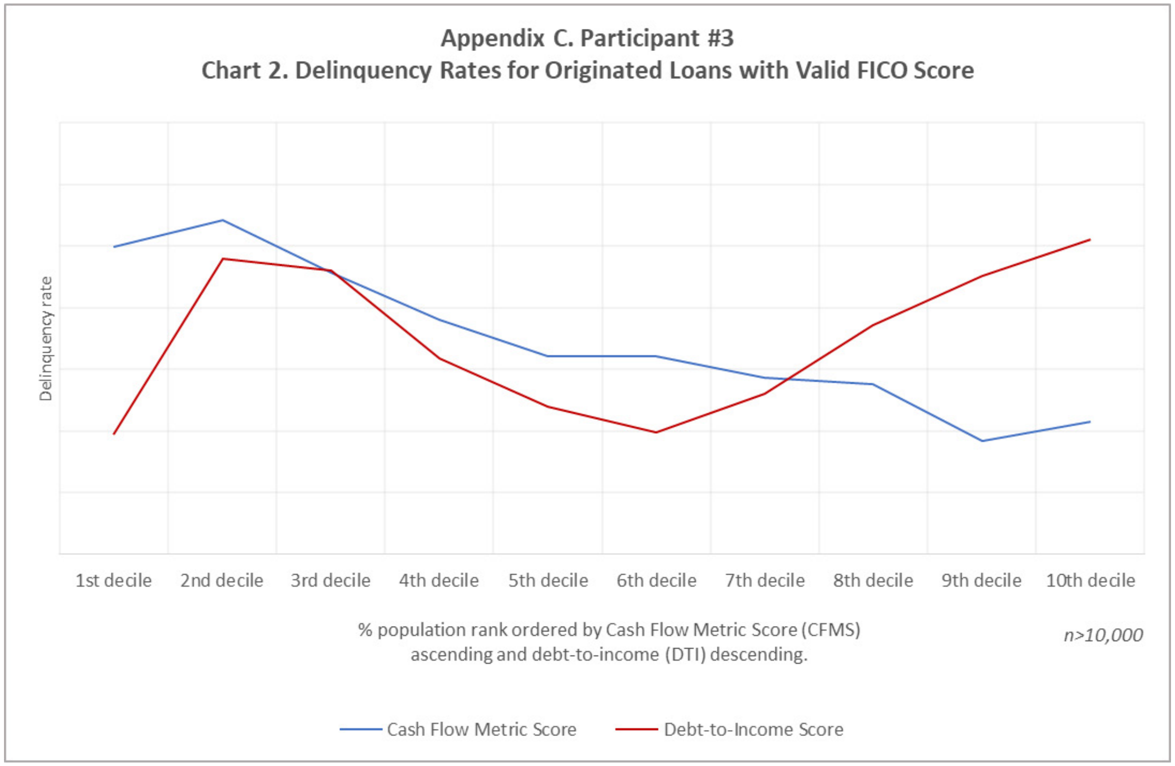
Chart 1.	Delinquency Rates for Originated Loans
Chart 2.	Delinquency Rates for Originated Loans With Valid FICO Score
Chart 3.	Delinquency Rates for Originated Loans Without Valid FICO Score
Chart 4.	Weighted Median Yearly Income

50 All of the Tables in Appendix C were created by Participant 3, and CRA has not validated the content.



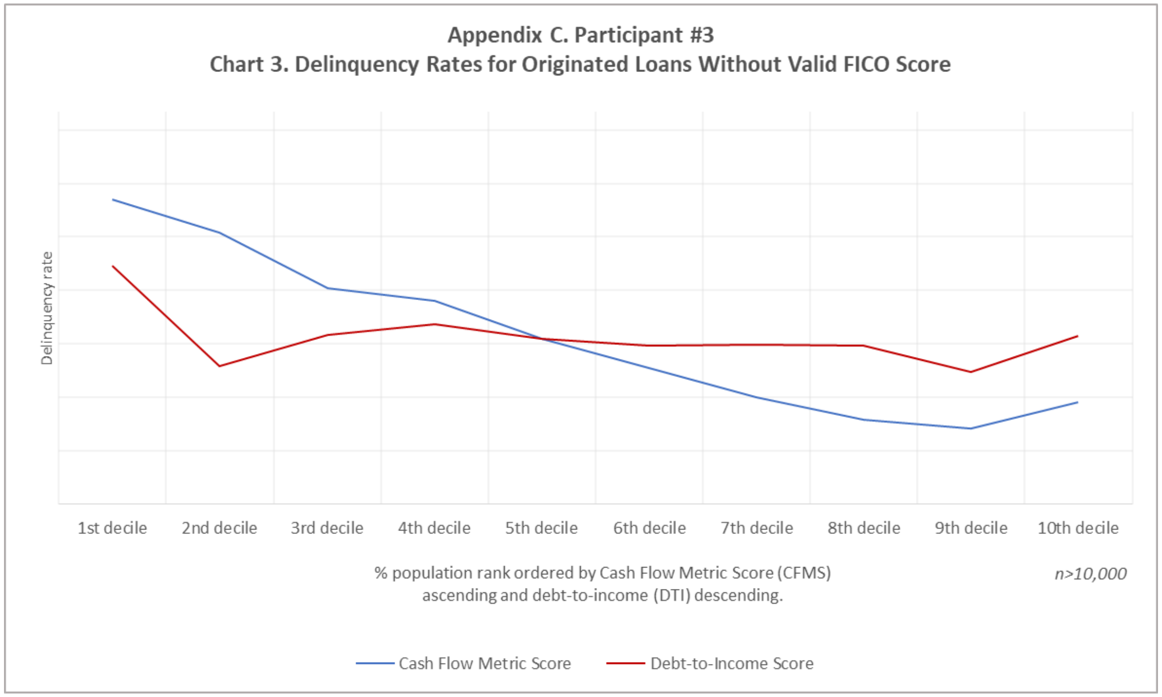
	Cash Flow Metric Score (CFMS)	Debt-to-income (DTI)
AUC	0.532	0.5125
95% confidence interval	± 0.002	± 0.002

Chart and Table created and reported by Participant 3



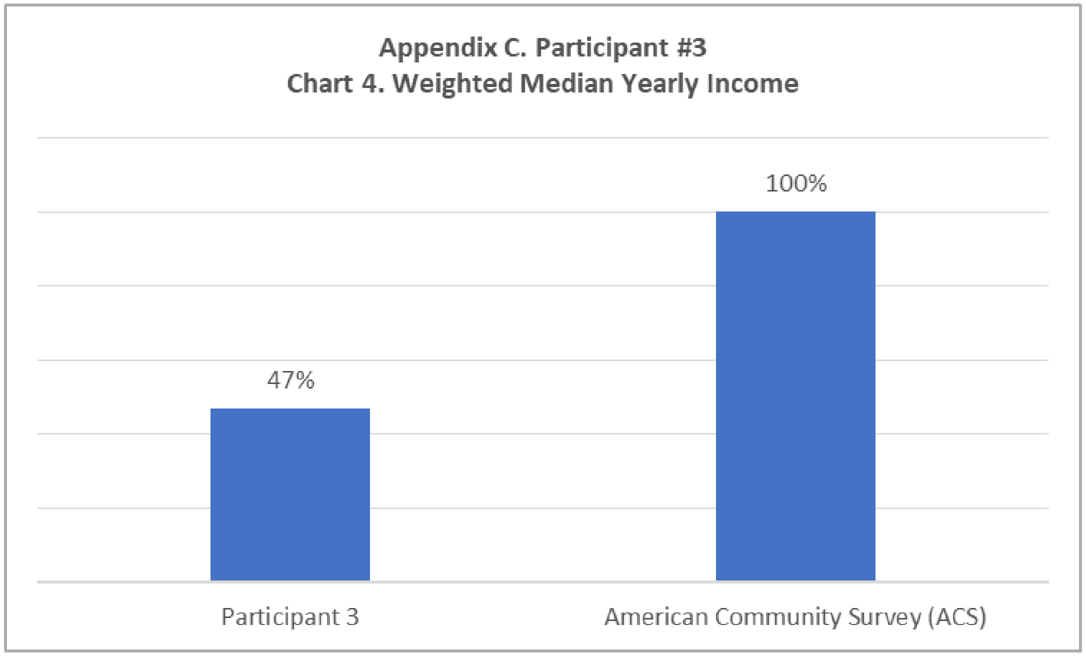
	Cash Flow Metric Score (CFMS)	Debt-to-income (DTI)
AUC	0.523	0.508
95% confidence interval	± 0.002	± 0.002

Chart and Table created and reported by Participant 3



	Cash Flow Metric Score (CFMS)	Debt-to-income (DTI)
AUC	0.537	0.507
95% confidence interval	± 0.002	± 0.002

Chart and Table created and reported by Participant 3



$$\text{Lender weighted median income} = \frac{\sum(\text{number of customers in zip code}_i * \text{ customer gross median income in zip code}_i)}{\sum(\text{number of customers in zip code}_i)}$$

$$\text{ACS weighted median income} = \frac{\sum(\text{number of customers living in zip code}_i * \text{ACS median income in zip code}_i)}{\sum(\text{number of customers inzip code}_i)}$$

Chart created and reported by Participant 3

APPENDIX D: Participant 4

Appendix D. Participant #4

Appendix D. Participant #4
Table 1. Data Diagnostics: All Applications

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Application Date	Approved	33,102	0	0	10/21/16	12/4/16	5/6/17	12/17/17	7/15/18	11/21/18	12/26/18	12/5/17
	Declined	53,161	0	0	10/18/16	12/13/16	7/5/17	2/9/18	7/16/18	11/16/18	12/27/18	1/6/18
	Other	25	0	0	12/29/16	2/5/17	11/4/17	11/10/17	11/27/17	12/6/18	12/19/18	11/27/17
	All	86,288	0	0	10/18/16	12/9/16	6/14/17	1/18/18	7/15/18	11/18/18	12/27/18	12/25/17
Fraud Score	Approved	33,102	12,887	0	159	484	608	686	754	827	949	675
	Declined	53,161	12,984	0	85	449	575	657	726	805	948	646
	Other	25	14	0	353	353	451	638	728	808	808	606
	All	86,288	25,885	0	85	459	585	664	736	814	949	656
Bank Behavior Score	Approved	33,102	12,690	0	164	591	708	772	819	907	975	761
	Declined	53,161	12,480	0	92	539	671	745	788	893	967	730
	Other	25	15	0	564	564	657	733	798	895	895	729
	All	86,288	25,185	0	92	554	684	756	799	900	975	740
Traditional Credit Probability #1	Approved	33,102	9,099	0	0.104	0.191	0.234	0.271	0.318	0.383	1.000	0.279
	Declined	53,161	53,046	0	0.165	0.195	0.243	0.287	0.335	0.392	0.446	0.289
	Other	25	21	0	0.258	0.258	0.277	0.308	0.335	0.349	0.349	0.306
	All	86,288	62,166	0	0.104	0.191	0.234	0.271	0.318	0.383	1.000	0.279
TPD	Approved	33,102	2,376	0	0.033	0.185	0.233	0.270	0.316	0.381	1.000	0.276
	Declined	53,161	28,192	0	0.102	0.225	0.295	0.354	0.404	0.444	0.761	0.347
	Other	25	15	0	0.258	0.258	0.279	0.302	0.337	0.444	0.444	0.314
	All	86,288	30,583	0	0.033	0.197	0.250	0.302	0.364	0.432	1.000	0.308
CFPD	Approved	33,102	0	0	0.119	0.203	0.250	0.288	0.324	0.373	0.630	0.287
	Declined	53,161	0	0	0.168	0.296	0.374	0.440	0.540	0.687	0.933	0.466
	Other	25	0	0	0.198	0.233	0.279	0.320	0.445	0.697	0.890	0.385
	All	86,288	0	0	0.119	0.226	0.298	0.368	0.470	0.639	0.933	0.397
Self-Reported Income	Approved	33,102	4,867	0	$1	$12,000	$22,000	$30,854	$45,000	$75,600	$10,000,000	$37,808
	Declined	53,161	714	0	$1	$10,000	$19,992	$28,000	$40,000	$68,000	$5,313,168	$32,723
	Other	25	15	0	$8,820	$8,820	$28,000	$33,500	$40,000	$75,000	$75,000	$35,775
	All	86,288	5,596	0	$1	$10,000	$20,000	$29,761	$40,000	$71,000	$10,000,000	$34,502
Number of Accounts	Approved	33,102	2	0	1	1	1	2	2	4	14	1.9
	Declined	53,161	507	0	1	1	1	2	2	4	21	1.8

	Other	25	0	0	1	1	1	1	2	4	4	1.8
	All	86,288	509	0	1	1	1	2	2	4	21	1.9
Cash Flow Metric #1	Approved	33,102	2,791	962	0	6	36	53	67	84	90	50
	Declined	53,161	17,513	347	0	31	61	75	85	90	90	70
	Other	25	5	0	13	13	32	53	73	86	90	51
	All	86,288	20,309	1,309	0	14	47	65	80	90	90	61
Cash Flow Metric #2	Approved	33,102	2,791	962	0%	7%	41%	60%	74%	93%	100%	57%
	Declined	53,161	17,513	347	0%	36%	71%	86%	96%	100%	100%	80%
	Other	25	5	0	14%	14%	36%	61%	83%	96%	100%	58%
	All	86,288	20,309	1,309	0%	17%	54%	74%	90%	100%	100%	69%
Cash Flow Metric #3	Approved	33,102	4	0	1	67	224	363	524	830	2,472	394
	Declined	53,161	523	3,591	0	0	35	123	281	575	5,208	186
	Other	25	0	1	0	3	101	377	610	980	1,460	416
	All	86,288	527	3,592	0	1	74	215	402	711	5,208	266
Cash Flow Metric #4	Approved	33,102	130	0	1	5	14	20	26	32	58	19
	Declined	53,161	1,167	3,554	0	0	4	10	18	28	52	12
	Other	25	0	1	0	1	12	21	30	34	35	19
	All	86,288	1,297	3,555	0	1	6	14	22	31	58	15
Cash Flow Metric #5	Approved	33,102	130	584	0	2	8	13	18	27	47	14
	Declined	53,161	1,167	11,567	0	0	1	3	9	18	46	6
	Other	25	0	3	0	0	5	14	22	30	30	14
	All	86,288	1,297	12,154	0	0	2	7	14	24	47	9
Cash Flow Metric #6	Approved	33,102	130	1,922	0	0	2	3	4	5	10	3
	Declined	53,161	4,721	18,170	0	0	0	1	2	4	9	1
	Other	25	1	5	0	0	1	3	4	6	6	3
	All	86,288	4,852	20,097	0	0	1	2	3	5	10	2
Cash Flow Metric #7	Approved	33,102	130	0	1	5	14	20	26	32	58	19
	Declined	53,161	1,167	3,554	0	0	4	10	18	28	52	12
	Other	25	0	1	0	1	12	21	30	35	35	19
	All	86,288	1,297	3,555	0	1	6	14	22	31	58	15

Appendix D. Participant #4 **Table 2. Difference of Means Tests: All Applications**[51]					
Variable	**Sample**	**#**	**Mean**	**T-Stat**	**P-Value**
Fraud Score	Approved	20,215	675	.	.
	Declined	40,177	646	7.63	0.000
Bank Behavior Score	Approved	20,412	761	.	.
	Declined	40,681	730	-0.65	0.516
Traditional Credit Probability #1	Approved	24,003	0.279	.	.
	Declined	115	0.289	-10.18	0.000
TPD	Approved	30,726	0.276	.	.
	Declined	24,969	0.347	-11.72	0.000
CFPD	Approved	33,102	0.287	.	.
	Declined	53,161	0.466	-18.80	0.000
Cash Flow Metric #1	Approved	30,311	50	.	.
	Declined	35,648	70	-13.71	0.000
Cash Flow Metric #2	Approved	30,311	56.6%	.	.
	Declined	35,648	79.9%	-14.08	0.000
Cash Flow Metric #3	Approved	33,098	394	.	.
	Declined	52,638	186	4.98	0.000
Cash Flow Metric #4	Approved	32,972	19.37	.	.
	Declined	51,994	11.52	5.22	0.000
Cash Flow Metric #5	Approved	32,972	13.57	.	.
	Declined	51,994	5.55	12.82	0.000
Cash Flow Metric #6	Approved	32,972	2.81	.	.
	Declined	48,440	1.29	12.17	0.000
Cash Flow Metric #7	Approved	32,972	19.36	.	.
	Declined	51,994	11.55	5.26	0.000

[51] The significance test tests the difference in means between the approved and declined populations using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% level. Counts in this table are of non-missing values of the indicated variable.

Appendix D. Participant #4
Table 3. Data Diagnostics: Originated Loans

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Application Date	Non-Default	20,885	0	0	10/21/16	12/8/16	6/2/17	12/29/17	7/7/18	10/24/18	12/12/18	12/11/17
	Default	3,931	0	0	10/26/16	12/7/16	6/9/17	2/5/18	7/15/18	10/16/18	12/5/18	12/21/17
	Default Unknown	1,137	0	0	11/30/16	11/24/18	12/5/18	12/12/18	12/18/18	12/24/18	12/26/18	12/7/18
	All	25,953	0	0	10/21/16	12/9/16	6/10/17	1/24/18	7/26/18	11/23/18	12/26/18	12/29/17
Fraud Score	Non-Default	20,885	6,534	0	162	485	610	686	754	826	949	675
	Default	3,931	1,227	0	159	467	587	668	737	813	929	659
	Default Unknown	1,137	67	0	292	542	654	722	783	853	915	712
	All	25,953	7,828	0	159	484	609	686	754	826	949	675
Bank Behavior Score	Non-Default	20,885	6,388	0	164	594	708	770	811	907	975	759
	Default	3,931	1,185	0	262	580	703	770	824	910	965	761
	Default Unknown	1,137	157	0	383	613	743	816	875	907	935	796
	All	25,953	7,730	0	164	591	708	772	820	907	975	762
Traditional Credit Probability #1	Non-Default	20,885	3,463	0	0.104	0.190	0.232	0.268	0.315	0.382	1.000	0.277
	Default	3,931	650	0	0.119	0.202	0.244	0.282	0.332	0.388	1.000	0.290
	Default Unknown	1,137	1,137	0	.	.	.	.	.	.	.	.
	All	25,953	5,250	0	0.104	0.192	0.234	0.270	0.318	0.383	1.000	0.279
TPD	Non-Default	20,885	0	0	0.033	0.184	0.231	0.267	0.313	0.380	1.000	0.273
	Default	3,931	0	0	0.090	0.198	0.242	0.280	0.328	0.385	0.480	0.285
	Default Unknown	1,137	5	0	0.092	0.183	0.242	0.276	0.320	0.378	0.449	0.281
	All	25,953	5	0	0.033	0.186	0.233	0.269	0.316	0.381	1.000	0.275
CFPD	Non-Default	20,885	0	0	0.124	0.202	0.249	0.286	0.323	0.372	0.630	0.286
	Default	3,931	0	0	0.120	0.219	0.267	0.305	0.337	0.386	0.498	0.303
	Default Unknown	1,137	0	0	0.152	0.199	0.246	0.284	0.318	0.376	0.428	0.284
	All	25,953	0	0	0.120	0.203	0.251	0.289	0.325	0.375	0.630	0.288
Self-Reported Income	Non-Default	20,885	727	0	$1	$12,000	$22,000	$31,000	$45,000	$76,000	$4,200,000	$37,311
	Default	3,931	169	0	$20	$12,000	$22,000	$30,000	$44,000	$75,000	$10,000,000	$39,768
	Default Unknown	1,137	37	0	$2,000	$12,000	$24,000	$34,000	$50,000	$80,000	$208,000	$38,932
	All	25,953	933	0	$1	$12,000	$22,000	$31,000	$45,000	$76,000	$10,000,000	$37,752
Number of Accounts	Non-Default	20,885	1	0	1	1	1	2	2	4	13	1.9
	Default	3,931	0	0	1	1	1	2	2	4	10	1.9
	Default Unknown	1,137	0	0	1	1	1	2	2	4	10	1.9

	All	25,953	1	0	1	1	1	2	2	4	13	1.9
Cash Flow Metric #1	Non-Default	20,885	1,765	617	0	6	35	52	66	83	90	50
	Default	3,931	396	99	0	9	42	59	72	87	90	55
	Default Unknown	1,137	63	10	0	14	37	53	67	84	90	51
	All	25,953	2,224	726	0	7	36	53	67	84	90	50
Cash Flow Metric #2	Non-Default	20,885	1,765	617	0%	7%	40%	59%	74%	93%	100%	56%
	Default	3,931	396	99	0%	10%	48%	67%	80%	98%	100%	62%
	Default Unknown	1,137	63	10	0%	16%	41%	59%	76%	94%	100%	57%
	All	25,953	2,224	726	0%	8%	41%	60%	74%	93%	100%	57%
Cash Flow Metric #3	Non-Default	20,885	2	0	1	71	228	368	530	829	2,421	397
	Default	3,931	0	0	1	45	201	338	510	833	2,472	376
	Default Unknown	1,137	0	0	1	91	279	432	607	938	1,439	460
	All	25,953	2	0	1	66	225	367	530	835	2,472	397
Cash Flow Metric #4	Non-Default	20,885	94	0	1	5	14	20	26	32	52	20
	Default	3,931	12	0	1	4	12	19	25	32	58	19
	Default Unknown	1,137	0	0	1	7	16	22	28	34	47	22
	All	25,953	106	0	1	5	14	20	26	32	58	19
Cash Flow Metric #5	Non-Default	20,885	94	313	0	2	8	14	19	27	47	14
	Default	3,931	12	116	0	1	7	12	17	26	46	12
	Default Unknown	1,137	0	18	0	3	9	14	20	27	47	15
	All	25,953	106	447	0	2	8	13	18	27	47	14
Cash Flow Metric #6	Non-Default	20,885	94	1,079	0	0	2	3	4	5	10	3
	Default	3,931	12	353	0	0	2	3	3	5	9	3
	Default Unknown	1,137	0	64	0	0	2	3	4	5	9	3
	All	25,953	106	1,496	0	0	2	3	4	5	10	3
Cash Flow Metric #7	Non-Default	20,885	94	0	1	5	14	20	26	32	53	20
	Default	3,931	12	0	1	4	12	19	25	32	58	19
	Default Unknown	1,137	0	0	1	7	16	22	28	34	47	22
	All	25,953	106	0	1	5	14	20	26	32	58	19

Appendix D. Participant #4

Table 4. Difference of Means Tests: Originated Loans[52]

Variable	Sample	#	Mean	T-Stat	P-Value
Fraud Score	Non-Default	14,351	675	.	.
	Default	2,704	659	7.63	0.000
Bank Behavior Score	Non-Default	14,497	759	.	.
	Default	2,746	761	-0.65	0.516
Traditional Credit Probability #1	Non-Default	17,422	0.277	.	.
	Default	3,281	0.290	-10.18	0.000
TPD	Non-Default	20,885	0.273	.	.
	Default	3,931	0.285	-11.72	0.000
CFPD	Non-Default	20,885	0.286	.	.
	Default	3,931	0.303	-18.80	0.000
Cash Flow Metric #1	Non-Default	19,120	49.6	.	.
	Default	3,535	55.2	-13.71	0.000
Cash Flow Metric #2	Non-Default	19,120	55.6%	.	.
	Default	3,535	62.0%	-14.08	0.000
Cash Flow Metric #3	Non-Default	20,883	397	.	.
	Default	3,931	376	4.98	0.000
Cash Flow Metric #4	Non-Default	20,791	19.52	.	.
	Default	3,919	18.73	5.22	0.000
Cash Flow Metric #5	Non-Default	20,791	13.86	.	.
	Default	3,919	12.21	12.82	0.000
Cash Flow Metric #6	Non-Default	20,791	2.87	.	.
	Default	3,919	2.55	12.17	0.000
Cash Flow Metric #7	Non-Default	20,791	19.51	.	.
	Default	3,919	18.71	5.26	0.000

52 The significance test tests the difference in means between the default and non-default populations using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% level. Counts in this table are of non-missing values of the indicated variable.

Appendix D. Participant #4

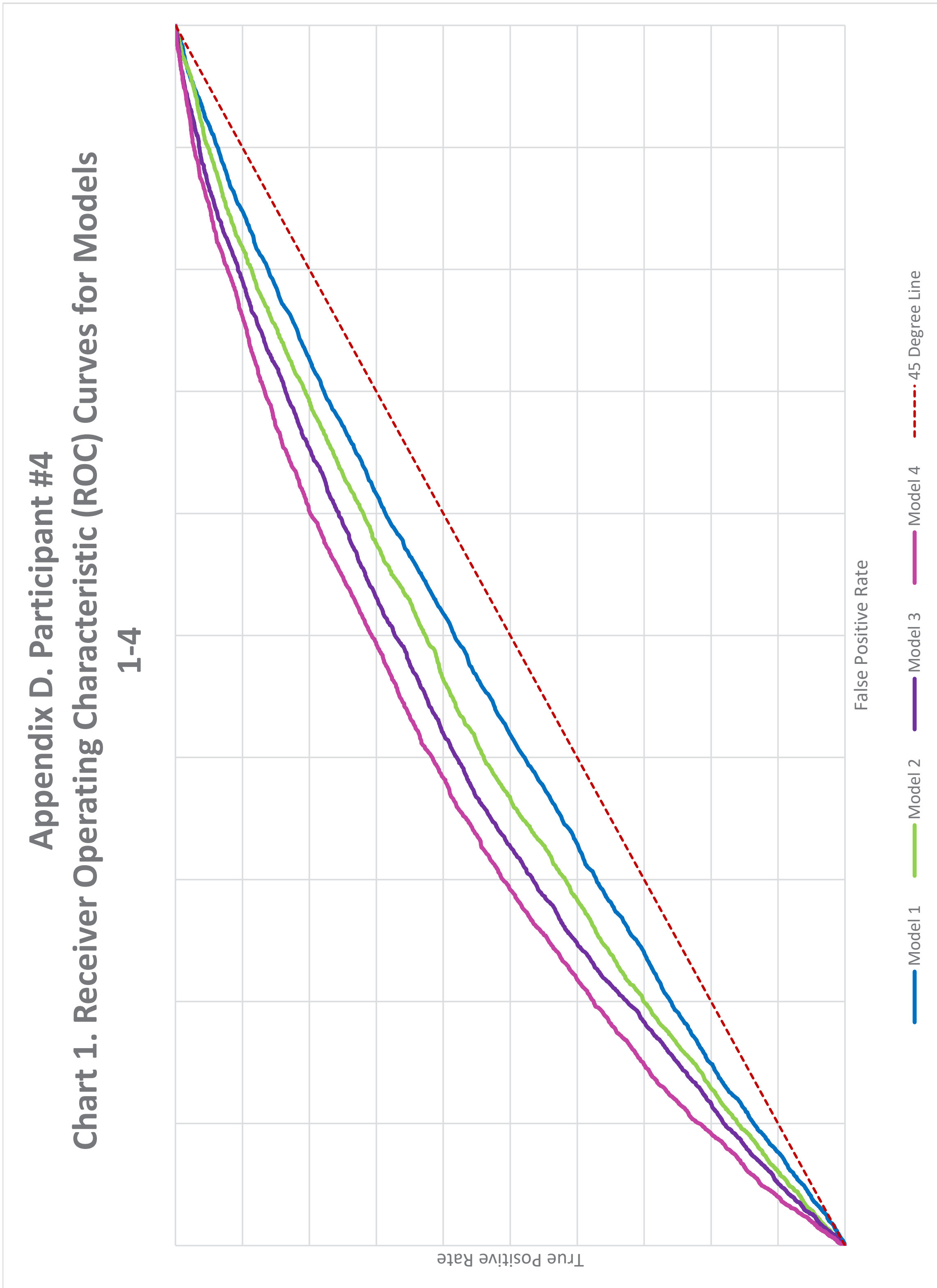
Table 5. Logistic Model for Default Specifications[53]

Variable	Comparison Group	TPD Model (Model 1)		CFPD Model (Model 2)		Combined Model (Model 3)		Exhaustive Model (Model 4)	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	--	26.14	0.000	.	.	117.01	0.000	81.98	0.000
CFPD	--	.	.	491.15	0.000	1,879.25	0.000	1,046.59	0.000
Fraud Score	--	.	.	.	.	.	.	1.00	0.000
Missing Fraud Score	Not Missing Fraud Score	.	.	.	.	.	.	0.40	0.000
Bank Behavior Score	--	.	.	.	.	.	.	1.00	0.942
Missing Bank Behavior Score	Not Missing Bank Behavior Score	.	.	.	.	.	.	1.02	0.935
Self-Reported Income	--	.	.	.	.	.	.	1.00	0.045
Missing Self-Reported Income	Not Missing Self-Reported Income	.	.	.	.	.	.	1.30	0.006
Number of Accounts	--	.	.	.	.	.	.	0.92	0.000
Missing Number of Accounts	Not Missing Number of Account	.	.	.	.	.	.	.	.
Cash Flow Metric #1	--	.	.	.	.	.	.	1.00	0.012
Missing Cash Flow Metric #1	Not Missing Cash Flow Balance #1	.	.	.	.	.	.	0.00	0.011
Cash Flow Metric #3	--	.	.	.	.	.	.	1.00	0.000
Cash Flow Metric #4	--	.	.	.	.	.	.	0.99	0.724
Cash Flow Metric #5	--	.	.	.	.	.	.	0.99	0.007
Cash Flow Metric #6	--	.	.	.	.	.	.	0.96	0.028

53 The dependent variable is a 0/1 indicator for default, with values of 1 indicating default and 0 indicating no default. This table only contains originations with a known default status. Percentiles are based on the population of originated loans with a known empirical default status.

Cash Flow Metric #7	--	.	.	.	.	.	.	1.01	0.783
Missing Cash Flow Metric #7	Not Missing Cash Flow Metric #7	.	.	.	.	.	.	.	.
Source Category #2	Source Category #1	.	.	.	.	.	.	0.97	0.908
Source Category #3		.	.	.	.	.	.	0.87	0.082
Source Category #4		.	.	.	.	.	.	1.08	0.525
Source Category #5		.	.	.	.	.	.	2.03	0.145
Source Category #6		.	.	.	.	.	.	.	.
Source Category #7		.	.	.	.	.	.	1.11	0.128
Source Category #8		.	.	.	.	.	.	0.88	0.625
Source Category #9		.	.	.	.	.	.	1.13	0.110
Source Category #10		.	.	.	.	.	.	1.13	0.500
Source Category #11		.	.	.	.	.	.	1.20	0.386
Source Category #12		.	.	.	.	.	.	1.42	0.018
State #2	State #1	.	.	.	.	.	.	0.81	0.258
State #3		.	.	.	.	.	.	0.63	0.020
State #4		.	.	.	.	.	.	0.36	0.169
State #5		.	.	.	.	.	.	1.30	0.565
State #6		.	.	.	.	.	.	0.59	0.065
State #7		.	.	.	.	.	.	0.52	0.018
State #8		.	.	.	.	.	.	0.86	0.643
State #9		.	.	.	.	.	.	0.58	0.007
State #10		.	.	.	.	.	.	0.79	0.601
State #11		.	.	.	.	.	.	0.77	0.249
State #12		.	.	.	.	.	.	0.75	0.170
State #13		.	.	.	.	.	.	0.75	0.201
State #14		.	.	.	.	.	.	0.47	0.063
State #15		.	.	.	.	.	.	0.96	0.832
State #16		.	.	.	.	.	.	0.79	0.495
State #17		.	.	.	.	.	.	0.97	0.905
State #18		.	.	.	.	.	.	0.75	0.140
State #19		.	.	.	.	.	.	0.68	0.037
State #20		.	.	.	.	.	.	0.52	0.105

State #21		.	.	.	.	.	.	0.56	0.116
State #22		.	.	.	.	.	.	0.88	0.528
State #23		.	.	.	.	.	.	1.29	0.704
Application Date: Month #2	Application Date: Month #1	.	.	.	.	.	.	1.60	0.297
Application Date: Month #3		.	.	.	.	.	.	1.26	0.602
Application Date: Month #4		.	.	.	.	.	.	1.26	0.604
Application Date: Month #5		.	.	.	.	.	.	0.99	0.982
Application Date: Month #6		.	.	.	.	.	.	1.33	0.527
Application Date: Month #7		.	.	.	.	.	.	1.38	0.478
Application Date: Month #8		.	.	.	.	.	.	1.28	0.581
Application Date: Month #9		.	.	.	.	.	.	1.72	0.226
Application Date: Month #10		.	.	.	.	.	.	1.40	0.463
Application Date: Month #11		.	.	.	.	.	.	1.24	0.635
Application Date: Month #12		.	.	.	.	.	.	1.42	0.450
Application Date: Month #13		.	.	.	.	.	.	1.13	0.795
Application Date: Month #14		.	.	.	.	.	.	1.03	0.942
Application Date: Month #15		.	.	.	.	.	.	1.26	0.613
Application Date: Month #16		.	.	.	.	.	.	0.97	0.942
Application Date: Month #17		.	.	.	.	.	.	0.96	0.933
Application Date: Month #18		.	.	.	.	.	.	1.49	0.385
Application Date: Month #19		.	.	.	.	.	.	1.56	0.333
Application Date: Month #20		.	.	.	.	.	.	1.65	0.271
Application Date: Month #21		.	.	.	.	.	.	1.62	0.289
Application Date: Month #22		.	.	.	.	.	.	1.76	0.214
Application Date: Month #23		.	.	.	.	.	.	1.38	0.481
Application Date: Month #24		.	.	.	.	.	.	1.33	0.547
Application Date: Month #25		.	.	.	.	.	.	1.17	0.736
Application Date: Month #26		.	.	.	.	.	.	0.85	0.736
Application Date: Month #27		.	.	.	.	.	.	0.50	0.294
Constant	-	0.08	0.000	0.03	0.000	0.01	0.000	0.02	0.000
Pseudo R-Squared		0.006		0.016		0.027		0.043	
AUC		0.559		0.592		0.620		0.650	
Num. of Observations		24,816		24,816		24,816		24,709	


Appendix D. Participant #4
Chart 1. Receiver Operating Characteristic (ROC) Curves for Models 1-4
True Positive Rate
False Positive Rate
Model 1
Model 2
Model 3
Model 4
45 Degree Line

Appendix D. Participant #4

Table 6. Default Frequency by CFPD Percentile and TPD Percentile[54]

Traditional Probability of Default	Cash Flow Based Probability of Default 95 - 100th	90 - 95th	85 - 90th	80 - 85th	75 - 80th	70 - 75th	65 - 70th	60 - 65th	55 - 60th	50 - 55th	45 - 50th	40 - 45th	35 - 40th	30 - 35th	25 - 30th	20 - 25th	15 - 20th	10 - 15th	5 - 10th	0 - 5th
95 - 100th	33.3	23.1	36.4	20.0	35.4	27.0	19.0	27.9	32.5	13.5	16.2	10.6	23.2	21.1	12.8	16.3	21.3	14.3	10.1	12.8
90 - 95th	28.6	40.0	36.6	39.6	28.3	32.5	20.5	22.2	18.8	18.3	24.7	22.8	23.9	23.5	11.1	22.8	18.4	17.0	4.9	20.5
85 - 90th	20.0	.	22.9	24.1	30.0	18.0	28.6	33.9	16.4	25.3	25.6	20.4	14.3	18.1	21.1	22.2	13.9	13.4	11.8	12.7
80 - 85th	.	.	22.5	27.3	21.8	23.3	18.3	17.3	28.1	22.7	11.5	28.0	16.9	11.6	11.3	7.5	11.7	12.5	16.9	15.9
75 - 80th	.	.	23.4	20.0	25.4	23.7	20.7	17.6	14.8	20.5	16.2	13.9	8.8	13.8	16.2	10.7	13.2	16.7	11.3	6.1
70 - 75th	0.0	17.4	26.8	29.7	16.5	15.4	23.1	19.3	31.9	6.9	15.6	18.6	25.0	22.4	15.4	21.7	10.0	9.5	9.7	12.5
65 - 70th	.	9.1	25.8	21.9	17.2	22.9	20.5	20.8	11.4	18.7	9.5	10.9	9.4	16.0	16.7	20.0	21.6	8.7	8.3	2.0
60 - 65th	21.1	25.8	16.5	36.1	25.4	20.8	25.0	26.3	15.7	4.8	10.3	20.0	9.4	17.7	18.5	10.5	7.7	10.2	13.7	7.6
55 - 60th	25.9	27.8	21.5	20.8	27.1	12.3	20.9	28.2	19.7	14.7	16.4	18.1	14.3	14.5	11.3	10.2	8.9	7.3	15.4	12.3
50 - 55th	25.5	20.5	25.4	21.6	15.2	20.7	17.8	17.5	6.5	20.7	9.4	16.7	10.0	20.0	8.1	11.7	14.8	6.3	4.2	5.1
45 - 50th	20.3	32.3	24.7	18.2	20.8	18.6	22.7	22.2	12.3	9.1	10.8	9.1	16.4	10.8	19.7	14.0	10.2	3.6	7.0	3.3
40 - 45th	34.6	20.0	28.2	30.0	19.7	15.8	20.6	17.3	10.5	8.8	5.1	21.3	15.6	8.2	13.5	9.3	10.6	7.7	8.0	11.4
35 - 40th	23.4	16.8	20.4	22.4	15.5	12.5	15.2	18.6	15.7	12.1	14.8	17.5	5.9	14.5	10.7	13.0	8.5	8.8	12.8	8.8
30 - 35th	30.5	24.8	18.2	21.5	19.3	13.8	15.3	14.5	11.5	8.9	17.9	14.1	10.5	7.4	14.3	10.3	12.2	7.7	3.4	3.3
25 - 30th	19.0	24.2	15.1	15.9	19.1	19.0	7.4	16.7	15.6	1.7	10.0	17.2	16.0	11.5	13.9	12.5	7.7	7.1	1.8	13.7
20 - 25th	21.3	21.6	15.4	9.7	12.1	13.6	17.6	16.9	16.9	8.0	11.8	13.2	17.2	11.5	13.4	9.3	5.8	7.4	3.9	6.1
15 - 20th	27.7	25.6	19.4	5.9	15.5	16.9	8.3	7.3	11.1	11.1	18.4	14.3	5.3	7.5	7.8	10.0	3.7	8.6	6.0	1.5
10 - 15th	20.6	21.1	16.7	14.0	4.4	10.7	13.6	13.0	6.3	7.2	11.7	19.6	7.8	5.5	10.5	9.1	9.4	13.1	4.8	7.1
5 - 10th	21.1	16.0	8.3	22.2	17.6	13.6	6.0	13.2	4.4	14.8	9.3	8.1	12.7	7.1	6.7	12.3	5.7	4.3	3.2	5.8
0 - 5th	19.6	18.8	16.7	13.3	7.5	6.9	11.3	12.0	12.9	7.0	4.0	12.5	6.0	2.5	7.7	4.9	3.4	4.3	2.9	3.9

54 Cells are shaded based on values. Green indicates values close to the lowest default frequency, yellow indicates values close to the median default frequency, and red indicates values close to the highest default frequency. Cells with fewer than 5 loans are excluded from this heat map. Percentiles are based on the population of originated loans with a known empirical default status.

Appendix D. Participant #4

Table 7. Summary of Actions Taken[55]

	All Applications	Approved Applications		Declined Applications		Other Applications		Originated Loans		Defaulted Loans	
	Count	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
All	86,288	33,102	38.36%	53,161	61.61%	25	0.03%	24,816	28.76%	3,931	15.84%

55 The percentages in the "Defaulted Loans" column are calculated out of originated loans.

Appendix D. Participant #4

Table 8. Difference of Means Tests Within Demographic Group: Originated Loans[56]

Variable	Demographic Group	Status	Count	Mean	T-Stat	P-Value
Fraud Score	Originated Loans	Default	2,704	658.7	.	.
		No Default	14,351	675.5	.	.
		All	17,055	672.8	7.6	0.000
	African American 75%	Default	326	649.4	.	.
		No Default	1,435	659.6	1.6	0.110
	Hispanic 75%	Default	646	661.1	.	.
		No Default	3,962	681.3	4.4	0.000
	Asian 75%	Default	57	691.5	.	.
		No Default	342	706.5	1.1	0.269
	Non-Hispanic White 75%	Default	605	660.6	.	.
		No Default	3,280	673.0	2.6	0.008
	Other or Missing BISG	Default	1,070	657.2	.	.
		No Default	5,332	674.9	5.1	0.000
	Female 75%	Default	1,336	652.6	.	.
		No Default	7,286	670.2	5.5	0.000
	Male 75%	Default	1,124	667.3	.	.
		No Default	5,832	683.1	4.7	0.000
	Gender Probabilities < 75% or Missing	Default	244	652.2	.	.
		No Default	1,233	670.3	2.6	0.010
Bank Behavior Score	Originated Loans	Default	2,746	760.8	.	.
		No Default	14,497	759.5	.	.
		All	17,243	759.7	-0.6	0.516
	African American 75%	Default	338	747.8	.	.
		No Default	1,459	746.5	-0.2	0.832
	Hispanic 75%	Default	647	776.7	.	.
		No Default	3,978	770.4	-1.7	0.087
	Asian 75%	Default	60	763.6	.	.
		No Default	352	766.0	0.2	0.866
	Non-Hispanic White 75%	Default	605	758.9	.	.
		No Default	3,301	754.7	-0.9	0.392

56 This table is restricted to originated loans with a known default status. T-tests assume unequal variances and are conducted on the population that defaulted and the population that did not default. Yellow highlighting indicates a difference between the default and no default groups that is statistically significant at the 95% confidence level (P-value < 0.05). Highlighting is shown regardless of the direction of the difference. Counts displayed are the counts of non-missing values for each variable, by demographic group and status.

	Other or Missing BISG	Default	1,096	756.4	.	.
		No Default	5,407	757.3	0.3	0.775
	Female 75%	Default	1,350	757.6	.	.
		No Default	7,392	758.5	0.3	0.760
	Male 75%	Default	1,141	765.0	.	.
		No Default	5,864	761.6	-1.1	0.269
	Gender Probabilities < 75% or Missing	Default	255	759.3	.	.
		No Default	1,241	754.8	-0.7	0.514
Traditional Credit Probability #1	Originated Loans	Default	3,281	0.290	.	.
		No Default	17,422	0.277	.	.
		All	20,703	0.279	-10.2	0.000
	African American 75%	Default	394	0.295	.	.
		No Default	1,846	0.276	-4.5	0.000
	Hispanic 75%	Default	700	0.287	.	.
		No Default	4,221	0.280	-2.8	0.005
	Asian 75%	Default	78	0.290	.	.
		No Default	393	0.280	-1.4	0.166
	Non-Hispanic White 75%	Default	824	0.288	.	.
		No Default	4,386	0.273	-5.4	0.000
	Other or Missing BISG	Default	1,285	0.290	.	.
		No Default	6,576	0.277	-7.1	0.000
	Female 75%	Default	1,619	0.292	.	.
		No Default	8,896	0.276	-8.0	0.000
	Male 75%	Default	1,338	0.288	.	.
		No Default	6,997	0.277	-5.8	0.000
	Gender Probabilities < 75% or Missing	Default	324	0.288	.	.
		No Default	1,529	0.279	-2.5	0.013
TPD	Originated Loans	Default	3,931	0.285	.	.
		No Default	20,885	0.273	.	.
		All	24,816	0.275	-11.7	0.000
	African American 75%	Default	468	0.289	.	.
		No Default	2,126	0.274	-4.8	0.000
	Hispanic 75%	Default	877	0.282	.	.
		No Default	5,317	0.275	-3.6	0.000
	Asian 75%	Default	86	0.289	.	.
		No Default	493	0.277	-1.9	0.063
	Non-Hispanic White 75%	Default	939	0.284	.	.
		No Default	5,069	0.270	-6.2	0.000
	Other or Missing BISG	Default	1,561	0.287	.	.
		No Default	7,880	0.274	-8.0	0.000
	Female 75%	Default	1,924	0.287	.	.

		No Default	10,667	0.273	-9.2	0.000
	Male 75%	Default	1,623	0.283	.	.
		No Default	8,402	0.273	-6.5	0.000
	Gender Probabilities < 75% or Missing	Default	384	0.286	.	.
		No Default	1,816	0.275	-3.4	0.001
CFPD	Originated Loans	Default	3,931	0.303	.	.
		No Default	20,885	0.286	.	.
		All	24,816	0.289	-18.8	0.000
	African American 75%	Default	468	0.305	.	.
		No Default	2,126	0.290	-5.8	0.000
	Hispanic 75%	Default	877	0.304	.	.
		No Default	5,317	0.285	-10.1	0.000
	Asian 75%	Default	86	0.303	.	.
		No Default	493	0.289	-2.4	0.017
	Non-Hispanic White 75%	Default	939	0.303	.	.
		No Default	5,069	0.285	-10.3	0.000
	Other or Missing BISG	Default	1,561	0.301	.	.
		No Default	7,880	0.287	-10.3	0.000
	Female 75%	Default	1,924	0.302	.	.
		No Default	10,667	0.287	-12.1	0.000
	Male 75%	Default	1,623	0.304	.	.
		No Default	8,402	0.285	-13.8	0.000
	Gender Probabilities < 75% or Missing	Default	384	0.301	.	.
		No Default	1,816	0.288	-4.5	0.000
Self-Reported Income	Originated Loans	Default	3,762	$39,768	.	.
		No Default	20,158	$37,311	.	.
		All	23,920	$37,698	-0.9	0.384
	African American 75%	Default	449	$33,197	.	.
		No Default	2,036	$33,021	-0.1	0.913
	Hispanic 75%	Default	839	$48,804	.	.
		No Default	5,136	$35,014	-1.1	0.265
	Asian 75%	Default	81	$38,693	.	.
		No Default	484	$41,078	0.7	0.459
	Non-Hispanic White 75%	Default	893	$39,389	.	.
		No Default	4,900	$39,375	0.0	0.993
	Other or Missing BISG	Default	1,500	$36,965	.	.
		No Default	7,602	$38,443	1.4	0.176
	Female 75%	Default	1,840	$40,276	.	.
		No Default	10,294	$34,461	-1.0	0.305
	Male 75%	Default	1,554	$39,435	.	.
		No Default	8,113	$41,550	2.1	0.036

	Gender Probabilities < 75% or Missing	Default	368	$38,635	.	.
		No Default	1,751	$34,428	-1.5	0.147
Number of Accounts	Originated Loans	Default	3,931	1.9	.	.
		No Default	20,884	1.9	.	.
		All	24,815	1.9	3.6	0.000
	African American 75%	Default	468	1.9	.	.
		No Default	2,125	1.9	-0.4	0.687
	Hispanic 75%	Default	877	1.8	.	.
		No Default	5,317	1.9	2.3	0.020
	Asian 75%	Default	86	2.1	.	.
		No Default	493	2.0	-0.5	0.618
	Non-Hispanic White 75%	Default	939	1.8	.	.
		No Default	5,069	1.9	2.6	0.010
	Other or Missing BISG	Default	1,561	1.9	.	.
		No Default	7,880	1.9	2.6	0.010
	Female 75%	Default	1,924	1.9	.	.
		No Default	10,666	1.9	3.1	0.002
	Male 75%	Default	1,623	1.8	.	.
		No Default	8,402	1.9	1.6	0.112
	Gender Probabilities < 75% or Missing	Default	384	1.9	.	.
		No Default	1,816	2.0	1.4	0.155
Cash Flow Metric #1	Originated Loans	Default	3,535	55.2	.	.
		No Default	19,120	49.6	.	.
		All	22,655	50.5	-13.7	0.000
	African American 75%	Default	415	57.8	.	.
		No Default	1,918	51.8	-5.3	0.000
	Hispanic 75%	Default	784	55.0	.	.
		No Default	4,944	50.1	-5.8	0.000
	Asian 75%	Default	80	56.6	.	.
		No Default	451	47.5	-3.5	0.001
	Non-Hispanic White 75%	Default	859	54.8	.	.
		No Default	4,619	48.0	-8.5	0.000
	Other or Missing BISG	Default	1,397	54.6	.	.
		No Default	7,188	49.7	-7.2	0.000
	Female 75%	Default	1,732	55.0	.	.
		No Default	9,784	50.2	-8.4	0.000
	Male 75%	Default	1,452	54.6	.	.
		No Default	7,694	48.7	-9.1	0.000
	Gender Probabilities < 75% or Missing	Default	351	58.1	.	.
		No Default	1,642	50.1	-6.5	0.000
	Originated Loans	Default	3,535	61.96%	.	.

Cash Flow Metric #2		No Default	19,120	55.57%	.	.
		All	22,655	56.57%	-14.1	0.000
	African American 75%	Default	415	64.71%	.	.
		No Default	1,918	57.99%	-5.3	0.000
	Hispanic 75%	Default	784	61.78%	.	.
		No Default	4,944	56.14%	-6.0	0.000
	Asian 75%	Default	80	63.30%	.	.
		No Default	451	53.41%	-3.4	0.001
	Non-Hispanic White 75%	Default	859	61.75%	.	.
		No Default	4,619	53.89%	-8.8	0.000
	Other or Missing BISG	Default	1,397	61.30%	.	.
		No Default	7,188	55.75%	-7.4	0.000
	Female 75%	Default	1,732	61.74%	.	.
		No Default	9,784	56.24%	-8.6	0.000
	Male 75%	Default	1,452	61.51%	.	.
		No Default	7,694	54.59%	-9.6	0.000
	Gender Probabilities < 75% or Missing	Default	351	64.91%	.	.
		No Default	1,642	56.15%	-6.4	0.000
Cash Flow Metric #3	Originated Loans	Default	3,931	376.1	.	.
		No Default	20,883	397.4	.	.
		All	24,814	394.0	5.0	0.000
	African American 75%	Default	468	346.6	.	.
		No Default	2,125	363.5	1.5	0.126
	Hispanic 75%	Default	877	375.3	.	.
		No Default	5,316	399.7	2.7	0.006
	Asian 75%	Default	86	356.8	.	.
		No Default	493	389.4	1.2	0.220
	Non-Hispanic White 75%	Default	939	398.1	.	.
		No Default	5,069	413.8	1.7	0.086
	Other or Missing BISG	Default	1,561	373.2	.	.
		No Default	7,880	394.9	3.1	0.002
	Female 75%	Default	1,924	367.0	.	.
		No Default	10,665	380.5	2.3	0.022
	Male 75%	Default	1,623	388.5	.	.
		No Default	8,402	423.1	4.9	0.000
	Gender Probabilities < 75% or Missing	Default	384	369.3	.	.
		No Default	1,816	377.6	0.7	0.515
Cash Flow Metric #4	Originated Loans	Default	3,919	18.7	.	.
		No Default	20,791	19.5	.	.
		All	24,710	19.4	5.2	0.000
	African American 75%	Default	465	18.4	.	.

		No Default	2,113	19.0	1.4	0.171
	Hispanic 75%	Default	875	18.7	.	.
		No Default	5,295	20.2	4.6	0.000
	Asian 75%	Default	86	18.5	.	.
		No Default	492	19.4	1.0	0.330
	Non-Hispanic White 75%	Default	938	19.1	.	.
		No Default	5,046	19.2	0.1	0.900
	Other or Missing BISG	Default	1,555	18.6	.	.
		No Default	7,845	19.4	3.5	0.001
	Female 75%	Default	1,919	18.8	.	.
		No Default	10,612	19.4	2.9	0.004
	Male 75%	Default	1,617	18.6	.	.
		No Default	8,374	19.7	4.8	0.000
	Gender Probabilities < 75% or Missing	Default	383	19.3	.	.
		No Default	1,805	19.4	0.2	0.813
Cash Flow Metric #5	Originated Loans	Default	3,919	12.2	.	.
		No Default	20,791	13.9	.	.
		All	24,710	13.6	12.8	0.000
	African American 75%	Default	465	11.7	.	.
		No Default	2,113	13.1	3.9	0.000
	Hispanic 75%	Default	875	12.3	.	.
		No Default	5,295	14.2	7.0	0.000
	Asian 75%	Default	86	11.9	.	.
		No Default	492	14.0	2.8	0.006
	Non-Hispanic White 75%	Default	938	12.6	.	.
		No Default	5,046	14.0	5.3	0.000
	Other or Missing BISG	Default	1,555	12.1	.	.
		No Default	7,845	13.7	7.9	0.000
	Female 75%	Default	1,919	12.3	.	.
		No Default	10,612	13.7	7.7	0.000
	Male 75%	Default	1,617	12.1	.	.
		No Default	8,374	14.1	9.9	0.000
	Gender Probabilities < 75% or Missing	Default	383	12.2	.	.
		No Default	1,805	13.5	3.3	0.001
Cash Flow Metric #6	Originated Loans	Default	3,919	2.6	.	.
		No Default	20,791	2.9	.	.
		All	24,710	2.8	12.2	0.000
	African American 75%	Default	465	2.4	.	.
		No Default	2,113	2.7	4.0	0.000
	Hispanic 75%	Default	875	2.6	.	.

		No Default	5,295	2.9	5.4	0.000
	Asian 75%	Default	86	2.5	.	.
		No Default	492	3.0	2.8	0.005
	Non-Hispanic White 75%	Default	938	2.6	.	.
		No Default	5,046	3.0	5.9	0.000
	Other or Missing BISG	Default	1,555	2.5	.	.
		No Default	7,845	2.9	7.7	0.000
	Female 75%	Default	1,919	2.6	.	.
		No Default	10,612	2.8	7.4	0.000
	Male 75%	Default	1,617	2.5	.	.
		No Default	8,374	2.9	9.5	0.000
	Gender Probabilities < 75% or Missing	Default	383	2.6	.	.
		No Default	1,805	2.8	2.9	0.004
Cash Flow Metric #7	Originated Loans	Default	3,919	18.7	.	.
		No Default	20,791	19.5	.	.
		All	24,710	19.4	5.3	0.000
	African American 75%	Default	465	18.4	.	.
		No Default	2,113	19.0	1.3	0.187
	Hispanic 75%	Default	875	18.7	.	.
		No Default	5,295	20.2	4.6	0.000
	Asian 75%	Default	86	18.6	.	.
		No Default	492	19.4	0.9	0.376
	Non-Hispanic White 75%	Default	938	19.1	.	.
		No Default	5,046	19.2	0.3	0.762
	Other or Missing BISG	Default	1,555	18.6	.	.
		No Default	7,845	19.4	3.4	0.001
	Female 75%	Default	1,919	18.8	.	.
		No Default	10,612	19.4	2.9	0.004
	Male 75%	Default	1,617	18.5	.	.
		No Default	8,374	19.7	4.9	0.000
	Gender Probabilities < 75% or Missing	Default	383	19.3	.	.
		No Default	1,805	19.3	0.2	0.870

Appendix D. Participant #4

Table 9. Logistic Model for Default Results Within Demographic Group[57]

Demographic Group	Count	TPD (Model 1) AUC	CFPD (Model 2) AUC	Combined (Model 3) AUC	All Variables (Model 4) AUC
African American 75%	2,594	0.568	0.584	0.620	0.670
Hispanic 75%	6,194	0.537	0.602	0.621	0.672
Asian 75%	579	0.568	0.583	0.619	0.764
Non-Hispanic White 75%	6,008	0.564	0.603	0.628	0.676
Other or Missing BISG Probability	9,441	0.565	0.581	0.615	0.652
Female 75%	12,591	0.567	0.584	0.618	0.650
Male 75%	10,025	0.552	0.606	0.630	0.660
Gender Probabilities < 75% or Missing	2,200	0.553	0.575	0.595	0.693
All Originations	24,816	0.559	0.592	0.620	0.650

57 The ROC analyses are restricted to the race/ethnicity or gender group listed and uses an indicator for "default" as the reference variable and the listed score as the rating. The analysis is based on originated loans with a known empirical default status.

Appendix D. Participant #4 Table 10. Model 1 Specification Within Race / Ethnicity Group								
	African American 75%		Hispanic 75%		Asian 75%		Non-Hispanic White 75%	
Control Variable	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	53.32	0.000	8.91	0.000	30.00	0.060	30.78	0.000
Constant	0.07	0.000	0.09	0.000	0.07	0.000	0.07	0.000
Pseudo R-Squared	0.009		0.002		0.006		0.007	
AUC	0.568		0.537		0.568		0.564	
Num. of Observations	2,594		6,194		579		6,008	

Appendix D. Participant #4 Table 11. Model 1 Specification Within Gender Group				
	Male 75%		Female 75%	
Control Variable	Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	16.98	0.000	38.12	0.000
Constant	0.09	0.000	0.07	0.000
Pseudo R-Squared	0.004		0.008	
AUC	0.552		0.567	
Num. of Observations	10,025		12,591	

Appendix D. Participant #4 Table 12. Model 2 Specification Within Race / Ethnicity Group								
	Non-Hispanic White 75%		African American 75%		Hispanic 75%		Asian 75%	
Control Variable	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
CFPD	271.56	0.000	1,180.95	0.000	362.50	0.016	879.10	0.000
Constant	0.04	0.000	0.02	0.000	0.03	0.000	0.03	0.000
Pseudo R-Squared	0.013		0.021		0.013		0.019	
AUC	0.584		0.602		0.583		0.603	
Num. of Observations	2,594		6,194		579		6,008	

Appendix D. Participant #4 Table 13. Model 2 Specification Within Gender Group				
	Male 75%		Female 75%	
Control Variable	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
CFPD	1,390.10	0.000	270.16	0.000
Constant	0.02	0.000	0.03	0.000
Pseudo R-Squared	0.021		0.013	
AUC	0.606		0.584	
Num. of Observations	10,025		12,591	

Appendix D. Participant #4 Table 14. Model 3 Specification Within Race / Ethnicity Group								
	Non-Hispanic White 75%		African American 75%		Hispanic 75%		Asian 75%	
Control Variable	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	235.17	0.000	53.84	0.000	176.52	0.009	121.29	0.000
CFPD	1,324.42	0.000	3,886.62	0.000	2,236.54	0.004	2,995.09	0.000
Constant	0.01	0.000	0.00	0.000	0.00	0.000	0.00	0.000
Pseudo R-Squared	0.029		0.028		0.026		0.031	
AUC	0.620		0.621		0.619		0.628	
Num. of Observations	2,594		6,194		579		6,008	

Appendix D. Participant #4 Table 15. Model 3 Specification Within Gender Group				
	Male 75%		Female 75%	
Control Variable	Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	96.87	0.000	153.08	0.000
CFPD	5,077.79	0.000	1,121.61	0.000
Constant	0.00	0.000	0.01	0.000
Pseudo R-Squared	0.032		0.026	
AUC	0.630		0.618	
Num. of Observations	10,025		12,591	

Appendix D. Participant #4

Table 16. Model 4 Specification Within Race / Ethnicity Group

Variable	Comparison Group	Non-Hispanic White 75%		African American 75%		Hispanic 75%		Asian 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	--	207.43	0.000	37.57	0.000	51.57	0.113	125.29	0.000
CFPD	--	272.08	0.000	3,186.86	0.000	118.40	0.160	2,227.73	0.000
Fraud Score	--	1.00	0.102	1.00	0.000	1.00	0.908	1.00	0.008
Missing Fraud Score	Not Missing Fraud Score	0.65	0.372	0.28	0.000	1.41	0.787	0.45	0.033
Bank Behavior Score	--	1.00	0.730	1.00	0.344	1.00	0.816	1.00	0.658
Missing Bank Behavior Score	Not Missing Bank Behavior Score	0.86	0.787	1.86	0.196	1.10	0.952	1.43	0.415
Self-Reported Income	--	1.00	0.667	1.00	0.053	1.00	0.966	1.00	0.536
Missing Self-Reported Income	Not Missing Self-Reported Income	1.23	0.467	1.16	0.476	3.23	0.124	1.58	0.015
Number of Accounts	--	0.99	0.843	0.91	0.010	1.03	0.827	0.90	0.002
Missing Number of Accounts	Not Missing Number of Account	.	.	.	.	.	.	.	.
Cash Flow Metric #1	--	1.01	0.049	1.00	0.179	1.02	0.036	1.01	0.034
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	0.00	0.048	0.00	0.179	0.00	0.035	0.00	0.034
Cash Flow Metric #3	--	1.00	0.609	1.00	0.000	1.00	0.245	1.00	0.092
Cash Flow Metric #4	--	0.93	0.394	1.00	0.941	0.77	0.207	1.12	0.062
Cash Flow Metric #5	--	1.00	0.819	0.99	0.379	1.00	0.998	0.98	0.147
Cash Flow Metric #6	--	0.96	0.462	0.99	0.784	0.92	0.522	0.98	0.634
Cash Flow Metric #7	--	1.06	0.467	0.98	0.759	1.26	0.271	0.90	0.091
Missing Cash Flow Metric #7	Not Missing Cash Flow Metric #7	.	.	.	.	.	.	.	.
Source Category #2	Source Category #1	0.78	0.746	0.59	0.470	.	.	1.25	0.692
Source Category #3		0.96	0.870	0.77	0.134	0.94	0.905	0.85	0.347
Source Category #4		1.01	0.983	0.96	0.905	2.32	0.221	1.05	0.826
Source Category #5		7.62	0.226	.	.	.	.	0.87	0.904

Source Category #7		1.11	0.627	1.04	0.778	1.33	0.520	1.07	0.638
Source Category #8		0.78	0.743	0.24	0.148	.	.	0.61	0.414
Source Category #9		1.11	0.647	1.15	0.388	0.69	0.440	1.08	0.608
Source Category #10		1.32	0.616	0.94	0.889	0.37	0.517	1.03	0.932
Source Category #11		1.13	0.829	1.79	0.187	8.16	0.112	1.49	0.323
Source Category #12		1.94	0.143	1.44	0.240	3.23	0.355	1.11	0.753
State #2	State #1	0.73	0.363	1.33	0.663	0.90	0.931	0.60	0.112
State #3		0.63	0.243	0.70	0.616	0.68	0.851	0.51	0.056
State #4		.	.	.	.	.	.	.	.
State #5		.	.	3.00	0.391	.	.	1.06	0.926
State #6		0.16	0.023	0.67	0.679	3.70	0.473	0.90	0.823
State #7		0.40	0.267	.	.	.	.	0.29	0.008
State #8		.	.	0.77	0.838	.	.	0.83	0.701
State #9		0.59	0.155	0.39	0.438	0.65	0.792	0.53	0.088
State #10		.	.	.	.	.	.	0.59	0.367
State #11		0.30	0.233	.	.	0.37	0.546	0.84	0.641
State #12		0.73	0.461	.	.	2.65	0.609	0.61	0.166
State #13		0.67	0.307	.	.	.	.	0.72	0.452
State #14		.	.	0.87	0.868	.	.	0.57	0.475
State #15		0.73	0.389	2.02	0.379	1.27	0.854	0.84	0.590
State #16		.	.	1.46	0.761	.	.	0.48	0.226
State #17		.	.	1.51	0.601	2.71	0.510	0.97	0.920
State #18		0.71	0.354	1.68	0.562	1.59	0.735	0.58	0.116
State #19		0.63	0.192	1.00	0.998	1.42	0.769	0.48	0.024
State #20		.	.	2.62	0.559	2.71	0.566	0.30	0.049
State #21		.	.	0.66	0.748	.	.	0.34	0.075
State #22		0.57	0.199	.	.	.	.	0.79	0.498
State #23		.	.	.	.	.	.	.	.
Application Date: Month #2	Application Date: Month #1	3.55	0.291	2.70	0.341	.	.	1.19	0.824
Application Date: Month #3		2.30	0.491	2.55	0.363	0.19	0.053	0.98	0.982

Application Date: Month #4		2.91	0.375	2.34	0.415	0.62	0.552	1.11	0.899
Application Date: Month #5		0.70	0.789	2.44	0.396	.	.	0.82	0.806
Application Date: Month #6		2.28	0.506	3.43	0.240	0.30	0.236	1.27	0.766
Application Date: Month #7		2.28	0.501	2.12	0.473	.	.	1.12	0.891
Application Date: Month #8		2.97	0.367	3.22	0.257	0.30	0.281	1.20	0.818
Application Date: Month #9		3.87	0.254	4.50	0.144	1.45	0.615	1.36	0.702
Application Date: Month #10		2.16	0.520	3.43	0.237	0.81	0.806	1.16	0.855
Application Date: Month #11		1.60	0.696	2.41	0.399	1.27	0.794	1.02	0.978
Application Date: Month #12		2.05	0.550	3.46	0.235	1.14	0.891	1.22	0.809
Application Date: Month #13		1.29	0.834	1.95	0.525	.	.	1.03	0.976
Application Date: Month #14		1.73	0.655	2.70	0.346	0.30	0.311	1.11	0.900
Application Date: Month #15		2.20	0.513	2.14	0.467	0.24	0.264	1.36	0.710
Application Date: Month #16		2.67	0.413	2.03	0.502	0.57	0.544	0.57	0.518
Application Date: Month #17		0.94	0.961	2.20	0.457	0.49	0.506	0.85	0.852
Application Date: Month #18		3.68	0.273	3.40	0.241	0.56	0.564	0.74	0.727
Application Date: Month #19		3.12	0.339	2.60	0.360	0.97	0.973	1.56	0.594
Application Date: Month #20		2.45	0.450	3.89	0.190	1.21	0.823	1.51	0.617
Application Date: Month #21		3.97	0.242	4.25	0.162	0.31	0.221	1.37	0.701
Application Date: Month #22		3.03	0.347	3.68	0.207	0.77	0.766	1.71	0.514
Application Date: Month #23		2.17	0.494	2.97	0.297	0.10	0.070	1.26	0.778
Application Date: Month #24		1.51	0.715	2.85	0.327	0.04	0.061	1.21	0.825
Application Date: Month #25		2.50	0.414	2.34	0.427	0.06	0.103	1.25	0.797
Application Date: Month #26		1.26	0.840	1.67	0.635	0.08	0.134	1.03	0.971
Application Date: Month #27		.	.	0.57	0.707	.	.	0.82	0.884
Constant	-	0.01	0.001	0.00	0.000	0.01	0.068	0.01	0.000
Pseudo R-Squared		0.058		0.056		0.152		0.059	
AUC		0.670		0.672		0.764		0.676	
Num. of Observations		2,571		6,128		514		5,978	

Appendix D. Participant #4

Table 17. Model 4 Specification Within Gender Group

Variable	Comparison Group	Male 75%		Female 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value
TPD	--	75.96	0.000	104.64	0.000
CFPD	--	2,611.70	0.000	689.01	0.000
Fraud Score	--	1.00	0.000	1.00	0.000
Missing Fraud Score	Not Missing Fraud Score	0.41	0.001	0.37	0.000
Bank Behavior Score	--	1.00	0.682	1.00	0.691
Missing Bank Behavior Score	Not Missing Bank Behavior Score	1.04	0.916	1.05	0.858
Self-Reported Income	--	1.00	0.867	1.00	0.080
Missing Self-Reported Income	Not Missing Self-Reported Income	1.28	0.096	1.30	0.054
Number of Accounts	--	0.95	0.064	0.91	0.000
Missing Number of Accounts	Not Missing Number of Account	.	.	.	.
Cash Flow Metric #1	--	1.00	0.800	1.00	0.125
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	0.01	0.795	0.00	0.126
Cash Flow Metric #3	--	1.00	0.035	1.00	0.000
Cash Flow Metric #4	--	1.02	0.608	0.98	0.611
Cash Flow Metric #5	--	0.98	0.060	0.98	0.038
Cash Flow Metric #6	--	0.94	0.021	0.97	0.187
Cash Flow Metric #7	--	0.97	0.545	1.02	0.586
Missing Cash Flow Metric #7	Not Missing Cash Flow Metric #7	.	.	.	.
Source Category #2	Source Category #1	0.94	0.894	0.91	0.818
Source Category #3		0.88	0.329	0.82	0.089
Source Category #4		1.26	0.219	0.97	0.849
Source Category #5		2.79	0.100	0.76	0.806
Source Category #7		1.15	0.211	1.04	0.722

Source Category #8		0.41	0.156	0.98	0.943
Source Category #9		1.18	0.170	1.06	0.610
Source Category #10		1.01	0.975	1.29	0.306
Source Category #11		1.04	0.899	1.27	0.441
Source Category #12		1.83	0.011	1.21	0.362
State #2	State #1	0.96	0.909	0.66	0.066
State #3		0.75	0.451	0.48	0.004
State #4		.	.	0.39	0.366
State #5		1.40	0.622	0.76	0.742
State #6		0.88	0.812	0.41	0.019
State #7		0.51	0.180	0.51	0.063
State #8		0.94	0.907	0.69	0.376
State #9		0.69	0.355	0.48	0.004
State #10		1.11	0.871	0.54	0.353
State #11		0.75	0.507	0.70	0.224
State #12		0.72	0.413	0.76	0.288
State #13		0.73	0.465	0.61	0.082
State #14		0.63	0.473	0.41	0.115
State #15		1.08	0.831	0.81	0.361
State #16		1.05	0.922	0.63	0.376
State #17		1.42	0.376	0.71	0.238
State #18		0.99	0.975	0.57	0.024
State #19		0.79	0.520	0.55	0.010
State #20		0.54	0.366	0.42	0.135
State #21		0.59	0.425	0.34	0.059
State #22		1.14	0.736	0.72	0.214
State #23		2.90	0.182	.	.
Application Date: Month #2	Application Date: Month #1	3.58	0.205	1.41	0.589
Application Date: Month #3		2.87	0.294	1.29	0.688
Application Date: Month #4		2.64	0.337	1.35	0.640

Application Date: Month #5		2.56	0.355	0.91	0.888
Application Date: Month #6		2.28	0.418	1.59	0.471
Application Date: Month #7		3.37	0.230	1.29	0.690
Application Date: Month #8		2.45	0.376	1.45	0.561
Application Date: Month #9		4.00	0.169	1.65	0.433
Application Date: Month #10		2.85	0.303	1.55	0.497
Application Date: Month #11		3.02	0.278	1.23	0.753
Application Date: Month #12		3.71	0.199	1.44	0.578
Application Date: Month #13		2.16	0.453	1.39	0.618
Application Date: Month #14		2.82	0.312	0.98	0.976
Application Date: Month #15		3.23	0.250	1.15	0.831
Application Date: Month #16		2.12	0.464	1.05	0.944
Application Date: Month #17		2.54	0.364	1.04	0.951
Application Date: Month #18		3.51	0.218	1.56	0.495
Application Date: Month #19		3.38	0.231	1.89	0.330
Application Date: Month #20		3.07	0.270	1.89	0.325
Application Date: Month #21		3.82	0.186	1.69	0.418
Application Date: Month #22		3.98	0.172	1.86	0.332
Application Date: Month #23		2.78	0.316	1.66	0.433
Application Date: Month #24		2.67	0.343	1.34	0.664
Application Date: Month #25		2.33	0.414	1.39	0.627
Application Date: Month #26		1.59	0.655	0.93	0.917
Application Date: Month #27		.	.	0.87	0.866
Constant	-	0.01	0.000	0.02	0.000
Pseudo R-Squared		0.051		0.044	
AUC		0.660		0.650	
Num. of Observations		9,959		12,529	

APPENDIX E: Participant 5

Appendix E. Participant #5

Appendix E. Participant #5
Table 1. Data Diagnostics: All Applications

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Annual Income	Approved	9,790	0	0	$500	$11,000	$28,000	$46,000	$75,000	$156,000	$1,000,000,000	$164,046
	Declined	220,162	0	454	$0	$10,000	$27,500	$40,000	$60,000	$115,000	$1,308,888,832	$82,140
	All	229,952	0	454	$0	$10,000	$27,500	$40,000	$61,000	$118,000	$1,308,888,832	$85,627
Pre-Qualification DTI	Approved	9,790	0	1	0.00	0.02	0.13	0.23	0.34	0.50	0.60	0.24
	Declined	220,162	1,450	40	0.00	0.02	0.14	0.24	0.37	0.72	4,944.00	0.62
	All	229,952	1,450	41	0.00	0.02	0.14	0.24	0.37	0.70	4,944.00	0.60
Pre-Qualification Cash Flow Score	Approved	9,790	393	0	0.00	0.03	0.06	0.10	0.13	0.18	0.77	0.10
	Declined	220,162	70,277	0	0.00	0.06	0.13	0.20	0.30	0.46	0.98	0.23
	All	229,952	70,670	0	0.00	0.06	0.12	0.19	0.29	0.45	0.98	0.22
Pre-Qualification Vantage Score	Approved	9,790	345	0	600	606	627	655	698	755	834	666
	Declined	220,162	14,592	0	300	449	517	560	608	670	837	561
	All	229,952	14,937	0	300	450	518	564	614	680	837	566
Total Tradelines at Application	Approved	9,790	1,226	0	2	2	5	10	21	49	269	16
	Declined	220,162	22,320	0	2	3	7	14	24	45	282	18
	All	229,952	23,546	0	2	3	7	14	24	45	282	18
Total Inquiries at Application	Approved	9,790	1,226	366	0	1	3	7	13	32	297	11
	Declined	220,162	22,320	4,210	0	1	5	10	17	37	760	13
	All	229,952	23,546	4,576	0	1	5	10	17	37	760	13
Application Vantage Score	Approved	9,790	342	0	600	606	627	655	698	755	834	666
	Declined	220,162	214,463	0	524	608	636	665	711	779	834	676
	All	229,952	214,805	0	524	606	630	659	703	763	834	670
APR Given	Approved	9,790	0	0	9.74	17.74	20.74	20.74	20.74	20.74	20.74	20.25
	Declined	220,162	220,162	0	.	.	.	.	.	.	.	.
	All	229,952	220,162	0	9.74	17.74	20.74	20.74	20.74	20.74	20.74	20.25
Cash Flow Metric #1	Approved	9,790	0	0	$500	$500	$1,000	$1,500	$3,000	$6,000	$10,000	$2,183
	Declined	220,162	220,162	0	.	.	.	.	.	.	.	.
	All	229,952	220,162	0	$500	$500	$1,000	$1,500	$3,000	$6,000	$10,000	$2,183
Current Balance	Approved	9,790	45	3,434	-$944	$0	$0	$153	$728	$2,157	$10,335	$529
	Declined	220,162	220,162	0	.	.	.	.	.	.	.	.

	All	229,952	220,207	3,434	-$944	$0	$0	$153	$728	$2,157	$10,335	$529
Cash Flow Metric #2	Approved	9,790	45	3,434	-94.4%	0.0%	0.0%	10.7%	58.2%	99.2%	341.0%	30.1%
	Declined	220,162	220,162	0	.	.	.	.	.	.	.	.
	All	229,952	220,207	3,434	-94.4%	0.0%	0.0%	10.7%	58.2%	99.2%	341.0%	30.1%
Number of Days Past Due	Approved	9,790	45	9,565	0	0	0	0	0	0	133	1
	Declined	220,162	220,162	0	.	.	.	.	.	.	.	.
	All	229,952	220,207	9,565	0	0	0	0	0	0	133	1

Appendix E. Participant #5 Table 2. Difference of Means Tests: All Applications[58]					
Variable	**Sample**	**#**	**Mean**	**T-Stat**	**P-Value**
Annual Income	Approved	9,790	$164,046	.	.
	Declined	220,162	$82,140	-0.80	0.425
Pre-Qualification DTI	Approved	9,790	0.24	.	.
	Declined	218,712	0.62	7.34	0.000
Pre-Qualification Cash Flow Score	Approved	9,397	0.10	.	.
	Declined	149,885	0.23	191.57	0.000
Pre-Qualification Vantage Score	Approved	9,445	666	.	.
	Declined	205,570	561	-203.66	0.000
Total Tradelines at Application	Approved	8,564	16	.	.
	Declined	197,842	18	8.65	0.000
Total Inquiries at Application	Approved	8,564	11	.	.
	Declined	197,842	13	18.47	0.000
Application Vantage Score	Approved	9,448	666	.	.
	Declined	5,699	676	12.72	0.000
APR Given	Approved	9,790	20.25	.	.
	Declined	0	.	.	.
Cash Flow Metric #1	Approved	9,790	$2,183	.	.
	Declined	0	.	.	.
Current Balance	Approved	9,745	$529	.	.
	Declined	0	.	.	.
Cash Flow Metric #2	Approved	9,745	30.1%	.	.
	Declined	0	.	.	.

58 The significance test tests the difference in means between approved applicants and declined applicants, using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% confidence level. Counts in this table are of non-missing values of the indicated variable.

Appendix E. Participant #5
Table 3. Data Diagnostics: Originations

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Annual Income	Not Past Due	8,571	0	0	$500	$12,000	$28,800	$47,000	$75,200	$155,000	$1,000,000,000	$178,252
	Past Due	180	0	0	$850	$12,500	$25,000	$44,883	$73,500	$150,000	$850,000	$61,654
	All	8,751	0	0	$500	$12,000	$28,800	$47,000	$75,000	$155,000	$1,000,000,000	$175,853
Pre-Qualification DTI	Not Past Due	8,571	0	1	0.00	0.02	0.13	0.23	0.34	0.50	0.60	0.24
	Past Due	180	0	0	0.00	0.02	0.14	0.24	0.37	0.52	0.60	0.26
	All	8,751	0	1	0.00	0.02	0.13	0.23	0.34	0.50	0.60	0.24
Pre-Qualification Cash Flow Score	Not Past Due	8,571	313	0	0.00	0.03	0.06	0.10	0.13	0.17	0.77	0.10
	Past Due	180	22	0	0.01	0.04	0.07	0.10	0.13	0.23	0.69	0.11
	All	8,751	335	0	0.00	0.03	0.06	0.10	0.13	0.17	0.77	0.10
Pre-Qualification Vantage Score	Not Past Due	8,571	298	0	600	606	627	654	696	754	834	665
	Past Due	180	6	0	600	602	614	640	683	754	801	656
	All	8,751	304	0	600	605	627	654	696	754	834	665
Total Tradelines at Application	Not Past Due	8,571	1,067	0	2	2	5	10	21	49	269	16
	Past Due	180	23	0	2	2	5	11	22	44	71	15
	All	8,751	1,090	0	2	2	5	10	21	49	269	16
Total Inquiries at Application	Not Past Due	8,571	1,067	311	0	1	3	7	14	32	297	11
	Past Due	180	23	5	0	1	5	10	17	41	162	14
	All	8,751	1,090	316	0	1	4	7	14	32	297	11
Application Vantage Score	Not Past Due	8,571	295	0	600	606	627	654	696	754	834	665
	Past Due	180	6	0	600	602	614	640	683	754	801	656
	All	8,751	301	0	600	605	627	654	696	754	834	665
APR Given	Not Past Due	8,571	0	0	9.74	17.74	20.74	20.74	20.74	20.74	20.74	20.28
	Past Due	180	0	0	9.74	17.74	20.74	20.74	20.74	20.74	20.74	20.10
	All	8,751	0	0	9.74	17.74	20.74	20.74	20.74	20.74	20.74	20.28
Cash Flow Metric #1	Not Past Due	8,571	0	0	$500	$500	$1,000	$1,500	$2,500	$6,000	$10,000	$2,174
	Past Due	180	0	0	$500	$500	$750	$1,500	$2,500	$6,000	$10,000	$2,063
	All	8,751	0	0	$500	$500	$1,000	$1,500	$2,500	$6,000	$10,000	$2,172
Current Balance	Not Past Due	8,571	0	2,464	-$944	$0	$0	$222	$754	$2,157	$9,960	$559
	Past Due	180	0	1	-$3	$217	$748	$1,118	$2,523	$5,782	$10,335	$1,913
	All	8,751	0	2,465	-$944	$0	$0	$235	$785	$2,295	$10,335	$586
	Not Past Due	8,571	0	2,464	-94.4%	0.0%	0.0%	15.6%	62.0%	98.7%	188.8%	32.2%

Cash Flow Metric #2	Past Due	180	0	1	-0.2%	14.4%	97.4%	100.3%	102.1%	108.4%	341.0%	92.7%
	All	8,751	0	2,465	-94.4%	0.0%	0.0%	16.6%	65.7%	99.4%	341.0%	33.4%
Number of Days Past Due	Not Past Due	8,571	0	8,571	0	0	0	0	0	0	0	0
	Past Due	180	0	0	13	13	13	13	41	103	133	29
	All	8,751	0	8,571	0	0	0	0	0	0	133	1

Appendix E. Participant #5

Table 4. Difference of Means Tests: Originations[59]

Variable	Sample	#	Mean	T-Stat	P-Value
Annual Income	Not Past Due	8,571	$178,252	.	.
	Past Due	180	$61,654	1.00	0.318
Pre-Qualification DTI	Not Past Due	8,571	0.24	.	.
	Past Due	180	0.26	-1.38	0.169
Pre-Qualification Cash Flow Score	Not Past Due	8,258	0.10	.	.
	Past Due	158	0.11	-1.46	0.146
Pre-Qualification Vantage Score	Not Past Due	8,273	665	.	.
	Past Due	174	656	2.26	0.025
Total Tradelines at Application	Not Past Due	7,504	16	.	.
	Past Due	157	15	0.92	0.358
Total Inquiries at Application	Not Past Due	7,504	11	.	.
	Past Due	157	14	-2.38	0.018
Application Vantage Score	Not Past Due	8,276	665	.	.
	Past Due	174	656	2.28	0.024
APR Given	Not Past Due	8,571	20.28	.	.

59 The significance test tests the difference in means between applicants with a past due status (i.e. positive number of days past due) compared to applicants with a non-past due status (i.e. zero days past due), using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% confidence level. Counts in this table are of non-missing values of the indicated variable.

	Past Due	180	20.10	1.19	0.237
Cash Flow Metric #1	Not Past Due	8,571	$2,174	.	.
	Past Due	180	$2,063	0.79	0.429
Current Balance	Not Past Due	8,571	$559	.	.
	Past Due	180	$1,913	-9.55	0.000
Cash Flow Metric #2	Not Past Due	8,571	32.2%	.	.
	Past Due	180	92.7%	-24.67	0.000

Appendix E. Participant #5
Table 5. Logistic Models for Past Due Status Results[60]

Model	AUC
(1) Pre-Qualification Vantage Score	0.573
(2) Pre-Qualification Cash Flow Score	0.572
(3) Pre-Qualification Vantage Score and Cash Flow Score	0.659

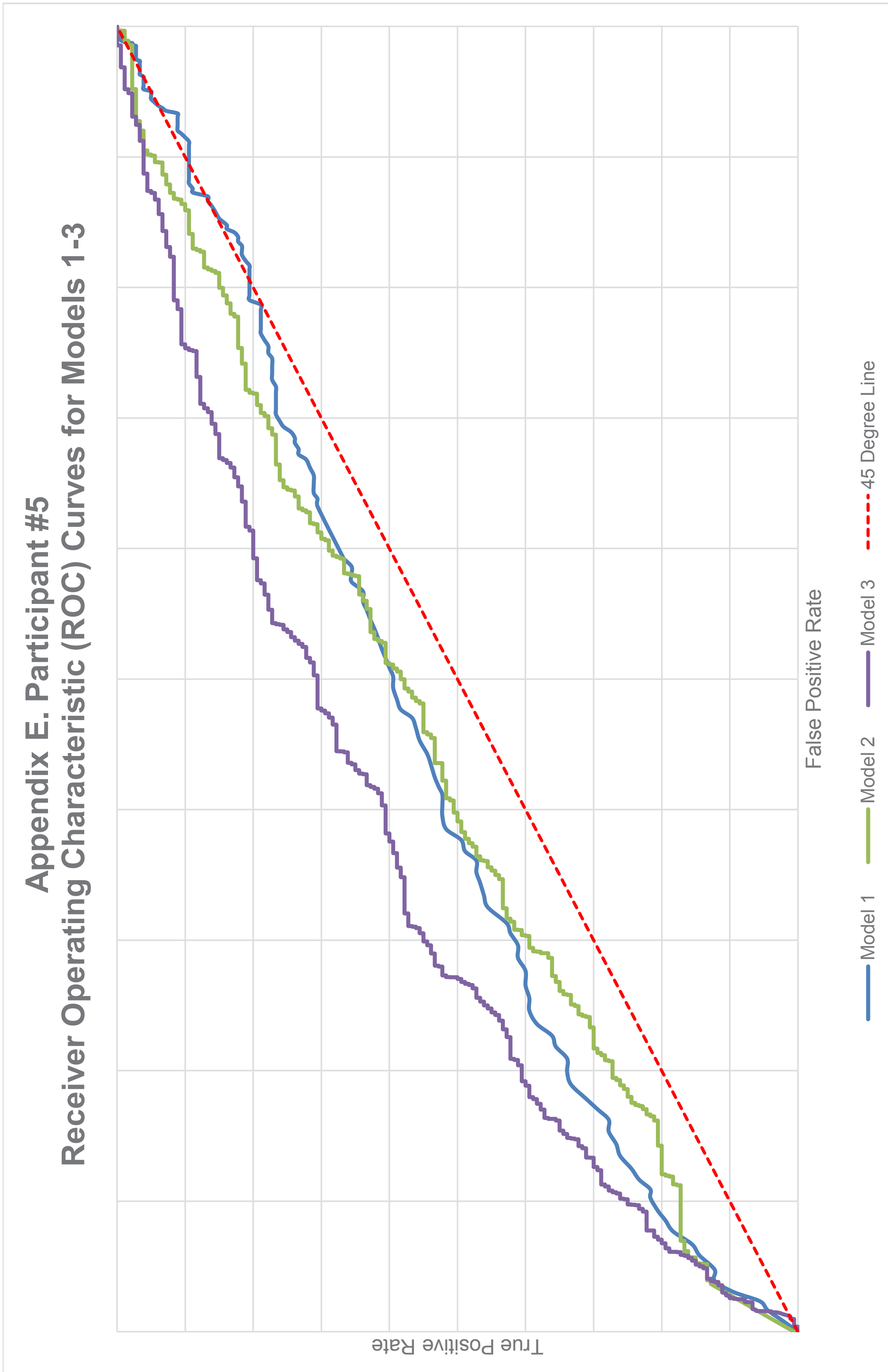
60 The dependent variable is a 0/1 indicator for past due, with values of 1 indicating past due status and 0 indicating non-past due status.

Appendix E. Participant #5

Table 6. Logistic Model for Past Due Status Specifications[61]

Control Variable	Pre-Qual. VS		Pre-Qual. CF		Pre-Qual. VS and CF	
	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Pre-Qualification Cash Flow Score (CF)	.	.	9.62	0.07	66.22	0.01
Missing Flag, Pre-Qualification Cash Flow Score (CF)	.	.	4.65	0.00	13.08	0.00
Pre-Qualification Vantage Score (VS)	1.00	0.04	.	.	0.99	0.00
Missing Flag, Pre-Qualification Vantage Score (VS)	0.06	0.04	.	.	0.00	0.00
Constant	0.35	0.44	0.02	0.00	12.40	0.09
Pseudo R-Squared	0.004		0.015		0.031	
AUC	0.573		0.572		0.659	
Sample Size	8,751		8,751		8,751	

61 The dependent variable is a 0/1 indicator for past due, with values of 1 indicating past due status and 0 indicating non-past due status.



Page 95 of 161

Appendix E. Participant #5

Table 7. Past Due Frequency by Cash Flow and Vantage Score Percentile, 10 Deciles[62]

Vantage Score	Cash Flow Score 0 - 10th	10 - 20th	20 - 30th	30 - 40th	40 - 50th	50 - 60th	60 - 70th	70 - 80th	80 - 90th	90 - 100th
0 - 10th	3.5%	6.6%	5.0%	2.6%	8.2%	1.1%	5.0%	3.8%	5.1%	0.0%
10 - 20th	0.0%	1.3%	2.5%	4.6%	0.0%	3.2%	2.1%	3.7%	3.4%	10.7%
20 - 30th	1.6%	1.4%	2.8%	0.0%	1.3%	2.2%	0.0%	2.2%	3.0%	0.0%
30 - 40th	1.3%	1.2%	2.2%	0.9%	3.0%	0.9%	1.3%	4.0%	3.2%	0.0%
40 - 50th	2.1%	1.4%	1.1%	5.8%	0.0%	1.2%	2.2%	0.0%	1.0%	3.6%
50 - 60th	0.0%	2.2%	1.0%	0.0%	2.1%	1.0%	4.4%	4.0%	0.0%	0.0%
60 - 70th	1.4%	1.2%	0.0%	3.2%	0.0%	1.5%	1.7%	3.9%	0.0%	3.6%
70 - 80th	2.4%	0.0%	0.0%	1.1%	1.5%	2.6%	0.0%	1.7%	0.0%	0.8%
80 - 90th	2.2%	0.0%	0.0%	1.8%	1.8%	1.8%	0.0%	0.0%	0.0%	1.3%
90 - 100th	0.0%	1.2%	0.0%	0.0%	2.2%	0.0%	0.0%	0.0%	3.0%	2.2%

Appendix E. Participant #5

Table 8. Summary of Actions Taken[63]

	All Applications	Approved Applications		Denied Applications		Originated Loans		Past Due Loans	
	Count	Count	Percent	Count	Percent	Count	Percent	Count	Percent[1]
All	229,952	9,790	4.26%	220,162	95.74%	8,751	3.81%	180	2.06%

62 Cells are shaded based on values. Green indicates values close to the lowest default frequency, yellow indicates values close to the median default frequency, and red indicates values close to the highest default frequency. Cells with fewer than 5 loans are excluded from this heat map. Percentiles are based on the population of originated loans. 304 originated loans with a missing Pre-Qual. Vantage score and 335 originated loans with a missing Cash Flow Score were excluded from the frequency table.

63 The percentages in this column are calculated out of originated loans.

Appendix E. Participant #5 Table 9. Difference of Means Tests Within Demographic Group: Originated Loans[64]						
Variable	**Demographic Group**	**Sample**	**#**	**Mean**	**T-Stat**	**P-Value**
Annual Income	All Originations	Not Past Due	8,571	$178,251.67	.	.
		Past Due	180	$61,653.57	.	.
		Originated	8,751	$175,853.34	0.998	0.318
	African American 75%	Not Past Due	345	$62,333.98	.	.
		Past Due	15	$92,280.34	-0.550	0.591
	Hispanic 75%	Not Past Due	997	$51,459.38	.	.
		Past Due	19	$72,231.63	-1.623	0.122
	Asian 75%	Not Past Due	561	$70,528.97	.	.
		Past Due	6	$67,983.34	0.084	0.936
	Non-Hispanic White 75%	Not Past Due	4,118	$304,230.16	.	.
		Past Due	87	$55,716.50	1.023	0.306
	Other or Missing BISG	Not Past Due	2,550	$63,764.03	.	.
		Past Due	53	$58,222.64	0.979	0.332
Pre-Qualification DTI	All Originations	Not Past Due	8,571	0.24	.	.
		Past Due	180	0.26	.	.
		Originated	8,751	0.24	-1.381	0.169
	African American 75%	Not Past Due	345	0.24	.	.
		Past Due	15	0.19	1.135	0.274
	Hispanic 75%	Not Past Due	997	0.24	.	.
		Past Due	19	0.33	-2.500	0.022
	Asian 75%	Not Past Due	561	0.22	.	.
		Past Due	6	0.18	1.252	0.262
	Non-Hispanic White 75%	Not Past Due	4,118	0.24	.	.
		Past Due	87	0.26	-1.267	0.209
	Other or Missing BISG	Not Past Due	2,550	0.24	.	.
		Past Due	53	0.25	-0.322	0.749
Pre-Qualification Cash Flow Score	All Originations	Not Past Due	8,258	0.10	.	.
		Past Due	158	0.11	.	.
		Originated	8,416	0.10	-1.462	0.146
	African American 75%	Not Past Due	321	0.11	.	.
		Past Due	10	0.09	1.240	0.243
	Hispanic 75%	Not Past Due	973	0.11	.	.
		Past Due	17	0.10	0.182	0.858
	Asian 75%	Not Past Due	536	0.09	.	.
		Past Due	5	0.10	-0.567	0.599
		Not Past Due	3,966	0.10	.	.

64 T-tests assume unequal variances and are conducted on the past due and not past due populations. Yellow highlighting indicates a difference between the past due and not past due groups that is statistically significant at the 95% confidence level (P-value < 0.05). Highlighting is shown regardless of the direction of the difference. Counts displayed are the counts of non-missing values for each variable, by demographic group and status.

	Non-Hispanic White 75%	Past Due	82	0.11	-1.503	0.137
	Other or Missing BISG	Not Past Due	2,462	0.10	.	.
		Past Due	44	0.11	-0.720	0.475
Pre-Qualification Vantage Score	All Originations	Not Past Due	8,273	665	.	.
		Past Due	174	656	.	.
		Originated	8,447	665	2.262	0.025
	African American 75%	Not Past Due	335	661	.	.
		Past Due	13	681	-1.489	0.160
	Hispanic 75%	Not Past Due	967	660	.	.
		Past Due	18	641	1.776	0.093
	Asian 75%	Not Past Due	542	681	.	.
		Past Due	6	665	0.660	0.538
	Non-Hispanic White 75%	Not Past Due	3,968	664	.	.
		Past Due	84	646	3.457	0.001
	Other or Missing BISG	Not Past Due	2,461	665	.	.
		Past Due	53	670	-0.588	0.559
Total Tradelines at Application	All Originations	Not Past Due	7,504	16	.	.
		Past Due	157	15	.	.
		Originated	7,661	16	0.921	0.358
	African American 75%	Not Past Due	309	17	.	.
		Past Due	13	11	1.396	0.186
	Hispanic 75%	Not Past Due	849	15	.	.
		Past Due	15	28	-2.374	0.032
	Asian 75%	Not Past Due	493	13	.	.
		Past Due	6	9	2.033	0.084
	Non-Hispanic White 75%	Not Past Due	3,634	17	.	.
		Past Due	78	15	1.202	0.233
	Other or Missing BISG	Not Past Due	2,219	16	.	.
		Past Due	45	13	1.598	0.117
Total Inquiries at Application	All Originations	Not Past Due	7,504	11	.	.
		Past Due	157	14	.	.
		Originated	7,661	11	-2.382	0.018
	African American 75%	Not Past Due	309	11	.	.
		Past Due	13	9	1.278	0.218
	Hispanic 75%	Not Past Due	849	12	.	.
		Past Due	15	14	-1.131	0.276
	Asian 75%	Not Past Due	493	10	.	.
		Past Due	6	10	-0.078	0.941
	Non-Hispanic White 75%	Not Past Due	3,634	10	.	.
		Past Due	78	15	-2.336	0.022
	Other or Missing BISG	Not Past Due	2,219	11	.	.
		Past Due	45	16	-1.177	0.245
Application Vantage Score	All Originations	Not Past Due	8,276	665	.	.
		Past Due	174	656	.	.
		Originated	8,450	665	2.279	0.024
	African American 75%	Not Past Due	335	661	.	.
		Past Due	13	681	-1.490	0.160

	Hispanic 75%	Not Past Due	967	660	.	.
		Past Due	18	641	1.771	0.094
	Asian 75%	Not Past Due	543	681	.	.
		Past Due	6	665	0.659	0.539
	Non-Hispanic White 75%	Not Past Due	3,969	664	.	.
		Past Due	84	646	3.461	0.001
	Other or Missing BISG	Not Past Due	2,462	665	.	.
		Past Due	53	670	-0.567	0.573
APR Given	All Originations	Not Past Due	8,571	20.28	.	.
		Past Due	180	20.10	.	.
		Originated	8,751	20.28	1.186	0.237
	African American 75%	Not Past Due	345	20.35	.	.
		Past Due	15	19.54	1.995	0.064
	Hispanic 75%	Not Past Due	997	20.53	.	.
		Past Due	19	20.42	0.493	0.628
	Asian 75%	Not Past Due	561	19.60	.	.
		Past Due	6	20.24	-1.240	0.265
	Non-Hispanic White 75%	Not Past Due	4,118	20.29	.	.
		Past Due	87	20.29	-0.008	0.994
	Other or Missing BISG	Not Past Due	2,550	20.30	.	.
		Past Due	53	19.82	1.522	0.134
Cash Flow Metric #1	All Originations	Not Past Due	8,571	$2,174.13	.	.
		Past Due	180	$2,062.50	.	.
		Originated	8,751	$2,171.84	0.792	0.429
	African American 75%	Not Past Due	345	$2,198.55	.	.
		Past Due	15	$2,066.67	0.317	0.756
	Hispanic 75%	Not Past Due	997	$1,802.91	.	.
		Past Due	19	$2,092.11	-0.887	0.386
	Asian 75%	Not Past Due	561	$2,837.34	.	.
		Past Due	6	$2,750.00	0.059	0.956
	Non-Hispanic White 75%	Not Past Due	4,118	$2,117.53	.	.
		Past Due	87	$1,718.39	2.645	0.010
	Other or Missing BISG	Not Past Due	2,550	$2,261.47	.	.
		Past Due	53	$2,537.74	-0.830	0.410
Current Balance	All Originations	Not Past Due	8,571	$558.57	.	.
		Past Due	180	$1,913.09	.	.
		Originated	8,751	$586.43	-9.549	0.000
	African American 75%	Not Past Due	345	$850.30	.	.
		Past Due	15	$2,039.62	-3.005	0.009
	Hispanic 75%	Not Past Due	997	$527.83	.	.
		Past Due	19	$1,631.54	-4.136	0.001
	Asian 75%	Not Past Due	561	$414.29	.	.
		Past Due	6	$2,466.52	-1.327	0.242
	Non-Hispanic White 75%	Not Past Due	4,118	$552.00	.	.
		Past Due	87	$1,607.74	-7.150	0.000
	Other or Missing BISG	Not Past Due	2,550	$573.46	.	.
		Past Due	53	$2,416.79	-5.315	0.000
	All Originations	Not Past Due	8,571	32.2%	.	.

Cash Flow Metric #2		Past Due	180	92.7%	.	.
		Originated	8,751	33.4%	-24.669	0.000
	African American 75%	Not Past Due	345	42.1%	.	.
		Past Due	15	98.8%	-20.628	0.000
	Hispanic 75%	Not Past Due	997	33.5%	.	.
		Past Due	19	85.3%	-7.328	0.000
	Asian 75%	Not Past Due	561	21.7%	.	.
		Past Due	6	79.7%	-3.850	0.012
	Non-Hispanic White 75%	Not Past Due	4,118	32.8%	.	.
		Past Due	87	95.1%	-15.322	0.000
	Other or Missing BISG	Not Past Due	2,550	31.7%	.	.
		Past Due	53	91.2%	-15.290	0.000

Appendix E. Participant #5

Table 10. Logistic Model for Past Due Results Within Demographic Group[65]

Demographic Group	Count	Model 1 AUC	Model 2 AUC	Model 3 AUC
All Originations	8,751	0.573	0.572	0.659
African American 75%	360	0.667	0.672	0.689
Hispanic 75%	1,016	0.663	0.557	0.731
Asian 75%	567	0.587	0.649	0.693
Non-Hispanic White 75%	4,205	0.632	0.555	0.665
Other or Missing BISG	2,603	0.508	0.595	0.616

65 The ROC analyses are restricted to the Race/Ethnicity or gender group listed and uses an indicator for "past due" as the reference variable and the listed score as the rating. The estimation samples may differ slightly from the displayed count based on missing values and perfect prediction among the set of predictor variables.

Appendix E. Participant #5

Table 11. Model 1 Specification Within Race/Ethnicity Group[66]

Control Variable	African American 75%		Hispanic 75%		Asian 75%		Non-Hispanic White 75%	
	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Pre-Qualification Cash Flow Score (CF)	.	.	.	.	.	.	.	.
Missing Flag, Pre-Qualification Cash Flow Score (CF)	.	.	.	.	.	.	.	.
Pre-Qualification Vantage Score (VS)	1.01	0.09	0.99	0.16	0.99	0.51	0.99	0.01
Missing Flag, Pre-Qualification Vantage Score (VS)	1376.36	0.04	0.00	0.19	.	.	0.00	0.01
Constant	0.00	0.01	91.79	0.45	0.63	0.94	12.39	0.26
Pseudo R-Squared	0.041		0.023		0.008		0.016	
AUC	0.667		0.663		0.587		0.632	
Sample Size	360		1,016		548		4,205	

[66] The dependent variable is a 0/1 indicator for past due, with values of 1 indicating past due status and 0 indicating non-past due status.

Appendix E. Participant #5

Table 12. Model 2 Specification Within Race/Ethnicity Group[67]

Control Variable	African American 75%		Hispanic 75%		Asian 75%		Non-Hispanic White 75%	
	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Pre-Qualification Cash Flow Score (CF)	0.00	0.29	0.39	0.86	9.01	0.47	25.88	0.03
Missing Flag, Pre-Qualification Cash Flow Score (CF)	3.24	0.18	4.32	0.12	5.27	0.15	2.24	0.10
Pre-Qualification Vantage Score (VS)	.	.	.	.	.	.	.	.
Missing Flag, Pre-Qualification Vantage Score (VS)	.	.	.	.	.	.	.	.
Constant	0.06	0.00	0.02	0.00	0.01	0.00	0.01	0.00
Pseudo R-Squared	0.076		0.015		0.020		0.006	
AUC	0.672		0.557		0.649		0.555	
Sample Size	360		1,016		567		4,205	

67 The dependent variable is a 0/1 indicator for past due, with values of 1 indicating past due status and 0 indicating non-past due status.

Appendix E. Participant #5 Table 13. Model 3 Specification Within Race/Ethnicity Group[68]								
	African American 75%		**Hispanic 75%**		**Asian 75%**		**Non-Hispanic White 75%**	
Control Variable	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
Pre-Qualification Cash Flow Score (CF)	0.03	0.62	39.44	0.64	18.78	0.50	374.02	0.00
Missing Flag, Pre-Qualification Cash Flow Score (CF)	6.16	0.12	54.65	0.00	12.69	0.05	9.46	0.00
Pre-Qualification Vantage Score (VS)	1.00	0.86	0.98	0.02	0.99	0.25	0.99	0.00
Missing Flag, Pre-Qualification Vantage Score (VS)	3.43	0.74	0.00	0.04	.	.	0.00	0.00
Constant	0.07	0.43	69983.84	0.12	16.94	0.67	145.65	0.03
Pseudo R-Squared	0.102		0.067		0.042		0.033	
AUC	0.689		0.731		0.693		0.665	
Sample Size	360		1,016		548		4,205	

68 The dependent variable is a 0/1 indicator for past due, with values of 1 indicating past due status and 0 indicating non-past due status.

APPENDIX F: Participant 6

Appendix F. Participant #6

Appendix F. Participant #6

Table 1. Data Diagnostics: All Applications

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Date Difference #1	Approved	3,994	0	46	1	3	11	21	36	85	418	30
	Declined	1,566	0	63	1	1	4	11	28	87	1,039	25
	In Progress	586	586	0	.	.	.	.	.	.	.	.
	Withdrawn	7,285	0	221	-314	2	13	28	63	609	1,405	91
	Approved/Declined	5,560	0	109	1	2	8	19	34	86	1,039	29
	All	13,431	586	330	-314	2	10	23	48	377	1,405	64
FICO score	Approved	3,994	307	0	431	522	598	643	687	758	847	642
	Declined	1,566	377	0	423	474	521	572	640	729	822	584
	In Progress	586	582	0	543	543	584	655	696	706	706	640
	Withdrawn	7,285	2,814	0	402	489	561	624	679	755	850	622
	Approved/Declined	5,560	684	0	423	498	575	632	680	752	847	628
	All	13,431	4,080	0	402	494	568	629	680	754	850	625
BK score	Approved	3,994	231	1	1	118	186	370	607	872	993	412
	Declined	1,566	342	0	2	37	137	175	393	740	993	278
	In Progress	586	582	0	158	158	271	468	658	762	762	464
	Withdrawn	7,285	2,699	0	2	71	154	310	584	834	993	378
	Approved/Declined	5,560	573	1	1	88	156	316	571	840	993	379
	All	13,431	3,854	1	1	79	155	313	579	840	993	379
# of open accounts on credit report	Approved	3,994	3,565	225	1	1	1	1	2	4	10	2
	Declined	1,566	1,463	60	1	1	1	2	3	4	7	2
	In Progress	586	582	3	3	3	3	3	3	3	3	3
	Withdrawn	7,285	6,634	377	1	1	1	1	3	5	11	2
	Approved/Declined	5,560	5,028	285	1	1	1	1	2	4	10	2
	All	13,431	12,244	665	1	1	1	1	2	4	11	2
	Approved	3,994	164	3,432	$39	$1,122	$13,949	$41,650	$136,028	$353,535	$931,802	$95,478
	Declined	1,566	324	1,149	$1,014	$5,327	$22,633	$73,085	$169,191	$380,959	$520,195	$113,981

$ amount of unpaid balances on credit report	In Progress	586	582	0	$238	$238	$1,029	$73,295	$245,198	$345,626	$345,626	$123,113
	Withdrawn	7,285	2,599	4,100	$64	$939	$15,478	$52,017	$156,461	$392,005	$1,004,322	$109,146
	Approved/Declined	5,560	488	4,581	$39	$1,385	$14,788	$46,737	$144,619	$353,535	$931,802	$98,982
	All	13,431	3,669	8,681	$39	$1,149	$15,221	$48,487	$151,776	$377,032	$1,004,322	$104,581
$ amount of monthly payments on credit report	Approved	3,994	164	3,440	$3	$57	$344	$716	$1,658	$3,802	$34,580	$1,311
	Declined	1,566	324	1,151	$57	$96	$440	$1,045	$1,925	$3,471	$5,308	$1,349
	In Progress	586	582	0	$25	$25	$88	$1,376	$3,130	$3,659	$3,659	$1,609
	Withdrawn	7,285	2,599	4,123	$25	$53	$418	$915	$1,997	$4,208	$12,034	$1,395
	Approved/Declined	5,560	488	4,591	$3	$77	$354	$798	$1,734	$3,641	$34,580	$1,318
	All	13,431	3,669	8,714	$3	$56	$380	$856	$1,874	$4,052	$34,580	$1,361
$ Credit limit of revolving accounts on credit report	Approved	3,994	3,641	15	$9	$382	$3,863	$15,831	$41,026	$426,300	$3,294,300	$93,686
	Declined	1,566	1,490	3	$72	$365	$7,271	$26,741	$67,393	$275,100	$586,157	$58,194
	In Progress	586	582	0	$240	$240	$13,113	$52,684	$123,670	$167,958	$167,958	$68,392
	Withdrawn	7,285	6,796	28	$1	$212	$3,057	$15,447	$54,879	$307,217	$10,297,775	$94,274
	Approved/Declined	5,560	5,131	18	$9	$382	$4,370	$17,089	$42,330	$332,429	$3,294,300	$87,382
	All	13,431	12,509	46	$1	$254	$3,596	$16,222	$49,453	$321,925	$10,297,775	$90,922
$ unpaid balances of revolving accounts on credit report	Approved	3,994	3,565	81	$9	$241	$1,450	$4,697	$11,650	$41,707	$154,807	$11,096
	Declined	1,566	1,463	26	$69	$250	$1,512	$6,768	$15,858	$46,540	$68,775	$11,780
	In Progress	586	582	0	$238	$238	$1,029	$7,657	$26,903	$40,310	$40,310	$13,966
	Withdrawn	7,285	6,634	175	$1	$178	$1,109	$5,017	$13,432	$44,889	$411,911	$11,552
	Approved/Declined	5,560	5,028	107	$9	$250	$1,462	$5,112	$12,883	$41,707	$154,807	$11,220
	All	13,431	12,244	282	$1	$200	$1,302	$5,069	$13,139	$44,265	$411,911	$11,407
% utilization of revolving accounts on credit report	Approved	3,994	3,641	0	1.00%	4.00%	21.00%	48.00%	76.00%	100.00%	100.00%	48.30%
	Declined	1,566	1,490	0	2.00%	4.00%	14.00%	39.50%	70.50%	98.00%	100.00%	43.24%
	In Progress	586	582	0	7.00%	7.00%	12.00%	20.50%	61.50%	99.00%	99.00%	36.75%
	Withdrawn	7,285	6,796	0	1.00%	4.00%	21.00%	48.00%	79.00%	100.00%	100.00%	50.47%
	Approved/Declined	5,560	5,131	0	1.00%	4.00%	20.00%	46.00%	75.00%	100.00%	100.00%	47.40%
	All	13,431	12,509	0	1.00%	4.00%	21.00%	47.00%	77.00%	100.00%	100.00%	48.98%
Cash Flow Metric #1	Approved	3,994	129	1,928	$1	$200	$600	$1,100	$2,000	$4,702	$175,000	$1,751
	Declined	1,566	46	694	$30	$233	$725	$1,500	$3,000	$6,620	$45,000	$2,323

	In Progress	586	443	32	$87	$180	$700	$1,300	$2,250	$7,000	$18,000	$2,059
	Withdrawn	7,285	1,015	2,818	$1	$200	$700	$1,350	$2,500	$6,651	$350,000	$2,564
	Approved/Declined	5,560	175	2,622	$1	$200	$600	$1,200	$2,200	$5,597	$175,000	$1,922
	All	13,431	1,633	5,472	$1	$200	$650	$1,250	$2,400	$6,100	$350,000	$2,275
Cash Flow Metric #2	Approved	3,994	144	1,713	$1	$100	$458	$1,500	$5,500	$30,000	$828,154	$7,463
	Declined	1,566	51	685	$1	$100	$500	$2,000	$7,000	$45,000	$480,000	$10,232
	In Progress	586	464	30	$1	$50	$500	$2,000	$6,000	$40,000	$85,947	$7,026
	Withdrawn	7,285	1,111	2,611	$1	$100	$500	$2,000	$7,000	$35,000	$13,333,330	$13,808
	Approved/Declined	5,560	195	2,398	$1	$100	$500	$1,500	$6,000	$31,239	$828,154	$8,237
	All	13,431	1,770	5,039	$1	$100	$500	$1,700	$6,300	$32,830	$13,333,330	$11,218
Cash Flow Metric #3	Approved	3,994	138	2,975	$3	$31	$108	$300	$807	$3,000	$24,000	$802
	Declined	1,566	55	1,216	$1	$35	$112	$400	$1,000	$4,000	$22,000	$1,011
	In Progress	586	474	41	$25	$50	$225	$500	$1,200	$3,500	$10,000	$1,013
	Withdrawn	7,285	1,095	4,704	$1	$45	$200	$450	$1,000	$5,000	$31,800	$1,049
	Approved/Declined	5,560	193	4,191	$1	$31	$108	$315	$900	$3,475	$24,000	$854
	All	13,431	1,762	8,936	$1	$35	$154	$400	$1,000	$4,000	$31,800	$964
Cash Flow Metric #4	Approved	3,994	89	588	$1	$1,000	$4,000	$8,500	$20,000	$60,000	$2,151,820	$19,719
	Declined	1,566	27	259	$1	$725	$4,500	$11,000	$28,000	$108,333	$27,000,000	$71,262
	In Progress	586	396	22	$1	$300	$3,200	$7,800	$20,000	$70,000	$1,000,000	$22,780
	Withdrawn	7,285	784	942	$1	$600	$4,000	$10,000	$25,000	$95,000	$35,000,000	$40,347
	Approved/Declined	5,560	116	847	$1	$901	$4,000	$9,000	$21,500	$76,000	$27,000,000	$34,071
	All	13,431	1,296	1,811	$1	$750	$4,000	$10,000	$23,543	$85,000	$35,000,000	$37,267
Cash Flow Metric #5	Approved	3,994	116	1,532	$10	$55	$200	$500	$1,350	$6,500	$320,000	$1,766
	Declined	1,566	40	633	$1	$75	$200	$500	$1,384	$8,000	$59,000	$1,914
	In Progress	586	448	19	$15	$40	$173	$320	$800	$4,200	$23,895	$1,056
	Withdrawn	7,285	989	2,420	$1	$70	$200	$500	$1,282	$6,981	$68,221,000	$19,814
	Approved/Declined	5,560	156	2,165	$1	$60	$200	$500	$1,361	$7,000	$320,000	$1,807
	All	13,431	1,593	4,604	$1	$63	$200	$500	$1,300	$7,000	$68,221,000	$11,443
Cash Flow Metric #6	Approved	3,994	189	3,411	$2	$300	$700	$1,684	$3,333	$12,060	$138,000	$3,697
	Declined	1,566	73	1,288	$100	$400	$1,000	$2,000	$4,000	$24,000	$400,000	$7,083

	In Progress	586	515	28	$55	$340	$650	$2,000	$3,500	$7,000	$42,000	$3,734
	Withdrawn	7,285	1,343	5,088	$1	$200	$800	$1,800	$4,000	$20,000	$2,300,000	$8,371
	Approved/Declined	5,560	262	4,699	$2	$300	$800	$1,760	$3,500	$15,000	$400,000	$4,856
	All	13,431	2,120	9,815	$1	$220	$800	$1,800	$4,000	$18,000	$2,300,000	$6,831
Cash Flow Metric #7	Approved	3,994	140	1,566	$1	$431	$1,442	$2,500	$4,000	$8,000	$72,902	$3,204
	Declined	1,566	225	628	$1	$325	$1,400	$2,500	$4,352	$8,000	$300,000	$4,234
	In Progress	586	514	10	$50	$250	$1,000	$2,650	$5,000	$10,000	$100,000	$5,472
	Withdrawn	7,285	1,503	2,652	$1	$500	$1,500	$2,500	$4,361	$10,000	$1,000,000	$4,512
	Approved/Declined	5,560	365	2,194	$1	$400	$1,416	$2,500	$4,000	$8,000	$300,000	$3,449
	All	13,431	2,382	4,856	$1	$450	$1,500	$2,500	$4,147	$9,000	$1,000,000	$4,006
Cash Flow Metric #8	Approved	3,994	139	2,163	$8	$300	$1,200	$3,000	$6,500	$24,000	$889,573	$7,513
	Declined	1,566	50	851	$50	$400	$1,800	$3,750	$8,000	$30,000	$720,000	$9,004
	In Progress	586	468	45	$20	$200	$1,200	$2,800	$7,083	$27,000	$36,295	$6,116
	Withdrawn	7,285	1,076	3,390	$1	$400	$1,500	$3,400	$8,000	$32,000	$20,000,000	$16,939
	Approved/Declined	5,560	189	3,014	$8	$350	$1,300	$3,000	$7,000	$25,000	$889,573	$7,934
	All	13,431	1,733	6,449	$1	$400	$1,500	$3,200	$7,600	$29,983	$20,000,000	$12,745
Cash Flow Metric #9	Approved	3,994	89	608	$1	$868	$3,727	$8,000	$20,000	$60,000	$2,146,320	$19,397
	Declined	1,566	27	268	$1	$600	$4,000	$10,271	$27,000	$108,333	$27,000,000	$70,624
	In Progress	586	396	27	$1	$300	$2,800	$7,500	$20,000	$70,000	$1,000,000	$22,494
	Withdrawn	7,285	784	990	$1	$500	$4,000	$10,000	$25,000	$90,000	$35,000,000	$39,402
	Approved/Declined	5,560	116	876	$1	$750	$3,900	$9,000	$21,000	$75,000	$27,000,000	$33,650
	All	13,431	1,296	1,893	$1	$600	$4,000	$9,800	$22,500	$83,000	$35,000,000	$36,567
Cash Flow Metric #10	Approved	3,994	89	719	$1	$300	$1,450	$3,773	$10,992	$41,800	$1,641,465	$12,047
	Declined	1,566	27	313	$1	$240	$1,500	$4,528	$13,183	$55,345	$841,500	$14,762
	In Progress	586	396	36	$20	$100	$1,050	$3,874	$12,000	$42,600	$136,663	$10,262
	Withdrawn	7,285	783	1,229	$1	$250	$1,475	$4,200	$12,400	$50,984	$68,221,000	$41,566
	Approved/Declined	5,560	116	1,032	$1	$283	$1,459	$3,979	$11,350	$45,050	$1,641,465	$12,802
	All	13,431	1,295	2,297	$1	$250	$1,457	$4,085	$11,850	$47,980	$68,221,000	$28,177
Cash Flow Metric #11	Approved	3,994	132	2,090	$10	$60	$150	$300	$540	$1,512	$17,800	$514
	Declined	1,566	46	751	$10	$75	$200	$350	$750	$2,400	$80,000	$852

	In Progress	586	456	35	$25	$40	$160	$350	$600	$2,200	$6,300	$646
	Withdrawn	7,285	1,040	3,077	$1	$50	$200	$350	$680	$2,000	$49,500	$725
	Approved/Declined	5,560	178	2,841	$10	$60	$165	$300	$600	$1,857	$80,000	$617
	All	13,431	1,674	5,953	$1	$50	$189	$325	$600	$2,000	$80,000	$676
Cash Flow Metric #12	Approved	3,994	186	3,650	$1	$50	$200	$400	$600	$1,500	$6,000	$558
	Declined	1,566	61	1,429	$20	$100	$200	$400	$710	$2,000	$2,700	$559
	In Progress	586	552	28	$1	$1	$1	$175	$200	$300	$300	$142
	Withdrawn	7,285	1,271	5,736	$1	$50	$170	$378	$600	$1,500	$35,000	$858
	Approved/Declined	5,560	247	5,079	$1	$50	$200	$400	$650	$1,588	$6,000	$559
	All	13,431	2,070	10,843	$1	$50	$189	$400	$600	$1,500	$35,000	$715
Cash Flow Metric #13	Approved	3,994	124	1,200	$1	$25	$89	$200	$400	$1,116	$8,000	$337
	Declined	1,566	40	861	$14	$25	$100	$200	$422	$1,377	$5,000	$382
	In Progress	586	529	8	$1	$25	$100	$151	$500	$1,500	$2,300	$386
	Withdrawn	7,285	1,041	2,881	$1	$25	$100	$200	$471	$1,191	$24,000	$388
	Approved/Declined	5,560	164	2,061	$1	$25	$92	$200	$400	$1,188	$8,000	$346
	All	13,431	1,734	4,950	$1	$25	$100	$200	$436	$1,198	$24,000	$367
Cash Flow Metric #14	Approved	3,994	167	3,233	$1	$50	$140	$300	$520	$1,500	$12,000	$460
	Declined	1,566	55	1,296	$25	$50	$100	$250	$600	$1,500	$5,000	$472
	In Progress	586	542	27	$1	$1	$100	$207	$800	$2,000	$2,000	$512
	Withdrawn	7,285	1,183	5,112	$1	$50	$125	$300	$597	$2,000	$35,000	$607
	Approved/Declined	5,560	222	4,529	$1	$50	$125	$280	$520	$1,500	$12,000	$463
	All	13,431	1,947	9,668	$1	$50	$125	$300	$558	$1,600	$35,000	$542
Cash Flow Metric #15	Approved	3,994	150	1,938	$1	$750	$1,723	$2,700	$4,320	$8,083	$80,000	$3,745
	Declined	1,566	48	826	$1	$600	$1,600	$2,800	$4,800	$14,500	$135,000	$5,723
	In Progress	586	538	12	$1	$140	$1,500	$2,300	$3,000	$5,000	$40,000	$3,317
	Withdrawn	7,285	1,167	3,199	$1	$600	$1,600	$2,955	$4,500	$22,746	$350,000	$6,308
	Approved/Declined	5,560	198	2,764	$1	$708	$1,700	$2,734	$4,465	$8,720	$135,000	$4,272
	All	13,431	1,903	5,975	$1	$600	$1,650	$2,800	$4,500	$11,000	$350,000	$5,336
Cash Flow Metric #16	Approved	3,994	88	146	$1	$100	$200	$300	$450	$900	$10,400	$366
	Declined	1,566	27	98	$1	$100	$200	$300	$500	$1,000	$60,000	$461

	In Progress	586	514	1	$5	$57	$150	$300	$500	$1,000	$23,000	$695
	Withdrawn	7,285	856	365	$1	$100	$200	$300	$500	$1,200	$40,000	$470
	Approved/Declined	5,560	115	244	$1	$100	$200	$300	$500	$1,000	$60,000	$392
	All	13,431	1,485	610	$1	$100	$200	$300	$500	$1,000	$60,000	$435
Cash Flow Metric #17	Approved	3,994	115	600	$1	$300	$665	$1,000	$1,563	$3,000	$22,800	$1,260
	Declined	1,566	32	415	$11	$300	$670	$1,050	$1,617	$3,200	$29,000	$1,328
	In Progress	586	523	8	$1	$240	$500	$850	$1,500	$2,500	$2,902	$1,039
	Withdrawn	7,285	961	1,661	$1	$295	$650	$1,000	$1,600	$3,000	$60,000	$1,320
	Approved/Declined	5,560	147	1,015	$1	$300	$666	$1,000	$1,600	$3,000	$29,000	$1,277
	All	13,431	1,631	2,684	$1	$300	$650	$1,000	$1,600	$3,000	$60,000	$1,298
Cash Flow Metric #18	Approved	3,994	121	848	$1	$50	$125	$200	$355	$750	$4,000	$293
	Declined	1,566	39	457	$10	$58	$135	$200	$400	$900	$7,700	$318
	In Progress	586	530	6	$1	$50	$100	$195	$320	$700	$1,000	$250
	Withdrawn	7,285	1,038	1,801	$1	$60	$150	$200	$400	$900	$12,000	$327
	Approved/Declined	5,560	160	1,305	$1	$54	$125	$200	$373	$790	$7,700	$299
	All	13,431	1,728	3,112	$1	$58	$140	$200	$400	$800	$12,000	$313
Cash Flow Metric #19	Approved	3,994	3,658	196	$43	$80	$150	$250	$500	$1,575	$5,000	$484
	Declined	1,566	1,496	39	$2	$6	$100	$205	$680	$1,500	$2,000	$430
	In Progress	586	557	14	$1	$1	$100	$200	$500	$1,450	$1,450	$383
	Withdrawn	7,285	6,706	391	$1	$50	$150	$300	$520	$2,000	$3,000	$487
	Approved/Declined	5,560	5,154	235	$2	$60	$150	$250	$600	$1,500	$5,000	$474
	All	13,431	12,417	640	$1	$50	$150	$272	$540	$1,650	$5,000	$477
Cash Flow Metric #20	Approved	3,994	175	2,494	$1	$286	$700	$1,200	$2,295	$5,556	$30,000	$1,873
	Declined	1,566	69	983	$50	$225	$600	$1,100	$2,000	$5,000	$53,000	$1,796
	In Progress	586	557	15	$192	$192	$735	$1,198	$3,000	$5,000	$5,000	$1,908
	Withdrawn	7,285	1,324	4,059	$1	$258	$700	$1,300	$2,300	$7,330	$720,000	$3,120
	Approved/Declined	5,560	244	3,477	$1	$250	$686	$1,200	$2,200	$5,415	$53,000	$1,852
	All	13,431	2,125	7,551	$1	$250	$700	$1,200	$2,250	$6,000	$720,000	$2,494
Cash Flow Metric #21	Approved	3,994	169	2,837	$1	$721	$1,800	$2,800	$4,700	$9,000	$240,000	$4,254
	Declined	1,566	59	1,141	$55	$600	$1,558	$3,000	$4,500	$14,000	$140,000	$5,624

	In Progress	586	543	19	$500	$500	$1,800	$3,150	$5,217	$40,000	$40,000	$6,328
	Withdrawn	7,285	1,195	4,560	$1	$800	$1,875	$3,000	$5,000	$30,000	$2,083,000	$8,777
	Approved/Declined	5,560	228	3,978	$1	$698	$1,714	$2,900	$4,654	$9,068	$240,000	$4,624
	All	13,431	1,966	8,557	$1	$733	$1,800	$3,000	$4,999	$12,887	$2,083,000	$6,823
Cash Flow Metric #22	Approved	3,994	88	21	$9	$534	$1,316	$2,132	$3,260	$5,930	$62,000	$2,553
	Declined	1,566	27	34	$2	$305	$1,099	$1,964	$3,200	$5,750	$67,365	$2,451
	In Progress	586	514	1	$65	$255	$1,105	$2,200	$3,640	$5,750	$46,000	$3,019
	Withdrawn	7,285	855	118	$2	$370	$1,150	$2,001	$3,200	$6,150	$191,600	$2,584
	Approved/Declined	5,560	115	55	$2	$470	$1,265	$2,098	$3,250	$5,886	$67,365	$2,524
	All	13,431	1,484	174	$2	$400	$1,200	$2,050	$3,222	$6,050	$191,600	$2,559
Cash Flow Metric #23	Approved	3,994	88	30	$4	$1,290	$2,641	$4,100	$6,500	$12,000	$240,000	$5,474
	Declined	1,566	26	98	$1	$1,000	$2,400	$4,000	$6,768	$16,586	$300,000	$7,277
	In Progress	586	514	2	$176	$1,500	$3,000	$5,000	$8,600	$40,000	$100,000	$9,104
	Withdrawn	7,285	855	335	$1	$1,000	$2,500	$4,100	$7,000	$20,000	$2,084,000	$8,971
	Approved/Declined	5,560	114	128	$1	$1,151	$2,560	$4,008	$6,527	$12,679	$300,000	$5,962
	All	13,431	1,483	465	$1	$1,000	$2,500	$4,100	$6,733	$15,000	$2,084,000	$7,579
Cash Flow Metric #24	Approved	3,994	88	349	$1	$50	$120	$200	$300	$590	$4,800	$255
	Declined	1,566	26	140	$1	$25	$125	$213	$350	$681	$7,000	$284
	In Progress	586	514	9	$1	$50	$135	$200	$350	$850	$23,000	$655
	Withdrawn	7,285	855	673	$1	$20	$134	$225	$360	$700	$43,440	$306
	Approved/Declined	5,560	114	489	$1	$50	$125	$200	$325	$600	$7,000	$263
	All	13,431	1,483	1,171	$1	$40	$125	$200	$350	$628	$43,440	$289
Cash Flow Metric #25	Approved	3,994	132	1,639	$1	$120	$300	$450	$699	$1,373	$12,000	$576
	Declined	1,566	42	854	$25	$129	$300	$430	$650	$1,500	$4,000	$556
	In Progress	586	529	18	$1	$5	$245	$460	$800	$1,600	$2,002	$541
	Withdrawn	7,285	1,044	3,229	$1	$136	$302	$450	$661	$1,347	$30,000	$603
	Approved/Declined	5,560	174	2,493	$1	$125	$300	$448	$684	$1,400	$12,000	$572
	All	13,431	1,747	5,740	$1	$125	$300	$450	$680	$1,384	$30,000	$587

Appendix F. Participant #6

Table 2. Difference of Means Tests: All Applications[69]

Variable	Sample	#	Mean	T-Stat	P-Value
Date Difference #1	Approved	3,994	29.9	.	.
	Declined	1,566	24.9	3.93	0.000
FICO score	Approved	3,687	642.3	.	.
	Declined	1,189	583.6	22.81	0.000
BK score	Approved	3,763	412.1	.	.
	Declined	1,224	278.2	17.91	0.000
# of open accounts on credit report	Approved	429	1.80	.	.
	Declined	103	2.05	0.01	0.994
$ amount of unpaid balances on credit report	Approved	3,830	$95,478	.	.
	Declined	1,242	$113,981	0.93	0.354
$ amount of monthly payments on credit report	Approved	3,830	$1,311	.	.
	Declined	1,242	$1,349	1.82	0.069
$ Credit limit of revolving accounts on credit report	Approved	353	$93,686	.	.
	Declined	76	$58,194	1.66	0.098
$ unpaid balances of revolving accounts on credit report	Approved	429	$11,096	.	.
	Declined	103	$11,780	0.12	0.903
% utilization of revolving accounts on credit report	Approved	353	48.30%	.	.
	Declined	76	43.24%	1.28	0.202
Cash Flow Metric #1	Approved	3,865	$1,751	.	.
	Declined	1,520	$2,323	-4.47	0.000
Cash Flow Metric #2	Approved	3,850	$7,463	.	.
	Declined	1,515	$10,232	-2.00	0.046
Cash Flow Metric #3	Approved	3,856	$802	.	.
	Declined	1,511	$1,011	-0.50	0.615
Cash Flow Metric #4	Approved	3,905	$19,719	.	.
	Declined	1,539	$71,262	-2.10	0.036
Cash Flow Metric #5	Approved	3,878	$1,766	.	.
	Declined	1,526	$1,914	-0.36	0.721
Cash Flow Metric #6	Approved	3,805	$3,697	.	.
	Declined	1,493	$7,083	-1.99	0.047
Cash Flow Metric #7	Approved	3,854	$3,204	.	.
	Declined	1,341	$4,234	-1.31	0.191

69 The significance test tests the difference in means between the approved and declined populations using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% level. Counts in this table are of non-missing values of the indicated variable.

Cash Flow Metric #8	Approved	3,855	$7,513	.	.
	Declined	1,516	$9,004	-1.05	0.295
Cash Flow Metric #9	Approved	3,905	$19,397	.	.
	Declined	1,539	$70,624	-2.08	0.038
Cash Flow Metric #10	Approved	3,905	$12,047	.	.
	Declined	1,539	$14,762	-1.66	0.096
Cash Flow Metric #11	Approved	3,862	$514	.	.
	Declined	1,520	$852	-3.03	0.002
Cash Flow Metric #12	Approved	3,808	$558	.	.
	Declined	1,505	$559	-0.95	0.344
Cash Flow Metric #13	Approved	3,870	$337	.	.
	Declined	1,526	$382	5.42	0.000
Cash Flow Metric #14	Approved	3,827	$460	.	.
	Declined	1,511	$472	0.46	0.646
Cash Flow Metric #15	Approved	3,844	$3,745	.	.
	Declined	1,518	$5,723	-3.06	0.002
Cash Flow Metric #16	Approved	3,906	$366	.	.
	Declined	1,539	$461	-1.92	0.055
Cash Flow Metric #17	Approved	3,879	$1,260	.	.
	Declined	1,534	$1,328	2.58	0.010
Cash Flow Metric #18	Approved	3,873	$293	.	.
	Declined	1,527	$318	0.61	0.542
Cash Flow Metric #19	Approved	336	$484	.	.
	Declined	70	$430	0.22	0.825
Cash Flow Metric #20	Approved	3,819	$1,873	.	.
	Declined	1,497	$1,796	0.58	0.560
Cash Flow Metric #21	Approved	3,825	$4,254	.	.
	Declined	1,507	$5,624	-1.35	0.179
Cash Flow Metric #22	Approved	3,906	$2,553	.	.
	Declined	1,539	$2,451	1.90	0.058
Cash Flow Metric #23	Approved	3,906	$5,474	.	.
	Declined	1,540	$7,277	-3.15	0.002
Cash Flow Metric #24	Approved	3,906	$255	.	.
	Declined	1,540	$284	-2.94	0.003
Cash Flow Metric #25	Approved	3,862	$576	.	.
	Declined	1,524	$556	6.63	0.000

Appendix F. Participant #6

Table 3. Data Diagnostics: Originated Loans

Variable	Sample	#	# Missing	# Zero	Min	5th%	25th%	50th%	75th%	95th%	Max	Mean
Date Difference #1	Delinquent	517	0	7	1	4	12	22	37	85	296	31.1
	Not Delinquent	3,259	0	36	1	3	11	20	35	83	418	29.1
	All	3,776	0	43	1	3	11	21	35	83	418	29.4
FICO score	Delinquent	517	51	0	431	499	560	598	638	704	833	600.1
	Not Delinquent	3,259	232	0	439	527	604	648	689	759	847	646.4
	All	3,776	283	0	431	522	597	641	685	756	847	640.2
BK score	Delinquent	517	34	0	19	115	145	224	381	682	989	293.8
	Not Delinquent	3,259	180	1	1	117	199	399	619	881	993	424.5
	All	3,776	214	1	1	116	185	362	597	860	993	406.8
# of open accounts on credit report	Delinquent	517	501	9	1	1	1	1	3	4	4	1.9
	Not Delinquent	3,259	2,865	209	1	1	1	1	2	4	10	1.8
	All	3,776	3,366	218	1	1	1	1	2	4	10	1.8
$ amount of unpaid balances on credit report	Delinquent	517	23	480	$197	$197	$22,016	$68,763	$188,772	$642,103	$642,103	$132,554
	Not Delinquent	3,259	129	2,765	$39	$1,122	$13,909	$42,672	$130,066	$353,878	$931,802	$94,788
	All	3,776	152	3,245	$39	$891	$13,949	$42,854	$136,028	$374,023	$931,802	$96,183
$ amount of monthly payments on credit report	Delinquent	517	23	482	$3	$3	$341	$733	$1,873	$7,762	$7,762	$1,559
	Not Delinquent	3,259	129	2,771	$12	$56	$344	$716	$1,664	$3,802	$34,580	$1,309
	All	3,776	152	3,253	$3	$56	$344	$716	$1,664	$3,802	$34,580	$1,317
$ Credit limit of revolving accounts on credit report	Delinquent	517	507	0	$492	$492	$816	$18,774	$126,686	$344,771	$344,771	$75,529
	Not Delinquent	3,259	2,936	14	$9	$340	$3,854	$15,269	$38,727	$410,736	$3,294,300	$85,322
	All	3,776	3,443	14	$9	$340	$3,767	$15,269	$40,381	$410,736	$3,294,300	$85,015
$ unpaid balances of revolving accounts on credit report	Delinquent	517	501	6	$197	$197	$349	$2,255	$17,666	$72,402	$72,402	$12,691
	Not Delinquent	3,259	2,865	74	$9	$237	$1,434	$4,581	$11,389	$41,095	$154,807	$10,719
	All	3,776	3,366	80	$9	$232	$1,431	$4,533	$11,429	$41,644	$154,807	$10,778
	Delinquent	517	507	0	4.00%	4.00%	7.00%	27.50%	71.00%	95.00%	95.00%	36.20%
	Not Delinquent	3,259	2,936	0	1.00%	5.00%	22.00%	49.00%	76.00%	100.00%	100.00%	49.22%

% utilization of revolving accounts on credit report	All	3,776	3,443	0	1.00%	4.00%	21.00%	48.00%	76.00%	100.00%	100.00%	48.83%
Cash Flow Metric #1	Delinquent	517	18	274	$70	$200	$500	$875	$1,500	$2,700	$8,000	$1,131
	Not Delinquent	3,259	104	1,544	$1	$200	$600	$1,100	$2,000	$4,994	$50,000	$1,693
	All	3,776	122	1,818	$1	$200	$580	$1,069	$2,000	$4,583	$50,000	$1,624
Cash Flow Metric #2	Delinquent	517	21	239	$25	$75	$250	$750	$2,200	$15,000	$70,000	$3,354
	Not Delinquent	3,259	115	1,369	$1	$100	$500	$1,640	$6,000	$29,200	$828,154	$7,693
	All	3,776	136	1,608	$1	$100	$450	$1,500	$5,000	$27,040	$828,154	$7,144
Cash Flow Metric #3	Delinquent	517	19	412	$15	$29	$80	$200	$500	$2,000	$12,000	$586
	Not Delinquent	3,259	111	2,414	$1	$30	$143	$300	$800	$3,000	$24,000	$779
	All	3,776	130	2,826	$1	$30	$100	$300	$750	$3,000	$24,000	$759
Cash Flow Metric #4	Delinquent	517	11	94	$1	$1,000	$3,000	$5,687	$10,400	$40,000	$322,000	$11,356
	Not Delinquent	3,259	74	460	$1	$950	$4,000	$9,000	$20,000	$65,000	$1,600,000	$20,058
	All	3,776	85	554	$1	$950	$4,000	$8,100	$19,131	$60,000	$1,600,000	$18,915
Cash Flow Metric #5	Delinquent	517	17	230	$10	$50	$150	$350	$933	$4,413	$23,000	$1,064
	Not Delinquent	3,259	93	1,223	$1	$55	$200	$490	$1,350	$6,700	$320,000	$1,877
	All	3,776	110	1,453	$1	$54	$185	$450	$1,250	$6,215	$320,000	$1,778
Cash Flow Metric #6	Delinquent	517	25	426	$100	$200	$500	$1,000	$2,000	$5,000	$15,000	$1,693
	Not Delinquent	3,259	149	2,804	$2	$400	$800	$1,775	$4,000	$12,000	$138,000	$4,058
	All	3,776	174	3,230	$2	$300	$697	$1,500	$3,220	$12,000	$138,000	$3,639
Cash Flow Metric #7	Delinquent	517	25	201	$50	$400	$1,416	$2,260	$3,900	$7,000	$27,000	$2,893
	Not Delinquent	3,259	104	1,277	$1	$418	$1,400	$2,500	$4,000	$8,000	$72,902	$3,142
	All	3,776	129	1,478	$1	$418	$1,400	$2,500	$4,000	$8,000	$72,902	$3,108
Cash Flow Metric #8	Delinquent	517	20	316	$100	$300	$1,000	$2,000	$4,500	$15,000	$67,450	$4,628
	Not Delinquent	3,259	111	1,730	$1	$300	$1,200	$3,000	$6,500	$24,000	$450,000	$7,237
	All	3,776	131	2,046	$1	$300	$1,200	$2,900	$6,250	$23,916	$450,000	$6,942
Cash Flow Metric #9	Delinquent	517	11	96	$1	$700	$3,000	$5,000	$10,000	$40,000	$322,000	$11,139
	Not Delinquent	3,259	74	474	$1	$893	$3,975	$8,600	$20,000	$65,000	$1,600,000	$19,704
	All	3,776	85	570	$1	$825	$3,600	$8,000	$19,000	$60,000	$1,600,000	$18,578

Cash Flow Metric #10	Delinquent	517	11	112	$25	$200	$950	$2,362	$6,125	$24,598	$106,000	$6,165
	Not Delinquent	3,259	74	551	$1	$300	$1,500	$4,000	$11,250	$43,286	$920,400	$12,201
	All	3,776	85	663	$1	$286	$1,400	$3,650	$10,608	$40,907	$920,400	$11,416
Cash Flow Metric #11	Delinquent	517	18	289	$20	$50	$120	$209	$492	$1,200	$5,000	$390
	Not Delinquent	3,259	107	1,696	$1	$50	$150	$300	$550	$1,600	$17,800	$520
	All	3,776	125	1,985	$1	$50	$150	$300	$525	$1,500	$17,800	$504
Cash Flow Metric #12	Delinquent	517	23	473	$25	$30	$100	$300	$450	$750	$958	$330
	Not Delinquent	3,259	151	2,981	$1	$50	$200	$400	$600	$1,500	$6,000	$576
	All	3,776	174	3,454	$1	$50	$178	$379	$600	$1,469	$6,000	$541
Cash Flow Metric #13	Delinquent	517	16	230	$8	$25	$50	$101	$257	$817	$2,978	$235
	Not Delinquent	3,259	101	924	$1	$25	$97	$200	$400	$1,116	$8,000	$339
	All	3,776	117	1,154	$1	$25	$85	$200	$400	$1,095	$8,000	$327
Cash Flow Metric #14	Delinquent	517	23	432	$30	$42	$100	$200	$303	$800	$2,000	$264
	Not Delinquent	3,259	134	2,639	$1	$50	$145	$300	$520	$1,500	$12,000	$465
	All	3,776	157	3,071	$1	$50	$125	$255	$500	$1,471	$12,000	$442
Cash Flow Metric #15	Delinquent	517	19	248	$80	$548	$1,405	$2,400	$3,750	$6,240	$63,000	$3,272
	Not Delinquent	3,259	124	1,587	$1	$765	$1,739	$2,761	$4,262	$8,000	$80,000	$3,653
	All	3,776	143	1,835	$1	$725	$1,700	$2,660	$4,200	$8,000	$80,000	$3,600
Cash Flow Metric #16	Delinquent	517	11	27	$1	$100	$160	$250	$400	$650	$2,000	$294
	Not Delinquent	3,259	73	106	$1	$100	$200	$300	$490	$900	$10,400	$369
	All	3,776	84	133	$1	$100	$200	$300	$450	$800	$10,400	$359
Cash Flow Metric #17	Delinquent	517	15	101	$21	$250	$500	$805	$1,240	$2,237	$10,256	$986
	Not Delinquent	3,259	92	477	$1	$300	$688	$1,004	$1,600	$3,000	$22,800	$1,284
	All	3,776	107	578	$1	$300	$650	$1,000	$1,511	$2,950	$22,800	$1,245
Cash Flow Metric #18	Delinquent	517	16	137	$10	$59	$100	$200	$300	$675	$2,050	$244
	Not Delinquent	3,259	99	686	$1	$50	$130	$200	$370	$750	$4,000	$296
	All	3,776	115	823	$1	$50	$125	$200	$350	$728	$4,000	$290
Cash Flow Metric #19	Delinquent	517	491	15	$75	$75	$100	$170	$200	$608	$608	$201
	Not Delinquent	3,259	2,973	167	$43	$80	$150	$300	$600	$1,650	$5,000	$513
	All	3,776	3,464	182	$43	$80	$150	$263	$500	$1,500	$5,000	$487

Cash Flow Metric #20	Delinquent	517	19	322	$50	$190	$600	$877	$1,800	$4,397	$16,000	$1,409
	Not Delinquent	3,259	145	2,024	$1	$300	$721	$1,250	$2,400	$5,556	$30,000	$1,936
	All	3,776	164	2,346	$1	$291	$700	$1,200	$2,300	$5,415	$30,000	$1,862
Cash Flow Metric #21	Delinquent	517	22	411	$150	$750	$1,490	$2,490	$3,731	$7,000	$8,000	$2,835
	Not Delinquent	3,259	137	2,300	$1	$720	$1,750	$2,896	$4,678	$8,800	$240,000	$4,168
	All	3,776	159	2,711	$1	$720	$1,700	$2,779	$4,561	$8,400	$240,000	$4,045
Cash Flow Metric #22	Delinquent	517	11	5	$40	$400	$1,016	$1,660	$2,410	$4,535	$11,622	$1,927
	Not Delinquent	3,259	73	11	$9	$556	$1,350	$2,170	$3,287	$6,050	$62,000	$2,593
	All	3,776	84	16	$9	$518	$1,300	$2,097	$3,200	$5,800	$62,000	$2,502
Cash Flow Metric #23	Delinquent	517	11	6	$400	$1,000	$2,015	$3,258	$5,173	$10,018	$65,000	$4,312
	Not Delinquent	3,259	73	21	$4	$1,300	$2,731	$4,147	$6,500	$11,833	$240,000	$5,414
	All	3,776	84	27	$4	$1,200	$2,600	$4,000	$6,239	$11,619	$240,000	$5,263
Cash Flow Metric #24	Delinquent	517	11	57	$1	$50	$100	$200	$300	$600	$2,000	$238
	Not Delinquent	3,259	73	278	$1	$50	$120	$200	$300	$570	$4,800	$253
	All	3,776	84	335	$1	$50	$120	$200	$300	$575	$4,800	$251
Cash Flow Metric #25	Delinquent	517	14	219	$10	$100	$300	$420	$600	$1,281	$2,033	$503
	Not Delinquent	3,259	110	1,338	$1	$125	$300	$450	$700	$1,393	$12,000	$584
	All	3,776	124	1,557	$1	$115	$300	$450	$683	$1,373	$12,000	$573

Appendix F. Participant #6 Table 4. Difference of Means Tests: Originated Loans[70]					
Variable	**Sample**	**#**	**Mean**	**T-Stat**	**P-Value**
Date Difference #1	Delinquent	517	31.1	.	.
	Not Delinquent	3,259	29.1	-1.12	0.262
FICO score	Delinquent	466	600.1	.	.
	Not Delinquent	3,027	646.4	14.95	0.000
BK score	Delinquent	483	293.8	.	.
	Not Delinquent	3,079	424.5	13.35	0.000
# of open accounts on credit report	Delinquent	16	1.9	.	.
	Not Delinquent	394	1.8	0.12	0.905
$ amount of unpaid balances on credit report	Delinquent	494	$132,554	.	.
	Not Delinquent	3,130	$94,788	3.91	0.000
$ amount of monthly payments on credit report	Delinquent	494	$1,559	.	.
	Not Delinquent	3,130	$1,309	4.60	0.000
$ Credit limit of revolving accounts on credit report	Delinquent	10	$75,529	.	.
	Not Delinquent	323	$85,322	0.15	0.881
$ unpaid balances of revolving accounts on credit report	Delinquent	16	$12,691	.	.
	Not Delinquent	394	$10,719	0.16	0.871
% utilization of revolving accounts on credit report	Delinquent	10	36%	.	.
	Not Delinquent	323	49%	1.23	0.247
Cash Flow Metric #1	Delinquent	499	$1,131	.	.
	Not Delinquent	3,155	$1,693	6.75	0.000
Cash Flow Metric #2	Delinquent	496	$3,354	.	.
	Not Delinquent	3,144	$7,693	5.20	0.000
Cash Flow Metric #3	Delinquent	498	$586	.	.
	Not Delinquent	3,148	$779	2.48	0.013
Cash Flow Metric #4	Delinquent	506	$11,356	.	.
	Not Delinquent	3,185	$20,058	6.30	0.000
Cash Flow Metric #5	Delinquent	500	$1,064	.	.
	Not Delinquent	3,166	$1,877	3.80	0.000
Cash Flow Metric #6	Delinquent	492	$1,693	.	.
	Not Delinquent	3,110	$4,058	2.27	0.023
Cash Flow Metric #7	Delinquent	492	$2,893	.	.
	Not Delinquent	3,155	$3,142	1.30	0.193

[70] The significance test tests the difference in means between the delinquent and non-delinquent populations using Student's T-test, assuming unequal variance. Yellow highlighting indicates statistical significance at the 95% level. Counts in this table are of non-missing values of the indicated variable.

Cash Flow Metric #8	Delinquent	497	$4,628	.	.
	Not Delinquent	3,148	$7,237	4.30	0.000
Cash Flow Metric #9	Delinquent	506	$11,139	.	.
	Not Delinquent	3,185	$19,704	6.21	0.000
Cash Flow Metric #10	Delinquent	506	$6,165	.	.
	Not Delinquent	3,185	$12,201	6.88	0.000
Cash Flow Metric #11	Delinquent	499	$390	.	.
	Not Delinquent	3,152	$520	3.59	0.000
Cash Flow Metric #12	Delinquent	494	$330	.	.
	Not Delinquent	3,108	$576	1.83	0.067
Cash Flow Metric #13	Delinquent	501	$235	.	.
	Not Delinquent	3,158	$339	7.45	0.000
Cash Flow Metric #14	Delinquent	494	$264	.	.
	Not Delinquent	3,125	$465	4.53	0.000
Cash Flow Metric #15	Delinquent	498	$3,272	.	.
	Not Delinquent	3,135	$3,653	0.79	0.430
Cash Flow Metric #16	Delinquent	506	$294	.	.
	Not Delinquent	3,186	$369	6.88	0.000
Cash Flow Metric #17	Delinquent	502	$986	.	.
	Not Delinquent	3,167	$1,284	7.43	0.000
Cash Flow Metric #18	Delinquent	501	$244	.	.
	Not Delinquent	3,160	$296	5.05	0.000
Cash Flow Metric #19	Delinquent	26	$201	.	.
	Not Delinquent	286	$513	3.16	0.002
Cash Flow Metric #20	Delinquent	498	$1,409	.	.
	Not Delinquent	3,114	$1,936	2.94	0.003
Cash Flow Metric #21	Delinquent	495	$2,835	.	.
	Not Delinquent	3,122	$4,168	5.53	0.000
Cash Flow Metric #22	Delinquent	506	$1,927	.	.
	Not Delinquent	3,186	$2,593	9.36	0.000
Cash Flow Metric #23	Delinquent	506	$4,312	.	.
	Not Delinquent	3,186	$5,414	4.41	0.000
Cash Flow Metric #24	Delinquent	506	$238	.	.
	Not Delinquent	3,186	$253	1.98	0.048
Cash Flow Metric #25	Delinquent	503	$503	.	.
	Not Delinquent	3,149	$584	2.75	0.006

Appendix F. Participant #6

Table 5. Logistic Model for Delinquency Specifications[71]

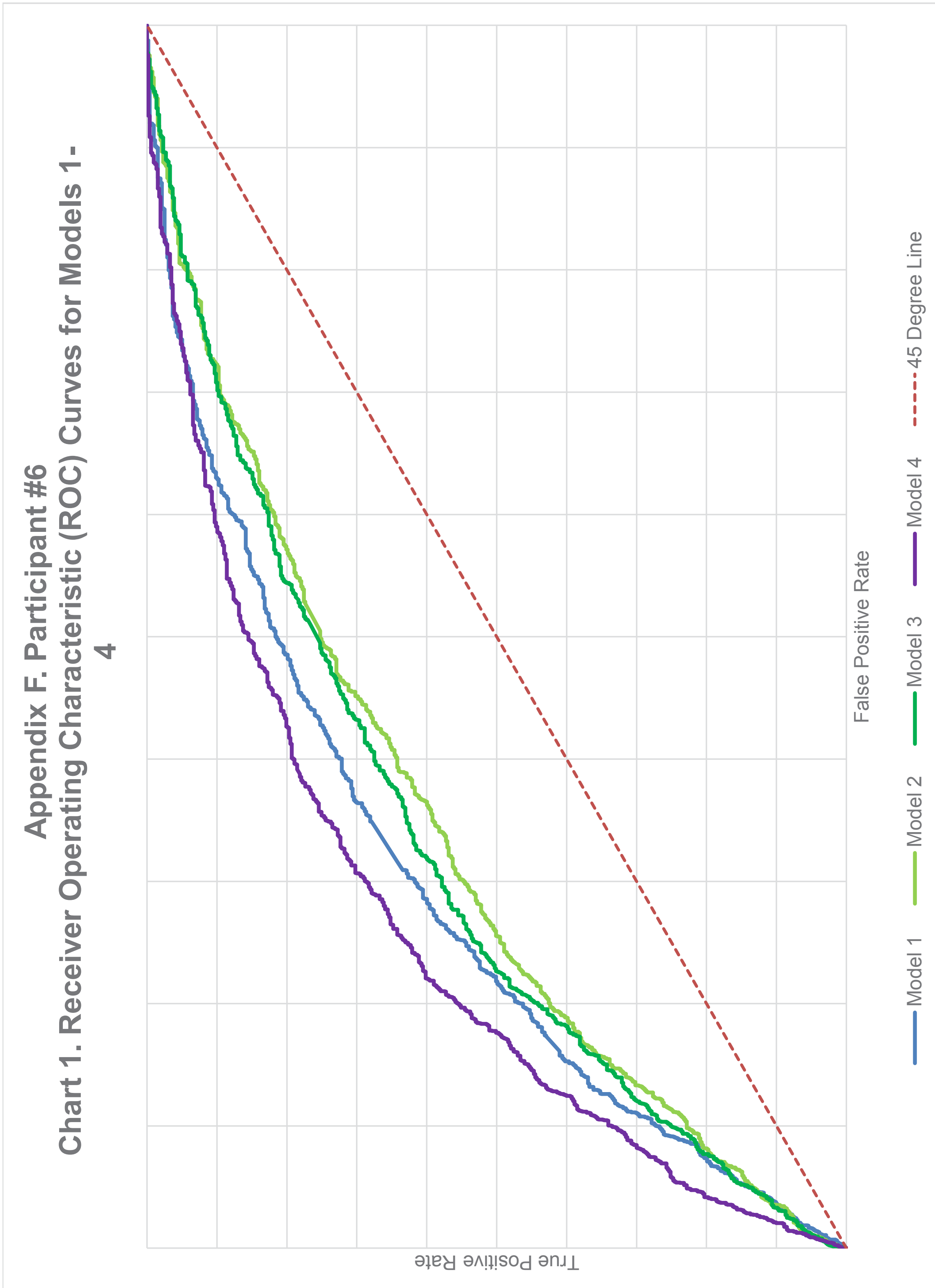
		Model 1		Model 2		Model 3		Model 4	
Control Variable	**Comparison Group**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**	**Odds Ratio**	**P-Value**
Hard Pull Not Available	Hard Pull Available	0.67	0.33	.	.	.	.	0.72	0.43
FICO score	--	0.99	0.00	.	.	.	.	0.99	0.00
Missing FICO score	Not Missing FICO score	2.42	0.00	.	.	.	.	1.95	0.03
BK score	--	1.00	0.00	.	.	.	.	1.00	0.00
Missing BK score	Not Missing BK score	0.77	0.56	.	.	.	.	0.82	0.66
# of open accounts on credit report	--	1.11	0.52	.	.	.	.	1.12	0.52
Missing # of open accounts on credit report	Not Missing # of open accounts on credit report	4.51	0.00	.	.	.	.	4.62	0.00
$ amount of unpaid balances on credit report	--	1.00	0.01	.	.	.	.	1.00	0.01
Missing $ amount of unpaid balances on credit report	Not Missing $ amount of unpaid balances on credit report	.	.	.	.	.	.	.	.
$ amount of monthly payments on credit report	--	1.00	0.50	.	.	.	.	1.00	0.48
Missing $ amount of monthly payments on credit report	Not Missing $ amount of monthly payments on credit report	.	.	.	.	.	.	.	.
$ Credit limit of revolving accounts on credit report	--	1.00	0.18	.	.	.	.	1.00	0.10

71 The dependent variable is a 0/1 indicator for delinquent, with values of 1 indicating delinquent and 0 indicating not delinquent. Model 1 includes traditional credit fields that were pulled from the credit bureau. Model 2 includes all cash flow fields whose delinquent population mean was statistically different from the not delinquent population mean (see table 4). Model 3 includes all cash flow fields. Model 4 includes all credit bureau and cash flow fields. Many cash flow variables' units have been transformed so their associated odds ratios are more interpretable.

Missing $ Credit limit of revolving accounts on credit report	Not Missing $ Credit limit of revolving accounts on credit report	1.80	0.41	.	.	.	.	1.77	0.37
$ unpaid balances of revolving accounts on credit report	--	1.00	0.40	.	.	.	.	1.00	0.09
Missing $ unpaid balances of revolving accounts on credit report	Not Missing $ unpaid balances of revolving accounts on credit report	.	.	.	.	.	.	.	.
% utilization of revolving accounts on credit report	--	0.38	0.53	.	.	.	.	0.39	0.51
Missing % utilization of revolving accounts on credit report	Not Missing % utilization of revolving accounts on credit report	.	.	.	.	.	.	.	.
Cash Flow Metric #1	--	.	.	0.83	0.01	0.84	0.02	0.85	0.03
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	.	.	1.54	0.47	1.73	0.39	1.69	0.47
Cash Flow Metric #2	--	.	.	0.94	0.19	0.95	0.28	0.93	0.10
Missing Cash Flow Metric #2	Not Missing Cash Flow Metric #2	.	.	2.05	0.28	2.02	0.32	2.67	0.14
Cash Flow Metric #3	--	.	.	1.02	0.90	1.02	0.90	1.08	0.51
Missing Cash Flow Metric #3	Not Missing Cash Flow Metric #3	.	.	1.07	0.93	1.10	0.91	1.18	0.84
Cash Flow Metric #4	--	.	.	1.00	0.44	0.99	0.36	1.00	0.42
Missing Cash Flow Metric #4	Not Missing Cash Flow Metric #4	.	.	0.00	0.00	0.00	0.00	0.00	0.00
Cash Flow Metric #5	--	.	.	0.91	0.11	0.92	0.18	0.90	0.08
Missing Cash Flow Metric #5	Not Missing Cash Flow Metric #5	.	.	1.73	0.50	1.87	0.43	1.32	0.74
Cash Flow Metric #6	--	.	.	0.99	0.60	0.99	0.62	0.98	0.55
Missing Cash Flow Metric #6	Not Missing Cash Flow Metric #6	.	.	2.38	0.20	3.41	0.16	2.65	0.24
Cash Flow Metric #8	--	.	.	0.95	0.30	0.96	0.39	0.94	0.17
Missing Cash Flow Metric #8	Not Missing Cash Flow Metric #8	.	.	1.61	0.42	1.53	0.44	1.88	0.24
Cash Flow Metric #9	--	.	.	.	.	.	.	.	.
Missing Cash Flow Metric #9	Not Missing Cash Flow Metric #9	.	.	.	.	.	.	.	.
Cash Flow Metric #10	--	.	.	1.05	0.26	1.04	0.34	1.07	0.13
Missing Cash Flow Metric #10	Not Missing Cash Flow Metric #10	.	.	.	.	.	.	.	.
Cash Flow Metric #11	--	.	.	1.00	0.97	1.00	0.97	0.94	0.64
Missing Cash Flow Metric #11	Not Missing Cash Flow Metric #11	.	.	1.17	0.82	1.06	0.94	1.21	0.82

Cash Flow Metric #13	--	.	.	0.79	0.69	3.58	0.29	4.28	0.24
Missing Cash Flow Metric #13	Not Missing Cash Flow Metric #13	.	.	1.46	0.59	1.13	0.89	0.94	0.94
Cash Flow Metric #14	--	.	.	1.36	0.59	5.75	0.14	4.67	0.22
Missing Cash Flow Metric #14	Not Missing Cash Flow Metric #14	.	.	1.89	0.38	3.90	0.27	3.53	0.28
Cash Flow Metric #16	--	.	.	1.02	0.98	4.74	0.18	4.54	0.22
Missing Cash Flow Metric #16	Not Missing Cash Flow Metric #16	.	.	3,764,226.46	0.00	346,558,531.41	0.00	80,276,220.57	0.00
Cash Flow Metric #17	--	.	.	1.56	0.33	6.96	0.08	5.63	0.14
Missing Cash Flow Metric #17	Not Missing Cash Flow Metric #17	.	.	0.48	0.41	0.11	0.11	0.12	0.15
Cash Flow Metric #18	--	.	.	1.23	0.69	5.35	0.14	3.65	0.29
Missing Cash Flow Metric #18	Not Missing Cash Flow Metric #18	.	.	1.18	0.80	0.98	0.98	0.81	0.78
Cash Flow Metric #20	--	.	.	0.94	0.16	1.30	0.40	1.40	0.23
Missing Cash Flow Metric #20	Not Missing Cash Flow Metric #20	.	.	0.25	0.01	0.23	0.02	0.24	0.01
Cash Flow Metric #21	--	.	.	0.92	0.03	1.26	0.44	1.33	0.30
Missing Cash Flow Metric #21	Not Missing Cash Flow Metric #21	.	.	0.72	0.70	0.47	0.46	0.44	0.39
Cash Flow Metric #22	--	.	.	0.52	0.14	0.12	0.06	0.16	0.11
Missing Cash Flow Metric #22	Not Missing Cash Flow Metric #22	.	.	.	.	.	.	.	.
Cash Flow Metric #23	--	.	.	1.01	0.38	0.73	0.30	0.68	0.16
Missing Cash Flow Metric #23	Not Missing Cash Flow Metric #23	.	.	.	.	.	.	.	.
Cash Flow Metric #24	--	.	.	3.64	0.01	15.97	0.02	12.47	0.04
Missing Cash Flow Metric #24	Not Missing Cash Flow Metric #24	.	.	.	.	.	.	.	.
Cash Flow Metric #25	--	.	.	2.08	0.11	9.12	0.05	7.55	0.09
Missing Cash Flow Metric #25	Not Missing Cash Flow Metric #25	.	.	0.12	0.00	0.08	0.00	0.10	0.00
Cash Flow Metric #19	--	.	.	0.48	0.30	0.57	0.38	0.63	0.54
Missing Cash Flow Metric #19	Not Missing Cash Flow Metric #19	.	.	2.23	0.00	2.32	0.00	1.44	0.14
Missing All Cash flow Metrics	Not Missing Any Cash flow Metrics	.	.	.	.	.	.	.	.
Cash Flow Metric #7	--	.	.	.	.	1.43	0.24	1.52	0.13
Missing Cash Flow Metric #7	Not Missing Cash Flow Metric #7	.	.	.	.	1.55	0.35	1.10	0.82
Cash Flow Metric #12	--	.	.	.	.	6.59	0.11	4.97	0.20
Missing Cash Flow Metric #12	Not Missing Cash Flow Metric #12	.	.	.	.	0.24	0.29	0.31	0.35
Cash Flow Metric #15	--	.	.	.	.	1.37	0.30	1.48	0.16

Missing Cash Flow Metric #15	Not Missing Cash Flow Metric #15	.	.	.	.	0.47	0.45	0.48	0.41
Cash Flow Data Quality Bucket A	Cash Flow Data Quality Bucket C	.	.	.	.	1.49	0.00	1.43	0.00
Cash Flow Data Quality Bucket B		.	.	.	.	1.03	0.97	1.18	0.85
Constant	--	11.59	0.00	0.18	0.00	0.13	0.00	3.99	0.13
Pseudo R Squared		0.090		0.055		0.064		0.128	
AUC		0.720		0.675		0.688		0.758	
Sample Size		3,776		3,776		3,776		3,776	



Appendix F. Participant #6
Chart 1. Receiver Operating Characteristic (ROC) Curves for Models 1-4

Page **124** of **161**

Appendix F. Participant #6

Table 6. Delinquency Frequency by FICO Score Percentile and Model 2's Predicted Probability of Delinquency Percentile[72]

FICO Score	Model 2's Predicted Probability of Delinquency: 95 - 100th	90 - 95th	85 - 90th	80 - 85th	75 - 80th	70 - 75th	65 - 70th	60 - 65th	55 - 60th	50 - 55th	45 - 50th	40 - 45th	35 - 40th	30 - 35th	25 - 30th	20 - 25th	15 - 20th	10 - 15th	5 - 10th	0 - 5th
0 - 5th	41.7	22.7	33.3	38.5	37.5	33.3	44.4	.	30.0	27.3	50.0	20.0	.	.	25.0	0.0	0.0	.	.	.
5 - 10th	25.0	52.9	77.8	27.3	12.5	11.1	22.2	25.0	45.5	33.3	33.3	16.7	22.2	0.0	0.0	33.3	.	12.5	20.0	.
10 - 15th	36.4	22.2	55.6	30.8	31.3	23.1	0.0	45.5	33.3	27.3	25.0	0.0	40.0	25.0	16.7	20.0	.	.	.	0.0
15 - 20th	44.4	28.6	33.3	36.4	27.3	30.0	27.3	16.7	27.3	37.5	10.0	8.3	14.3	12.5	20.0	20.0	33.3	.	.	0.0
20 - 25th	35.7	16.7	63.6	42.9	23.1	9.1	0.0	20.0	22.2	33.3	30.0	37.5	20.0	11.1	0.0	0.0	11.1	.	14.3	.
25 - 30th	50.0	.	8.3	12.5	40.0	23.1	15.4	8.3	0.0	28.6	20.0	16.7	11.1	0.0	0.0	20.0	0.0	0.0	11.1	20.0
30 - 35th	13.3	15.4	25.0	30.0	7.1	0.0	27.3	9.1	20.0	.	0.0	11.1	9.1	20.0	.	9.1	20.0	0.0	.	.
35 - 40th	42.9	36.4	42.9	25.0	40.0	40.0	0.0	9.1	14.3	11.1	16.7	0.0	22.2	25.0	16.7	0.0	0.0	.	.	0.0
40 - 45th	.	20.0	33.3	21.4	37.5	66.7	33.3	27.3	8.3	18.2	0.0	0.0	18.2	0.0	16.7	0.0	0.0	0.0	.	14.3
45 - 50th	14.3	.	0.0	18.2	25.0	14.3	20.0	0.0	10.0	25.0	0.0	16.7	12.5	14.3	20.0	16.7	12.5	0.0	0.0	0.0
50 - 55th	25.0	0.0	0.0	33.3	10.0	8.3	25.0	0.0	14.3	.	10.0	0.0	0.0	28.6	12.5	0.0	12.5	0.0	10.0	0.0
55 - 60th	25.0	20.0	0.0	14.3	8.3	0.0	0.0	.	9.1	20.0	9.1	0.0	0.0	0.0	16.7	0.0	16.7	13.3	0.0	.
60 - 65th	.	.	0.0	33.3	16.7	0.0	7.7	27.3	0.0	0.0	0.0	0.0	0.0	12.5	20.0	.	0.0	0.0	0.0	0.0
65 - 70th	20.0	40.0	.	33.3	42.9	0.0	0.0	25.0	0.0	0.0	0.0	18.2	12.5	15.4	0.0	0.0	0.0	0.0	0.0	12.5
70 - 75th	0.0	0.0	.	0.0	0.0	22.2	0.0	0.0	12.5	11.1	20.0	0.0	0.0	0.0	7.1	0.0	0.0	0.0	0.0	7.1
75 - 80th	.	28.6	0.0	.	.	16.7	.	0.0	12.5	9.1	0.0	0.0	0.0	10.0	12.5	0.0	8.3	0.0	0.0	0.0
80 - 85th	.	14.3	12.5	14.3	0.0	14.3	.	10.0	0.0	7.7	0.0	0.0	0.0	18.2	0.0	15.4	16.7	0.0	8.3	0.0
85 - 90th	.	.	.	.	.	11.1	10.0	11.1	0.0	0.0	0.0	30.0	0.0	0.0	5.6	0.0	0.0	18.8	7.7	0.0
90 - 95th	.	0.0	.	0.0	0.0	0.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0	9.1	0.0	0.0	0.0	7.1	0.0	0.0
95 - 100th	.	.	0.0	.	0.0	.	0.0	11.1	0.0	14.3	0.0	0.0	0.0	0.0	0.0	0.0	6.3	0.0	0.0	5.3

72 Cells are shaded based on values. Green indicates values close to the lowest delinquent frequency, yellow indicates values close to the median delinquent frequency, and red indicates values close to the highest delinquent frequency. Gray values indicate cells where there were fewer than 5 loans. Percentiles are based on the population of originated loans. 283 originated loans with a missing FICO score were excluded from the frequency table.

Appendix F. Participant #6 Table 7. Summary of Whether Applicant's Zip Code Population is at least 50% Minority, by Delinquency Status[73]									
	Delinquent			Not Delinquent			All		
Value	#	Row %	Col %	#	Row %	Col %	#	%	P-Val
Missing	6	14.6%	1.2%	35	85.4%	1.1%	41	1.1%	0.819
False	221	12.2%	42.7%	1,593	87.8%	48.9%	1,814	48.0%	0.010
True	290	15.1%	56.1%	1,631	84.9%	50.0%	1,921	50.9%	0.012
All	517	13.7%	100.0%	3,259	86.3%	100.0%	3,776	100.0%	.

Appendix F. Participant #6 Table 8. Summary of Whether Applicant's Zip Code Population is at least 80% Minority, by Delinquency Status									
	Delinquent			Not Delinquent			All		
Value	#	Row %	Col %	#	Row %	Col %	#	%	P-Val
Missing	6	14.6%	1.2%	35	85.4%	1.1%	41	1.1%	0.819
False	322	12.2%	62.3%	2,319	87.8%	71.2%	2,641	69.9%	0.000
True	189	17.3%	36.6%	905	82.7%	27.8%	1,094	29.0%	0.000
All	517	13.7%	100.0%	3,259	86.3%	100.0%	3,776	100.0%	.

Appendix F. Participant #6 Table 9. Summary of Whether Applicant's Income Exceeds Zip Code's Median Income, by Delinquency Status									
	Delinquent			Not Delinquent			All		
Value	#	Row %	Col %	#	Row %	Col %	#	%	P-Val
Missing	17	13.5%	3.3%	109	86.5%	3.3%	126	3.3%	1.000
False	330	14.7%	63.8%	1,911	85.3%	58.6%	2,241	59.3%	0.027
True	170	12.1%	32.9%	1,239	87.9%	38.0%	1,409	37.3%	0.028
All	517	13.7%	100.0%	3,259	86.3%	100.0%	3,776	100.0%	.

73 Missing demographic data is the result of invalid zip codes, zip codes outside of the 50 States, or zip codes that do not have an associated ZCTA (Zip Code Tabulation Area).

Appendix F. Participant #6

Table 10. Summary of Actions Taken[74]

	All Applications	Approved Applications		Declined Applications		Progress Applications		Withdrawn Applications		Originated Loans		Delinquent Loans	
	Count	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
All	13,431	3,994	29.74%	1,566	11.66%	586	4.36%	7,285	54.24%	3,776	28.11%	517	13.69%

74 The percentages in the "Delinquent Loans" column are calculated out of originated loans.

Appendix F. Participant #6

Table 11. Difference of Means Tests Within Demographic Group: Originated Loans[75]

Variable	Demographic Group	Status	Count	Mean	T-Stat	P-Value
Date Difference #1	Originated Loans	Delinquent	517	30.6	.	.
		Not Delinquent	3,259	28.8	.	.
		All	3,776	29.0	-1.12	0.262
	African American 75%	Delinquent	131	32.5	.	.
		Not Delinquent	397	29.7	-0.81	0.416
	Hispanic 75%	Delinquent	46	32.4	.	.
		Not Delinquent	339	28.2	-0.62	0.538
	Asian 75%	Delinquent	2	53.5	.	.
		Not Delinquent	55	38.0	-0.51	0.692
	Non-Hispanic White 75%	Delinquent	95	29.6	.	.
		Not Delinquent	637	31.1	0.46	0.646
	Other or Missing BISG	Delinquent	243	29.5	.	.
		Not Delinquent	1,831	27.6	-0.78	0.434
	Female	Delinquent	178	33.4	.	.
		Not Delinquent	1,053	29.5	-1.31	0.193
	Male	Delinquent	214	31.5	.	.
		Not Delinquent	1,446	30.6	-0.31	0.756
	Gender Unassigned	Delinquent	125	25.3	.	.
		Not Delinquent	760	24.4	-0.35	0.728
FICO score	Originated Loans	Delinquent	466	600.1	.	.
		Not Delinquent	3,027	646.4	.	.
		All	3,493	640.2	14.95	0.000
	African American 75%	Delinquent	115	587.0	.	.
		Not Delinquent	364	626.0	5.93	0.000
	Hispanic 75%	Delinquent	33	587.2	.	.
		Not Delinquent	299	647.6	4.55	0.000
	Asian 75%	Delinquent	2	634.0	.	.
		Not Delinquent	53	664.7	2.65	0.033
	Non-Hispanic White 75%	Delinquent	87	603.8	.	.
		Not Delinquent	616	659.4	7.76	0.000

[75] T-tests assume unequal variances and are conducted on the delinquent and non-delinquent populations. Yellow highlighting indicates a difference between the delinquent and non-delinquent groups that is statistically significant at the 95% confidence level (P-value < 0.05). Highlighting is shown regardless of the direction of the difference. Counts displayed are the counts of non-missing values for each variable, by demographic group and status.

	Other or Missing BISG	Delinquent	229	606.8	.	.
		Not Delinquent	1,695	645.2	9.03	0.000
	Female	Delinquent	169	596.3	.	.
		Not Delinquent	984	643.5	9.30	0.000
	Male	Delinquent	185	604.8	.	.
		Not Delinquent	1,354	647.8	8.47	0.000
	Gender Unassigned	Delinquent	112	598.1	.	.
		Not Delinquent	689	647.7	8.14	0.000
BK score	Originated Loans	Delinquent	483	293.8	.	.
		Not Delinquent	3,079	424.4	.	.
		All	3,562	406.7	13.35	0.000
	African American 75%	Delinquent	121	276.5	.	.
		Not Delinquent	379	371.2	4.65	0.000
	Hispanic 75%	Delinquent	35	284.6	.	.
		Not Delinquent	303	419.7	3.86	0.000
	Asian 75%	Delinquent	2	314.0	.	.
		Not Delinquent	53	458.5	3.61	0.005
	Non-Hispanic White 75%	Delinquent	90	291.9	.	.
		Not Delinquent	621	476.8	8.15	0.000
	Other or Missing BISG	Delinquent	235	304.6	.	.
		Not Delinquent	1,723	416.9	7.88	0.000
	Female	Delinquent	169	281.2	.	.
		Not Delinquent	1,000	417.4	8.20	0.000
	Male	Delinquent	195	298.6	.	.
		Not Delinquent	1,381	426.5	8.49	0.000
	Gender Unassigned	Delinquent	119	303.7	.	.
		Not Delinquent	698	430.2	6.18	0.000
# of open accounts on credit report	Originated Loans	Delinquent	16	0.8	.	.
		Not Delinquent	394	0.9	.	.
		All	410	0.8	0.12	0.905
	African American 75%	Delinquent	5	0.4	.	.
		Not Delinquent	75	0.9	1.86	0.103
	Hispanic 75%	Delinquent	0	.	.	.
		Not Delinquent	54	0.6	.	.
	Asian 75%	Delinquent	0	.	.	.
		Not Delinquent	6	0.3	.	.
	Non-Hispanic White 75%	Delinquent	4	1.0	.	.
		Not Delinquent	85	0.9	-0.07	0.948
	Other or Missing BISG	Delinquent	7	1.0	.	.
		Not Delinquent	174	0.9	-0.30	0.777
	Female	Delinquent	3	1.7	.	.

		Not Delinquent	127	0.9	-1.10	0.380
	Male	Delinquent	7	0.1	.	.
		Not Delinquent	170	0.8	3.98	0.002
	Gender Unassigned	Delinquent	6	1.2	.	.
		Not Delinquent	97	0.8	-0.56	0.598
$ amount of unpaid balances on credit report	Originated Loans	Delinquent	494	$3,757	.	.
		Not Delinquent	3,130	$11,054	.	.
		All	3,624	$10,059	3.91	0.000
	African American 75%	Delinquent	125	$3,999	.	.
		Not Delinquent	388	$13,926	2.99	0.003
	Hispanic 75%	Delinquent	38	$0	.	.
		Not Delinquent	314	$14,512	4.34	0.000
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$16,408	1.87	0.067
	Non-Hispanic White 75%	Delinquent	91	$3,623	.	.
		Not Delinquent	627	$13,891	2.40	0.017
	Other or Missing BISG	Delinquent	238	$4,311	.	.
		Not Delinquent	1,747	$8,610	1.41	0.160
	Female	Delinquent	172	$5,157	.	.
		Not Delinquent	1,012	$12,844	1.79	0.075
	Male	Delinquent	203	$1,874	.	.
		Not Delinquent	1,408	$10,619	4.53	0.000
	Gender Unassigned	Delinquent	119	$4,943	.	.
		Not Delinquent	710	$9,364	1.37	0.173
$ amount of monthly payments on credit report	Originated Loans	Delinquent	494	$38	.	.
		Not Delinquent	3,130	$150	.	.
		All	3,624	$135	4.60	0.000
	African American 75%	Delinquent	125	$28	.	.
		Not Delinquent	388	$157	4.18	0.000
	Hispanic 75%	Delinquent	38	$0	.	.
		Not Delinquent	314	$308	2.66	0.008
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$136	2.14	0.037
	Non-Hispanic White 75%	Delinquent	91	$42	.	.
		Not Delinquent	627	$205	2.87	0.004
	Other or Missing BISG	Delinquent	238	$48	.	.
		Not Delinquent	1,747	$101	1.49	0.137
	Female	Delinquent	172	$60	.	.
		Not Delinquent	1,012	$140	1.59	0.113
	Male	Delinquent	203	$14	.	.
		Not Delinquent	1,408	$155	4.57	0.000

	Gender Unassigned	Delinquent	119	$46	.	.
		Not Delinquent	710	$156	2.42	0.016
$ Credit limit of revolving accounts on credit report	Originated Loans	Delinquent	10	$75,529	.	.
		Not Delinquent	323	$81,624	.	.
		All	333	$81,441	0.15	0.881
	African American 75%	Delinquent	3	$16,307	.	.
		Not Delinquent	54	$23,812	0.50	0.654
	Hispanic 75%	Delinquent	0	.	.	.
		Not Delinquent	47	$43,791	.	.
	Asian 75%	Delinquent	0	.	.	.
		Not Delinquent	4	$61,881	.	.
	Non-Hispanic White 75%	Delinquent	3	$68,023	.	.
		Not Delinquent	76	$152,017	0.98	0.364
	Other or Missing BISG	Delinquent	4	$125,574	.	.
		Not Delinquent	142	$79,012	-0.58	0.598
	Female	Delinquent	3	$171,725	.	.
		Not Delinquent	103	$95,365	-0.81	0.490
	Male	Delinquent	3	$1,820	.	.
		Not Delinquent	140	$93,427	3.15	0.002
	Gender Unassigned	Delinquent	4	$58,663	.	.
		Not Delinquent	80	$43,278	-0.31	0.770
$ unpaid balances of revolving accounts on credit report	Originated Loans	Delinquent	16	$7,932	.	.
		Not Delinquent	394	$8,705	.	.
		All	410	$8,675	0.16	0.871
	African American 75%	Delinquent	5	$1,057	.	.
		Not Delinquent	75	$5,165	3.35	0.001
	Hispanic 75%	Delinquent	0	.	.	.
		Not Delinquent	54	$7,194	.	.
	Asian 75%	Delinquent	0	.	.	.
		Not Delinquent	6	$13,263	.	.
	Non-Hispanic White 75%	Delinquent	4	$4,538	.	.
		Not Delinquent	85	$12,246	1.53	0.184
	Other or Missing BISG	Delinquent	7	$14,781	.	.
		Not Delinquent	174	$8,814	-0.59	0.578
	Female	Delinquent	3	$27,964	.	.
		Not Delinquent	127	$11,391	-0.74	0.535
	Male	Delinquent	7	$348	.	.
		Not Delinquent	170	$8,700	6.65	0.000
	Gender Unassigned	Delinquent	6	$6,763	.	.
		Not Delinquent	97	$5,199	-0.37	0.725
	Originated Loans	Delinquent	10	36.20%	.	.

% utilization of revolving accounts on credit report		Not Delinquent	323	49.22%	.	.
		All	333	48.83%	1.23	0.247
	African American 75%	Delinquent	3	48.00%	.	.
		Not Delinquent	54	52.22%	0.16	0.886
	Hispanic 75%	Delinquent	0	.	.	.
		Not Delinquent	47	43.77%	.	.
	Asian 75%	Delinquent	0	.	.	.
		Not Delinquent	4	45.50%	.	.
	Non-Hispanic White 75%	Delinquent	3	15.67%	.	.
		Not Delinquent	76	50.37%	3.51	0.050
	Other or Missing BISG	Delinquent	4	42.75%	.	.
		Not Delinquent	142	49.38%	0.39	0.722
	Female	Delinquent	3	11.33%	.	.
		Not Delinquent	103	47.05%	6.38	0.005
	Male	Delinquent	3	49.33%	.	.
		Not Delinquent	140	48.53%	-0.07	0.950
	Gender Unassigned	Delinquent	4	45.00%	.	.
		Not Delinquent	80	53.24%	0.36	0.743
Cash Flow Metric #1	Originated Loans	Delinquent	499	$510	.	.
		Not Delinquent	3,155	$864	.	.
		All	3,654	$816	6.75	0.000
	African American 75%	Delinquent	124	$281	.	.
		Not Delinquent	383	$483	3.04	0.003
	Hispanic 75%	Delinquent	45	$801	.	.
		Not Delinquent	334	$859	0.37	0.713
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$1,789	5.79	0.000
	Non-Hispanic White 75%	Delinquent	92	$564	.	.
		Not Delinquent	619	$990	3.34	0.001
	Other or Missing BISG	Delinquent	236	$558	.	.
		Not Delinquent	1,764	$875	3.88	0.000
	Female	Delinquent	172	$566	.	.
		Not Delinquent	1,029	$980	4.24	0.000
	Male	Delinquent	211	$457	.	.
		Not Delinquent	1,412	$830	4.62	0.000
	Gender Unassigned	Delinquent	116	$522	.	.
		Not Delinquent	714	$766	2.58	0.010
Cash Flow Metric #2	Originated Loans	Delinquent	496	$1,738	.	.
		Not Delinquent	3,144	$4,343	.	.
		All	3,640	$3,988	5.20	0.000
	African American 75%	Delinquent	121	$970	.	.

		Not Delinquent	381	$3,111	1.54	0.125
	Hispanic 75%	Delinquent	45	$2,909	.	.
		Not Delinquent	330	$3,769	0.70	0.487
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$6,596	3.88	0.000
	Non-Hispanic White 75%	Delinquent	92	$1,054	.	.
		Not Delinquent	623	$4,891	5.37	0.000
	Other or Missing BISG	Delinquent	236	$2,190	.	.
		Not Delinquent	1,755	$4,453	2.85	0.005
	Female	Delinquent	172	$800	.	.
		Not Delinquent	1,024	$3,046	4.26	0.000
	Male	Delinquent	211	$2,301	.	.
		Not Delinquent	1,409	$5,402	3.23	0.001
	Gender Unassigned	Delinquent	113	$2,113	.	.
		Not Delinquent	711	$4,113	2.29	0.023
Cash Flow Metric #3	Originated Loans	Delinquent	498	$101	.	.
		Not Delinquent	3,148	$182	.	.
		All	3,646	$171	2.48	0.013
	African American 75%	Delinquent	124	$79	.	.
		Not Delinquent	381	$140	1.21	0.229
	Hispanic 75%	Delinquent	46	$25	.	.
		Not Delinquent	330	$159	3.70	0.000
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$328	3.10	0.003
	Non-Hispanic White 75%	Delinquent	91	$44	.	.
		Not Delinquent	620	$228	4.67	0.000
	Other or Missing BISG	Delinquent	235	$151	.	.
		Not Delinquent	1,763	$174	0.37	0.708
	Female	Delinquent	172	$116	.	.
		Not Delinquent	1,028	$127	0.15	0.881
	Male	Delinquent	212	$79	.	.
		Not Delinquent	1,412	$226	3.88	0.000
	Gender Unassigned	Delinquent	114	$119	.	.
		Not Delinquent	708	$172	0.94	0.350
Cash Flow Metric #4	Originated Loans	Delinquent	506	$9,246	.	.
		Not Delinquent	3,185	$17,161	.	.
		All	3,691	$16,076	6.30	0.000
	African American 75%	Delinquent	126	$8,003	.	.
		Not Delinquent	392	$13,176	1.41	0.159
	Hispanic 75%	Delinquent	46	$9,377	.	.
		Not Delinquent	338	$15,448	2.72	0.007

	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$23,443	6.80	0.000
	Non-Hispanic White 75%	Delinquent	94	$7,586	.	.
		Not Delinquent	627	$20,021	6.53	0.000
	Other or Missing BISG	Delinquent	238	$10,612	.	.
		Not Delinquent	1,773	$17,163	3.66	0.000
	Female	Delinquent	172	$7,495	.	.
		Not Delinquent	1,035	$13,576	4.05	0.000
	Male	Delinquent	212	$9,661	.	.
		Not Delinquent	1,424	$21,164	5.52	0.000
	Gender Unassigned	Delinquent	122	$10,994	.	.
		Not Delinquent	726	$14,420	1.18	0.240
Cash Flow Metric #5	Originated Loans	Delinquent	500	$575	.	.
		Not Delinquent	3,166	$1,152	.	.
		All	3,666	$1,073	3.80	0.000
	African American 75%	Delinquent	125	$411	.	.
		Not Delinquent	386	$583	1.57	0.118
	Hispanic 75%	Delinquent	46	$553	.	.
		Not Delinquent	332	$1,113	2.16	0.032
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$1,471	1.87	0.067
	Non-Hispanic White 75%	Delinquent	92	$562	.	.
		Not Delinquent	624	$1,168	2.59	0.011
	Other or Missing BISG	Delinquent	235	$676	.	.
		Not Delinquent	1,769	$1,268	2.25	0.024
	Female	Delinquent	172	$442	.	.
		Not Delinquent	1,031	$775	2.24	0.026
	Male	Delinquent	212	$642	.	.
		Not Delinquent	1,417	$1,454	2.62	0.009
	Gender Unassigned	Delinquent	116	$647	.	.
		Not Delinquent	718	$1,097	2.43	0.015
Cash Flow Metric #6	Originated Loans	Delinquent	492	$227	.	.
		Not Delinquent	3,110	$399	.	.
		All	3,602	$376	2.27	0.023
	African American 75%	Delinquent	121	$420	.	.
		Not Delinquent	380	$389	-0.16	0.869
	Hispanic 75%	Delinquent	45	$147	.	.
		Not Delinquent	320	$238	0.86	0.393
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$693	1.42	0.160

	Non-Hispanic White 75%	Delinquent	93	$227	.	.
		Not Delinquent	608	$708	1.78	0.075
	Other or Missing BISG	Delinquent	231	$144	.	.
		Not Delinquent	1,747	$314	2.82	0.005
	Female	Delinquent	171	$204	.	.
		Not Delinquent	1,009	$331	1.37	0.173
	Male	Delinquent	210	$211	.	.
		Not Delinquent	1,404	$544	2.47	0.014
	Gender Unassigned	Delinquent	111	$293	.	.
		Not Delinquent	697	$208	-0.58	0.563
Cash Flow Metric #7	Originated Loans	Delinquent	492	$1,711	.	.
		Not Delinquent	3,155	$1,870	.	.
		All	3,647	$1,849	1.30	0.193
	African American 75%	Delinquent	118	$1,160	.	.
		Not Delinquent	384	$1,671	1.79	0.073
	Hispanic 75%	Delinquent	46	$1,540	.	.
		Not Delinquent	333	$1,840	1.22	0.226
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$2,510	5.80	0.000
	Non-Hispanic White 75%	Delinquent	93	$1,786	.	.
		Not Delinquent	625	$2,011	0.85	0.395
	Other or Missing BISG	Delinquent	233	$2,009	.	.
		Not Delinquent	1,758	$1,849	-0.80	0.422
	Female	Delinquent	165	$1,259	.	.
		Not Delinquent	1,018	$1,370	0.72	0.473
	Male	Delinquent	206	$1,682	.	.
		Not Delinquent	1,416	$2,108	2.40	0.017
	Gender Unassigned	Delinquent	121	$2,377	.	.
		Not Delinquent	721	$2,108	-0.83	0.405
Cash Flow Metric #8	Originated Loans	Delinquent	497	$1,685	.	.
		Not Delinquent	3,148	$3,260	.	.
		All	3,645	$3,045	4.30	0.000
	African American 75%	Delinquent	124	$1,263	.	.
		Not Delinquent	381	$1,576	0.60	0.550
	Hispanic 75%	Delinquent	46	$1,233	.	.
		Not Delinquent	330	$3,194	3.13	0.002
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$4,264	4.42	0.000
	Non-Hispanic White 75%	Delinquent	92	$1,252	.	.
		Not Delinquent	619	$4,921	3.78	0.000
	Other or Missing BISG	Delinquent	233	$2,185	.	.

		Not Delinquent	1,764	$3,022	1.49	0.137
	Female	Delinquent	172	$1,819	.	.
		Not Delinquent	1,028	$3,104	1.73	0.084
	Male	Delinquent	212	$1,467	.	.
		Not Delinquent	1,413	$3,795	4.45	0.000
	Gender Unassigned	Delinquent	113	$1,891	.	.
		Not Delinquent	707	$2,417	0.90	0.371
Cash Flow Metric #9	Originated Loans	Delinquent	506	$9,026	.	.
		Not Delinquent	3,185	$16,771	.	.
		All	3,691	$15,709	6.21	0.000
	African American 75%	Delinquent	126	$7,600	.	.
		Not Delinquent	392	$12,799	1.42	0.156
	Hispanic 75%	Delinquent	46	$9,234	.	.
		Not Delinquent	338	$15,222	2.68	0.008
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$22,750	6.68	0.000
	Non-Hispanic White 75%	Delinquent	94	$7,361	.	.
		Not Delinquent	627	$19,334	6.71	0.000
	Other or Missing BISG	Delinquent	238	$10,473	.	.
		Not Delinquent	1,773	$16,853	3.57	0.000
	Female	Delinquent	172	$7,292	.	.
		Not Delinquent	1,035	$13,253	3.98	0.000
	Male	Delinquent	212	$9,452	.	.
		Not Delinquent	1,424	$20,628	5.42	0.000
	Gender Unassigned	Delinquent	122	$10,727	.	.
		Not Delinquent	726	$14,221	1.21	0.229
Cash Flow Metric #10	Originated Loans	Delinquent	506	$4,800	.	.
		Not Delinquent	3,185	$10,090	.	.
		All	3,691	$9,365	6.88	0.000
	African American 75%	Delinquent	126	$3,055	.	.
		Not Delinquent	392	$5,992	1.79	0.075
	Hispanic 75%	Delinquent	46	$5,780	.	.
		Not Delinquent	338	$9,324	2.02	0.046
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$15,300	5.57	0.000
	Non-Hispanic White 75%	Delinquent	94	$3,791	.	.
		Not Delinquent	627	$12,439	6.13	0.000
	Other or Missing BISG	Delinquent	238	$5,973	.	.
		Not Delinquent	1,773	$10,150	3.43	0.001
	Female	Delinquent	172	$3,982	.	.
		Not Delinquent	1,035	$8,278	3.87	0.000

	Male	Delinquent	212	$5,179	.	.
		Not Delinquent	1,424	$12,076	5.03	0.000
	Gender Unassigned	Delinquent	122	$5,296	.	.
		Not Delinquent	726	$8,778	2.74	0.007
Cash Flow Metric #11	Originated Loans	Delinquent	499	$164	.	.
		Not Delinquent	3,152	$240	.	.
		All	3,651	$230	3.59	0.000
	African American 75%	Delinquent	124	$96	.	.
		Not Delinquent	380	$159	2.39	0.018
	Hispanic 75%	Delinquent	46	$236	.	.
		Not Delinquent	330	$292	0.78	0.439
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	55	$484	3.80	0.000
	Non-Hispanic White 75%	Delinquent	92	$208	.	.
		Not Delinquent	620	$253	0.76	0.447
	Other or Missing BISG	Delinquent	235	$170	.	.
		Not Delinquent	1,767	$236	2.08	0.039
	Female	Delinquent	172	$137	.	.
		Not Delinquent	1,029	$244	3.90	0.000
	Male	Delinquent	212	$167	.	.
		Not Delinquent	1,411	$249	2.56	0.011
	Gender Unassigned	Delinquent	115	$198	.	.
		Not Delinquent	712	$218	0.34	0.733
Cash Flow Metric #12	Originated Loans	Delinquent	494	$14	.	.
		Not Delinquent	3,108	$24	.	.
		All	3,602	$22	1.83	0.067
	African American 75%	Delinquent	120	$16	.	.
		Not Delinquent	379	$19	0.22	0.828
	Hispanic 75%	Delinquent	45	$10	.	.
		Not Delinquent	320	$20	0.80	0.426
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$11	1.00	0.322
	Non-Hispanic White 75%	Delinquent	92	$30	.	.
		Not Delinquent	607	$21	-0.65	0.518
	Other or Missing BISG	Delinquent	235	$7	.	.
		Not Delinquent	1,748	$27	2.83	0.005
	Female	Delinquent	172	$1	.	.
		Not Delinquent	1,012	$7	2.31	0.021
	Male	Delinquent	211	$29	.	.
		Not Delinquent	1,403	$32	0.33	0.739

	Gender Unassigned	Delinquent	111	$6	.	.
		Not Delinquent	693	$31	2.13	0.034
Cash Flow Metric #13	Originated Loans	Delinquent	501	$127	.	.
		Not Delinquent	3,158	$240	.	.
		All	3,659	$224	7.45	0.000
	African American 75%	Delinquent	125	$75	.	.
		Not Delinquent	387	$168	5.56	0.000
	Hispanic 75%	Delinquent	45	$70	.	.
		Not Delinquent	329	$200	5.15	0.000
	Asian 75%	Delinquent	2	$120	.	.
		Not Delinquent	55	$366	2.50	0.115
	Non-Hispanic White 75%	Delinquent	94	$149	.	.
		Not Delinquent	619	$305	4.05	0.000
	Other or Missing BISG	Delinquent	235	$157	.	.
		Not Delinquent	1,768	$236	3.07	0.002
	Female	Delinquent	172	$154	.	.
		Not Delinquent	1,028	$261	3.42	0.001
	Male	Delinquent	211	$125	.	.
		Not Delinquent	1,419	$228	4.81	0.000
	Gender Unassigned	Delinquent	118	$92	.	.
		Not Delinquent	711	$233	5.83	0.000
Cash Flow Metric #14	Originated Loans	Delinquent	494	$33	.	.
		Not Delinquent	3,125	$72	.	.
		All	3,619	$67	4.53	0.000
	African American 75%	Delinquent	120	$28	.	.
		Not Delinquent	381	$60	2.11	0.035
	Hispanic 75%	Delinquent	45	$21	.	.
		Not Delinquent	324	$51	2.07	0.040
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$93	2.14	0.037
	Non-Hispanic White 75%	Delinquent	92	$48	.	.
		Not Delinquent	613	$81	1.22	0.226
	Other or Missing BISG	Delinquent	235	$33	.	.
		Not Delinquent	1,753	$75	3.58	0.000
	Female	Delinquent	172	$27	.	.
		Not Delinquent	1,017	$78	4.38	0.000
	Male	Delinquent	211	$33	.	.
		Not Delinquent	1,407	$70	2.82	0.005
	Gender Unassigned	Delinquent	111	$43	.	.
		Not Delinquent	701	$69	1.16	0.249
	Originated Loans	Delinquent	498	$1,643	.	.

Cash Flow Metric #15		Not Delinquent	3,135	$1,804	.	.
		All	3,633	$1,782	0.79	0.430
	African American 75%	Delinquent	123	$2,086	.	.
		Not Delinquent	383	$2,628	0.85	0.394
	Hispanic 75%	Delinquent	45	$925	.	.
		Not Delinquent	327	$1,474	2.37	0.020
	Asian 75%	Delinquent	2	$3,950	.	.
		Not Delinquent	55	$1,529	-4.58	0.049
	Non-Hispanic White 75%	Delinquent	92	$1,277	.	.
		Not Delinquent	612	$1,973	2.70	0.007
	Other or Missing BISG	Delinquent	236	$1,671	.	.
		Not Delinquent	1,758	$1,636	-0.12	0.904
	Female	Delinquent	172	$1,328	.	.
		Not Delinquent	1,021	$1,846	2.82	0.005
	Male	Delinquent	211	$1,748	.	.
		Not Delinquent	1,409	$1,903	0.45	0.654
	Gender Unassigned	Delinquent	115	$1,919	.	.
		Not Delinquent	705	$1,545	-0.67	0.502
Cash Flow Metric #16	Originated Loans	Delinquent	506	$279	.	.
		Not Delinquent	3,186	$357	.	.
		All	3,692	$346	6.88	0.000
	African American 75%	Delinquent	126	$242	.	.
		Not Delinquent	392	$278	2.00	0.046
	Hispanic 75%	Delinquent	46	$296	.	.
		Not Delinquent	338	$337	1.26	0.211
	Asian 75%	Delinquent	2	$200	.	.
		Not Delinquent	55	$370	1.59	0.313
	Non-Hispanic White 75%	Delinquent	94	$279	.	.
		Not Delinquent	627	$428	5.96	0.000
	Other or Missing BISG	Delinquent	238	$295	.	.
		Not Delinquent	1,774	$352	3.08	0.002
	Female	Delinquent	172	$268	.	.
		Not Delinquent	1,035	$337	3.63	0.000
	Male	Delinquent	212	$285	.	.
		Not Delinquent	1,425	$373	5.11	0.000
	Gender Unassigned	Delinquent	122	$283	.	.
		Not Delinquent	726	$351	2.87	0.005
Cash Flow Metric #17	Originated Loans	Delinquent	502	$788	.	.
		Not Delinquent	3,167	$1,090	.	.
		All	3,669	$1,049	7.43	0.000
	African American 75%	Delinquent	124	$671	.	.

		Not Delinquent	387	$864	2.51	0.013
	Hispanic 75%	Delinquent	46	$757	.	.
		Not Delinquent	333	$1,030	2.70	0.008
	Asian 75%	Delinquent	2	$1,075	.	.
		Not Delinquent	54	$1,302	0.64	0.610
	Non-Hispanic White 75%	Delinquent	94	$798	.	.
		Not Delinquent	624	$1,285	5.76	0.000
	Other or Missing BISG	Delinquent	236	$849	.	.
		Not Delinquent	1,769	$1,076	3.44	0.001
	Female	Delinquent	172	$817	.	.
		Not Delinquent	1,030	$1,120	4.76	0.000
	Male	Delinquent	212	$724	.	.
		Not Delinquent	1,423	$1,080	6.69	0.000
	Gender Unassigned	Delinquent	118	$861	.	.
		Not Delinquent	714	$1,068	1.88	0.062
Cash Flow Metric #18	Originated Loans	Delinquent	501	$178	.	.
		Not Delinquent	3,160	$232	.	.
		All	3,661	$225	5.05	0.000
	African American 75%	Delinquent	125	$156	.	.
		Not Delinquent	387	$200	2.17	0.031
	Hispanic 75%	Delinquent	45	$134	.	.
		Not Delinquent	330	$193	2.02	0.047
	Asian 75%	Delinquent	2	$275	.	.
		Not Delinquent	55	$206	-0.85	0.518
	Non-Hispanic White 75%	Delinquent	94	$198	.	.
		Not Delinquent	623	$269	3.01	0.003
	Other or Missing BISG	Delinquent	235	$189	.	.
		Not Delinquent	1,765	$234	2.65	0.008
	Female	Delinquent	172	$164	.	.
		Not Delinquent	1,028	$216	3.34	0.001
	Male	Delinquent	211	$178	.	.
		Not Delinquent	1,415	$229	2.88	0.004
	Gender Unassigned	Delinquent	118	$196	.	.
		Not Delinquent	717	$261	2.66	0.008
Cash Flow Metric #19	Originated Loans	Delinquent	26	$85	.	.
		Not Delinquent	286	$214	.	.
		All	312	$203	3.16	0.002
	African American 75%	Delinquent	8	$147	.	.
		Not Delinquent	58	$130	-0.22	0.831
	Hispanic 75%	Delinquent	0	.	.	.

		Not Delinquent	37	$123	.	.
	Asian 75%	Delinquent	0	.	.	.
		Not Delinquent	7	$279	.	.
	Non-Hispanic White 75%	Delinquent	6	$60	.	.
		Not Delinquent	58	$381	3.12	0.003
	Other or Missing BISG	Delinquent	12	$56	.	.
		Not Delinquent	126	$198	2.56	0.012
	Female	Delinquent	5	$40	.	.
		Not Delinquent	98	$203	2.48	0.020
	Male	Delinquent	10	$138	.	.
		Not Delinquent	125	$227	1.13	0.271
	Gender Unassigned	Delinquent	11	$57	.	.
		Not Delinquent	63	$203	2.39	0.019
Cash Flow Metric #20	Originated Loans	Delinquent	498	$498	.	.
		Not Delinquent	3,114	$678	.	.
		All	3,612	$653	2.94	0.003
	African American 75%	Delinquent	124	$522	.	.
		Not Delinquent	379	$675	1.34	0.180
	Hispanic 75%	Delinquent	45	$272	.	.
		Not Delinquent	322	$613	2.86	0.005
	Asian 75%	Delinquent	2	$700	.	.
		Not Delinquent	54	$718	0.02	0.984
	Non-Hispanic White 75%	Delinquent	92	$522	.	.
		Not Delinquent	611	$715	1.49	0.139
	Other or Missing BISG	Delinquent	235	$517	.	.
		Not Delinquent	1,748	$676	1.55	0.122
	Female	Delinquent	172	$586	.	.
		Not Delinquent	1,010	$733	1.15	0.251
	Male	Delinquent	211	$380	.	.
		Not Delinquent	1,403	$632	3.64	0.000
	Gender Unassigned	Delinquent	115	$581	.	.
		Not Delinquent	701	$689	0.81	0.419
Cash Flow Metric #21	Originated Loans	Delinquent	495	$481	.	.
		Not Delinquent	3,122	$1,097	.	.
		All	3,617	$1,013	5.53	0.000
	African American 75%	Delinquent	124	$604	.	.
		Not Delinquent	380	$871	1.09	0.278
	Hispanic 75%	Delinquent	45	$158	.	.
		Not Delinquent	322	$581	4.09	0.000
	Asian 75%	Delinquent	2	$0	.	.
		Not Delinquent	54	$610	2.92	0.005

	Non-Hispanic White 75%	Delinquent	90	$676	.	.
		Not Delinquent	614	$1,765	2.46	0.014
	Other or Missing BISG	Delinquent	234	$408	.	.
		Not Delinquent	1,752	$1,023	5.70	0.000
	Female	Delinquent	172	$558	.	.
		Not Delinquent	1,013	$1,492	3.38	0.001
	Male	Delinquent	211	$387	.	.
		Not Delinquent	1,410	$877	4.01	0.000
	Gender Unassigned	Delinquent	112	$540	.	.
		Not Delinquent	699	$971	2.73	0.007
Cash Flow Metric #22	Originated Loans	Delinquent	506	$1,908	.	.
		Not Delinquent	3,186	$2,584	.	.
		All	3,692	$2,491	9.36	0.000
	African American 75%	Delinquent	126	$1,737	.	.
		Not Delinquent	392	$2,132	3.06	0.002
	Hispanic 75%	Delinquent	46	$1,683	.	.
		Not Delinquent	338	$2,317	4.01	0.000
	Asian 75%	Delinquent	2	$2,382	.	.
		Not Delinquent	55	$2,842	1.23	0.305
	Non-Hispanic White 75%	Delinquent	94	$2,009	.	.
		Not Delinquent	627	$3,038	6.11	0.000
	Other or Missing BISG	Delinquent	238	$1,998	.	.
		Not Delinquent	1,774	$2,565	4.95	0.000
	Female	Delinquent	172	$1,941	.	.
		Not Delinquent	1,035	$2,581	5.61	0.000
	Male	Delinquent	212	$1,862	.	.
		Not Delinquent	1,425	$2,587	6.51	0.000
	Gender Unassigned	Delinquent	122	$1,940	.	.
		Not Delinquent	726	$2,579	3.95	0.000
Cash Flow Metric #23	Originated Loans	Delinquent	506	$4,261	.	.
		Not Delinquent	3,186	$5,378	.	.
		All	3,692	$5,225	4.41	0.000
	African American 75%	Delinquent	126	$4,293	.	.
		Not Delinquent	392	$5,741	1.91	0.057
	Hispanic 75%	Delinquent	46	$2,866	.	.
		Not Delinquent	338	$4,392	6.07	0.000
	Asian 75%	Delinquent	2	$4,650	.	.
		Not Delinquent	55	$5,343	1.49	0.165
	Non-Hispanic White 75%	Delinquent	94	$4,176	.	.
		Not Delinquent	627	$6,355	4.28	0.000

	Other or Missing BISG	Delinquent	238	$4,543	.	.
		Not Delinquent	1,774	$5,141	1.63	0.105
	Female	Delinquent	172	$3,706	.	.
		Not Delinquent	1,035	$5,366	4.89	0.000
	Male	Delinquent	212	$4,165	.	.
		Not Delinquent	1,425	$5,473	3.35	0.001
	Gender Unassigned	Delinquent	122	$5,210	.	.
		Not Delinquent	726	$5,209	0.00	0.999
Cash Flow Metric #24	Originated Loans	Delinquent	506	$211	.	.
		Not Delinquent	3,186	$231	.	.
		All	3,692	$228	1.98	0.048
	African American 75%	Delinquent	126	$219	.	.
		Not Delinquent	392	$216	-0.15	0.884
	Hispanic 75%	Delinquent	46	$201	.	.
		Not Delinquent	338	$236	1.05	0.297
	Asian 75%	Delinquent	2	$200	.	.
		Not Delinquent	55	$193	-0.07	0.954
	Non-Hispanic White 75%	Delinquent	94	$184	.	.
		Not Delinquent	627	$250	3.90	0.000
	Other or Missing BISG	Delinquent	238	$220	.	.
		Not Delinquent	1,774	$228	0.56	0.573
	Female	Delinquent	172	$211	.	.
		Not Delinquent	1,035	$224	0.75	0.454
	Male	Delinquent	212	$206	.	.
		Not Delinquent	1,425	$235	1.88	0.062
	Gender Unassigned	Delinquent	122	$220	.	.
		Not Delinquent	726	$233	0.66	0.511
Cash Flow Metric #25	Originated Loans	Delinquent	503	$284	.	.
		Not Delinquent	3,149	$336	.	.
		All	3,652	$329	2.75	0.006
	African American 75%	Delinquent	125	$337	.	.
		Not Delinquent	386	$340	0.07	0.941
	Hispanic 75%	Delinquent	46	$194	.	.
		Not Delinquent	325	$278	1.57	0.121
	Asian 75%	Delinquent	2	$513	.	.
		Not Delinquent	54	$321	-1.89	0.219
	Non-Hispanic White 75%	Delinquent	94	$312	.	.
		Not Delinquent	622	$386	1.57	0.118
	Other or Missing BISG	Delinquent	236	$260	.	.

		Not Delinquent	1,762	$328	2.56	0.011
	Female	Delinquent	172	$290	.	.
		Not Delinquent	1,024	$336	1.55	0.121
	Male	Delinquent	212	$281	.	.
		Not Delinquent	1,418	$327	1.45	0.149
	Gender Unassigned	Delinquent	119	$280	.	.
		Not Delinquent	707	$353	1.99	0.048

Appendix F. Participant #6

Table 12. Logistic Model for Default Results Within Demographic Group[76]

Demographic Group	Count	Model 1 AUC	Model 2 AUC	Model 3 AUC
African American 75%	528	0.712	0.752	0.766
Hispanic 75%	328	0.736	0.758	0.759
Non-Hispanic White 75%	732	0.775	0.766	0.802
Other or Missing BISG	2,074	0.694	0.667	0.684
Female	1,231	0.749	0.700	0.711
Male	1,660	0.716	0.684	0.702
Gender Unassigned	885	0.737	0.727	0.738
Originated Loans	3,776	0.720	0.675	0.688

[76] The ROC analyses are restricted to the race/ethnicity or gender group listed and uses an indicator for "delinquent" as the reference variable and the listed score as the rating. No model was run for the Asian 75% demographic group because it had fewer than 5 delinquent loans. The estimation samples may differ slightly from the displayed count based on missing values and perfect prediction among the set of predictor variables.

Appendix F. Participant #6

Table 13. Model 1 Specification Within Race / Ethnicity Group[77]

Control Variable	Comparison Group	African American 75%		Hispanic 75%		Non-Hispanic White 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Hard Pull Not Available	Hard Pull Available	1.03	0.973	0.84	0.821	2.39	0.479
FICO score	--	0.99	0.000	0.99	0.001	0.99	0.000
Missing FICO score	Not Missing FICO score	2.13	0.191	4.85	0.039	3.48	0.161
BK score	--	1.00	0.218	1.00	0.954	1.00	0.008
Missing BK score	Not Missing BK score	1.07	0.939	0.82	0.829	0.35	0.449
# of open accounts on credit report	--	0.55	0.144	.	.	0.95	0.896
Missing # of open accounts on credit report	Not Missing # of open accounts on credit report	52.38	0.001	.	.	9.06	0.059
$ amount of unpaid balances on credit report	--	1.00	0.368	.	.	1.00	0.159
Missing $ amount of unpaid balances on credit report	Not Missing $ amount of unpaid balances on credit report	.	.	.	.	.	.
$ amount of monthly payments on credit report	--	1.00	0.724	.	.	1.00	0.343
Missing $ amount of monthly payments on credit report	Not Missing $ amount of monthly payments on credit report	.	.	.	.	.	.
$ Credit limit of revolving accounts on credit report	--	1.00	0.902	.	.	1.00	0.646
Missing $ Credit limit of revolving accounts on credit report	Not Missing $ Credit limit of revolving accounts on credit report	0.51	0.603	.	.	3.88	0.595

77 No model was run for the Asian 75% demographic group because it had fewer than 5 delinquent loans.

$ unpaid balances of revolving accounts on credit report	--	1.00	0.136	.	.	1.00	0.310
Missing $ unpaid balances of revolving accounts on credit report	Not Missing $ unpaid balances of revolving accounts on credit report	.	.	.	.	.	.
% utilization of revolving accounts on credit report	--	1.15	0.947	.	.	0.00	0.073
Missing % utilization of revolving accounts on credit report	Not Missing % utilization of revolving accounts on credit report	.	.	.	.	.	.
Constant	--	15.83	0.066	785.18	0.004	395.27	0.002
Pseudo R Squared		0.100		0.113		0.141	
AUC		0.712		0.736		0.775	
Sample Size		528		331		732	

Appendix F. Participant #6

Table 14. Model 1 Specification Within Gender Group

Control Variable	Comparison Group	Female		Male	
		Odds Ratio	P-Value	Odds Ratio	P-Value
Hard Pull Not Available	Hard Pull Available	0.58	0.522	0.89	0.825
FICO score	--	0.99	0.000	0.99	0.000
Missing FICO score	Not Missing FICO score	0.38	0.331	3.06	0.005
BK score	--	1.00	0.045	1.00	0.007
Missing BK score	Not Missing BK score	3.73	0.244	0.82	0.718
# of open accounts on credit report	--	2.03	0.287	0.50	0.406
Missing # of open accounts on credit report	Not Missing # of open accounts on credit report	2,681,411.55	0.000	14.41	0.357
$ amount of unpaid balances on credit report	--	1.00	0.799	1.00	0.018
Missing $ amount of unpaid balances on credit report	Not Missing $ amount of unpaid balances on credit report	.	.	.	.
$ amount of monthly payments on credit report	--	1.00	0.001	1.00	0.467
Missing $ amount of monthly payments on credit report	Not Missing $ amount of monthly payments on credit report	.	.	.	.
$ Credit limit of revolving accounts on credit report	--	1.00	0.002	1.00	0.079
Missing $ Credit limit of revolving accounts on credit report	Not Missing $ Credit limit of revolving accounts on credit report	97.34	0.426	172,866,315,707,000,000,000.00	0.078
$ unpaid balances of revolving accounts on credit report	--	1.00	0.192	1.00	0.744
Missing $ unpaid balances of revolving accounts on credit report	Not Missing $ unpaid balances of revolving accounts on credit report	.	.	.	.

% utilization of revolving accounts on credit report	--	0.00	0.002	0.01	0.022
Missing % utilization of revolving accounts on credit report	Not Missing % utilization of revolving accounts on credit report	.	.	.	.
Constant	--	354.00	0.009	314.82	0.001
Pseudo R Squared		0.121		0.092	
AUC		0.749		0.716	
Sample Size		1,231		1,660	

Appendix F. Participant #6

Table 15. Model 2 Specification Within Race / Ethnicity Group[78]

Control Variable	Comparison Group	African American 75%		Hispanic 75%		Non-Hispanic White 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Metric #1	--	0.83	0.531	1.18	0.775	0.69	0.023
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	10,482.02	0.000	.	.	0.00	0.000
Cash Flow Metric #2	--	1.08	0.734	0.98	0.971	0.71	0.018
Missing Cash Flow Metric #2	Not Missing Cash Flow Metric #2	14.53	0.213	.	.	264,087,045.32	0.000
Cash Flow Metric #3	--	0.30	0.160	0.09	0.333	0.44	0.146
Missing Cash Flow Metric #3	Not Missing Cash Flow Metric #3	0.00	0.000	.	.	387.08	0.121
Cash Flow Metric #4	--	1.00	0.702	1.00	0.817	0.98	0.168
Missing Cash Flow Metric #4	Not Missing Cash Flow Metric #4	27,446,636,133,900,000,000.00	0.000	.	.	.	.
Cash Flow Metric #5	--	0.91	0.745	0.91	0.881	0.76	0.024
Missing Cash Flow Metric #5	Not Missing Cash Flow Metric #5	216,369,794,991,000,000.00	0.000	.	.	.	.
Cash Flow Metric #6	--	1.11	0.154	0.98	0.869	0.98	0.852
Missing Cash Flow Metric #6	Not Missing Cash Flow Metric #6	30,971,197,385,300,000,000.00	0.000	.	.	0.00	0.000
Cash Flow Metric #8	--	1.19	0.437	0.91	0.876	0.72	0.036
Missing Cash Flow Metric #8	Not Missing Cash Flow Metric #8	0.00	0.000	.	.	.	.
Cash Flow Metric #9	--	.	.	.	.	.	.
Missing Cash Flow Metric #9	Not Missing Cash Flow Metric #9	.	.	.	.	.	.
Cash Flow Metric #10	--	0.90	0.646	1.05	0.933	1.32	0.026
Missing Cash Flow Metric #10	Not Missing Cash Flow Metric #10	.	.	.	.	.	.
Cash Flow Metric #11	--	0.83	0.696	0.96	0.954	2.06	0.048
Missing Cash Flow Metric #11	Not Missing Cash Flow Metric #11	0.00	0.000	.	.	.	.

78 No model was run for the Asian 75% demographic group because it had fewer than 5 delinquent loans.

Cash Flow Metric #13	--	0.06	0.047	0.08	0.303	0.46	0.462
Missing Cash Flow Metric #13	Not Missing Cash Flow Metric #13	1,256,459.27	0.000	.	.	.	.
Cash Flow Metric #14	--	0.80	0.866	0.71	0.886	1.41	0.725
Missing Cash Flow Metric #14	Not Missing Cash Flow Metric #14	17,236,918,463.06	0.000	.	.	0.08	0.306
Cash Flow Metric #16	--	0.85	0.892	2.00	0.748	0.25	0.180
Missing Cash Flow Metric #16	Not Missing Cash Flow Metric #16	.	.	.	.	.	.
Cash Flow Metric #17	--	0.77	0.828	1.66	0.770	0.62	0.551
Missing Cash Flow Metric #17	Not Missing Cash Flow Metric #17	0.00	0.000	.	.	.	.
Cash Flow Metric #18	--	0.25	0.323	0.58	0.813	0.72	0.683
Missing Cash Flow Metric #18	Not Missing Cash Flow Metric #18	0.00	0.000	.	.	0.00	0.209
Cash Flow Metric #20	--	0.98	0.907	0.67	0.136	0.98	0.836
Missing Cash Flow Metric #20	Not Missing Cash Flow Metric #20	0.00	0.000	.	.	.	.
Cash Flow Metric #21	--	1.05	0.515	0.65	0.193	1.01	0.874
Missing Cash Flow Metric #21	Not Missing Cash Flow Metric #21	0.00	0.000	.	.	2,599,065,747,920.05	0.000
Cash Flow Metric #22	--	0.96	0.973	0.94	0.969	1.16	0.840
Missing Cash Flow Metric #22	Not Missing Cash Flow Metric #22	.	.	.	.	.	.
Cash Flow Metric #23	--	0.97	0.429	0.73	0.031	0.97	0.505
Missing Cash Flow Metric #23	Not Missing Cash Flow Metric #23	.	.	.	.	.	.
Cash Flow Metric #24	--	2.11	0.546	0.91	0.962	0.38	0.371
Missing Cash Flow Metric #24	Not Missing Cash Flow Metric #24	.	.	.	.	.	.
Cash Flow Metric #25	--	2.02	0.539	1.23	0.906	1.09	0.918
Missing Cash Flow Metric #25	Not Missing Cash Flow Metric #25	0.00	0.000	.	.	.	.
Cash Flow Metric #19	--	7.11	0.126	.	.	0.05	0.079
Missing Cash Flow Metric #19	Not Missing Cash Flow Metric #19	3.15	0.016	.	.	4.01	0.033
Missing All Cash flow Metrics	Not Missing Any Cash flow Metrics	.	.	.	.	.	.
Constant	--	0.17	0.000	0.51	0.142	0.29	0.068
Pseudo R Squared		0.148		0.129		0.158	
AUC		0.752		0.758		0.766	
Sample Size		528		328		732	

Appendix F. Participant #6
Table 16. Model 2 Specification Within Gender Group

Control Variable	Comparison Group	Female		Male	
		Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Metric #1	--	0.69	0.075	0.94	0.484
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	252.88	0.000	8,048.73	0.000
Cash Flow Metric #2	--	0.77	0.237	1.04	0.348
Missing Cash Flow Metric #2	Not Missing Cash Flow Metric #2	0.00	0.000	576.45	0.000
Cash Flow Metric #3	--	1.11	0.699	0.93	0.728
Missing Cash Flow Metric #3	Not Missing Cash Flow Metric #3	2,876.98	0.000	0.00	0.000
Cash Flow Metric #4	--	0.99	0.242	0.99	0.264
Missing Cash Flow Metric #4	Not Missing Cash Flow Metric #4	831,539,546,697,458.00	0.000	3,128,652,849.14	0.000
Cash Flow Metric #5	--	0.82	0.365	1.02	0.791
Missing Cash Flow Metric #5	Not Missing Cash Flow Metric #5	0.00	0.000	.	.
Cash Flow Metric #6	--	1.05	0.563	0.98	0.551
Missing Cash Flow Metric #6	Not Missing Cash Flow Metric #6	5.22	0.242	12.74	0.041
Cash Flow Metric #8	--	0.84	0.401	1.02	0.602
Missing Cash Flow Metric #8	Not Missing Cash Flow Metric #8	0.00	0.000	0.00	0.000
Cash Flow Metric #9	--	.	.	.	.
Missing Cash Flow Metric #9	Not Missing Cash Flow Metric #9	.	.	.	.
Cash Flow Metric #10	--	1.21	0.348	0.97	0.387
Missing Cash Flow Metric #10	Not Missing Cash Flow Metric #10	.	.	.	.
Cash Flow Metric #11	--	0.61	0.202	1.22	0.300
Missing Cash Flow Metric #11	Not Missing Cash Flow Metric #11	63.42	0.063	.	.
Cash Flow Metric #13	--	2.33	0.644	0.69	0.588
Missing Cash Flow Metric #13	Not Missing Cash Flow Metric #13	17.33	0.258	.	.
Cash Flow Metric #14	--	1.37	0.862	1.21	0.784

Missing Cash Flow Metric #14	Not Missing Cash Flow Metric #14	0.00	0.003	0.00	0.000
Cash Flow Metric #16	--	2.38	0.622	0.64	0.458
Missing Cash Flow Metric #16	Not Missing Cash Flow Metric #16	.	.	.	.
Cash Flow Metric #17	--	3.12	0.502	0.99	0.988
Missing Cash Flow Metric #17	Not Missing Cash Flow Metric #17	598.50	0.222	0.00	0.000
Cash Flow Metric #18	--	2.32	0.637	0.98	0.970
Missing Cash Flow Metric #18	Not Missing Cash Flow Metric #18	46.38	0.044	.	.
Cash Flow Metric #20	--	1.03	0.755	0.84	0.023
Missing Cash Flow Metric #20	Not Missing Cash Flow Metric #20	0.00	0.000	0.00	0.000
Cash Flow Metric #21	--	0.96	0.512	0.91	0.133
Missing Cash Flow Metric #21	Not Missing Cash Flow Metric #21	0.00	0.000	0.00	0.000
Cash Flow Metric #22	--	0.28	0.446	0.72	0.486
Missing Cash Flow Metric #22	Not Missing Cash Flow Metric #22	.	.	.	.
Cash Flow Metric #23	--	0.94	0.187	1.02	0.368
Missing Cash Flow Metric #23	Not Missing Cash Flow Metric #23	.	.	.	.
Cash Flow Metric #24	--	5.65	0.309	2.90	0.147
Missing Cash Flow Metric #24	Not Missing Cash Flow Metric #24	.	.	.	.
Cash Flow Metric #25	--	4.41	0.380	1.60	0.345
Missing Cash Flow Metric #25	Not Missing Cash Flow Metric #25	0.00	0.000	.	.
Cash Flow Metric #19	--	0.08	0.415	0.84	0.727
Missing Cash Flow Metric #19	Not Missing Cash Flow Metric #19	5.53	0.013	2.17	0.030
Missing All Cash flow Metrics	Not Missing Any Cash flow Metrics	.	.	.	.
Constant	--	0.13	0.000	0.17	0.000
Pseudo R Squared		0.079		0.073	
AUC		0.700		0.684	
Sample Size		1,231		1,660	

Appendix F. Participant #6

Table 17. Model 3 Specification Within Race / Ethnicity Group[79]

Control Variable	Comparison Group	African American 75%		Hispanic 75%		Non-Hispanic White 75%	
		Odds Ratio	P-Value	Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Metric #1	--	0.95	0.873	1.16	0.816	0.69	0.016
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	322.23	0.047	.	.	.	.
Cash Flow Metric #2	--	1.21	0.450	0.98	0.973	0.72	0.024
Missing Cash Flow Metric #2	Not Missing Cash Flow Metric #2	52.52	0.169	.	.	31,585,350.32	0.000
Cash Flow Metric #3	--	0.33	0.220	0.09	0.345	0.44	0.189
Missing Cash Flow Metric #3	Not Missing Cash Flow Metric #3	.	.	.	.	888.39	0.010
Cash Flow Metric #4	--	1.01	0.533	0.99	0.643	0.96	0.099
Missing Cash Flow Metric #4	Not Missing Cash Flow Metric #4	.	.	.	.	.	.
Cash Flow Metric #5	--	1.08	0.807	0.91	0.885	0.77	0.046
Missing Cash Flow Metric #5	Not Missing Cash Flow Metric #5	.	.	.	.	.	.
Cash Flow Metric #6	--	1.11	0.113	0.97	0.835	0.98	0.809
Missing Cash Flow Metric #6	Not Missing Cash Flow Metric #6	117,138,257,788.01	0.000	.	.	0.00	0.000
Cash Flow Metric #8	--	1.31	0.260	0.91	0.885	0.72	0.039
Missing Cash Flow Metric #8	Not Missing Cash Flow Metric #8	0.00	0.000	.	.	.	.
Cash Flow Metric #9	--	.	.	.	.	.	.
Missing Cash Flow Metric #9	Not Missing Cash Flow Metric #9	.	.	.	.	.	.
Cash Flow Metric #10	--	0.80	0.358	1.06	0.932	1.33	0.024
Missing Cash Flow Metric #10	Not Missing Cash Flow Metric #10	.	.	.	.	.	.
Cash Flow Metric #11	--	0.81	0.680	1.03	0.966	2.30	0.038

79 No model was run for the Asian 75% demographic group because it had fewer than 5 delinquent loans.

Missing Cash Flow Metric #11	Not Missing Cash Flow Metric #11	0.00	0.100	.	.	.	.
Cash Flow Metric #13	--	0.04	0.410	0.07	0.597	5.91	0.441
Missing Cash Flow Metric #13	Not Missing Cash Flow Metric #13	943.42	0.028	.	.	0.00	0.067
Cash Flow Metric #14	--	0.55	0.872	0.70	0.940	22.40	0.173
Missing Cash Flow Metric #14	Not Missing Cash Flow Metric #14	.	.	.	.	22.19	0.094
Cash Flow Metric #16	--	0.69	0.918	1.77	0.900	3.10	0.626
Missing Cash Flow Metric #16	Not Missing Cash Flow Metric #16	.	.	.	.	.	.
Cash Flow Metric #17	--	0.52	0.856	1.39	0.940	10.07	0.294
Missing Cash Flow Metric #17	Not Missing Cash Flow Metric #17	0.00	0.080	.	.	.	.
Cash Flow Metric #18	--	0.16	0.622	0.55	0.902	11.44	0.272
Missing Cash Flow Metric #18	Not Missing Cash Flow Metric #18	0.00	0.383	.	.	0.00	0.015
Cash Flow Metric #20	--	1.23	0.555	0.53	0.035	0.93	0.517
Missing Cash Flow Metric #20	Not Missing Cash Flow Metric #20	.	.	.	.	0.00	0.000
Cash Flow Metric #21	--	1.35	0.337	0.49	0.048	0.96	0.600
Missing Cash Flow Metric #21	Not Missing Cash Flow Metric #21	0.00	0.000	.	.	.	.
Cash Flow Metric #22	--	1.44	0.920	1.02	0.996	0.08	0.253
Missing Cash Flow Metric #22	Not Missing Cash Flow Metric #22	.	.	.	.	.	.
Cash Flow Metric #23	--	0.76	0.366	.	.	.	.
Missing Cash Flow Metric #23	Not Missing Cash Flow Metric #23	.	.	.	.	.	.
Cash Flow Metric #24	--	1.48	0.913	0.70	0.937	5.74	0.445
Missing Cash Flow Metric #24	Not Missing Cash Flow Metric #24	.	.	.	.	.	.
Cash Flow Metric #25	--	1.42	0.922	1.09	0.984	18.20	0.203
Missing Cash Flow Metric #25	Not Missing Cash Flow Metric #25	.	.	.	.	.	.
Cash Flow Metric #19	--	5.18	0.219	.	.	0.08	0.114
Missing Cash Flow Metric #19	Not Missing Cash Flow Metric #19	3.35	0.012	.	.	4.34	0.021
Missing All Cash flow Metrics	Not Missing Any Cash flow Metrics	.	.	.	.	.	.
Cash Flow Metric #7	--	1.15	0.642	0.79	0.188	1.11	0.185
Missing Cash Flow Metric #7	Not Missing Cash Flow Metric #7	2.14	0.185	.	.	1.32	0.774
Cash Flow Metric #12	--	0.57	0.885	1.05	0.991	23.24	0.210
Missing Cash Flow Metric #12	Not Missing Cash Flow Metric #12	1.52	0.865	.	.	.	.

Cash Flow Metric #15	--	1.29	0.416	0.71	0.028	0.86	0.055
Missing Cash Flow Metric #15	Not Missing Cash Flow Metric #15	0.02	0.069	.	.	0.12	0.173
Cash Flow Data Quality Bucket A	Cash Flow Data Quality Bucket C	1.23	0.383	0.86	0.709	2.00	0.006
Cash Flow Data Quality Bucket B		150,065,946,232.67	0.006	.	.	25,356,461.47	0.001
Constant	--	0.15	0.000	0.55	0.259	0.18	0.013
Pseudo R Squared		0.164		0.132		0.189	
AUC		0.766		0.759		0.802	
Sample Size		528		323		732	

Appendix F. Participant #6

Table 18. Model 3 Specification Within Gender Group

Control Variable	Comparison Group	Female		Male	
		Odds Ratio	P-Value	Odds Ratio	P-Value
Cash Flow Metric #1	--	0.71	0.097	0.96	0.644
Missing Cash Flow Metric #1	Not Missing Cash Flow Metric #1	.	.	75,636.58	0.000
Cash Flow Metric #2	--	0.78	0.271	1.05	0.272
Missing Cash Flow Metric #2	Not Missing Cash Flow Metric #2	0.00	0.000	0.04	0.073
Cash Flow Metric #3	--	1.12	0.686	0.91	0.649
Missing Cash Flow Metric #3	Not Missing Cash Flow Metric #3	280,674.02	0.000	0.00	0.000
Cash Flow Metric #4	--	0.99	0.162	0.98	0.194
Missing Cash Flow Metric #4	Not Missing Cash Flow Metric #4	408,407,484,922,759.00	0.000	.	.
Cash Flow Metric #5	--	0.84	0.421	1.02	0.690
Missing Cash Flow Metric #5	Not Missing Cash Flow Metric #5	.	.	.	.
Cash Flow Metric #6	--	1.03	0.723	0.98	0.660
Missing Cash Flow Metric #6	Not Missing Cash Flow Metric #6	4.82	0.224	24.27	0.029
Cash Flow Metric #8	--	0.85	0.439	1.03	0.512
Missing Cash Flow Metric #8	Not Missing Cash Flow Metric #8	0.00	0.000	0.00	0.000
Cash Flow Metric #9	--	.	.	.	.
Missing Cash Flow Metric #9	Not Missing Cash Flow Metric #9	.	.	.	.
Cash Flow Metric #10	--	1.20	0.385	0.97	0.348
Missing Cash Flow Metric #10	Not Missing Cash Flow Metric #10	.	.	.	.
Cash Flow Metric #11	--	0.62	0.221	1.27	0.255
Missing Cash Flow Metric #11	Not Missing Cash Flow Metric #11	.	.	.	.
Cash Flow Metric #13	--	1.91	0.739	66.16	0.020
Missing Cash Flow Metric #13	Not Missing Cash Flow Metric #13	.	.	.	.
Cash Flow Metric #14	--	1.02	0.990	105.24	0.009

Missing Cash Flow Metric #14	Not Missing Cash Flow Metric #14	.	.	.	.
Cash Flow Metric #16	--	1.92	0.730	59.17	0.018
Missing Cash Flow Metric #16	Not Missing Cash Flow Metric #16	.	.	.	.
Cash Flow Metric #17	--	2.61	0.595	90.10	0.007
Missing Cash Flow Metric #17	Not Missing Cash Flow Metric #17	.	.	.	.
Cash Flow Metric #18	--	1.87	0.741	85.03	0.010
Missing Cash Flow Metric #18	Not Missing Cash Flow Metric #18	.	.	.	.
Cash Flow Metric #20	--	1.30	0.681	1.08	0.889
Missing Cash Flow Metric #20	Not Missing Cash Flow Metric #20	0.00	0.000	0.00	0.000
Cash Flow Metric #21	--	1.21	0.763	1.16	0.771
Missing Cash Flow Metric #21	Not Missing Cash Flow Metric #21	0.00	0.000	.	.
Cash Flow Metric #22	--	0.34	0.548	0.01	0.004
Missing Cash Flow Metric #22	Not Missing Cash Flow Metric #22	.	.	.	.
Cash Flow Metric #23	--	0.75	0.649	0.79	0.643
Missing Cash Flow Metric #23	Not Missing Cash Flow Metric #23	.	.	.	.
Cash Flow Metric #24	--	4.36	0.419	281.49	0.002
Missing Cash Flow Metric #24	Not Missing Cash Flow Metric #24	.	.	.	.
Cash Flow Metric #25	--	3.62	0.475	141.13	0.003
Missing Cash Flow Metric #25	Not Missing Cash Flow Metric #25	0.02	0.349	0.17	0.687
Cash Flow Metric #19	--	0.09	0.433	0.75	0.605
Missing Cash Flow Metric #19	Not Missing Cash Flow Metric #19	5.39	0.015	2.33	0.020
Missing All Cash flow Metrics	Not Missing Any Cash flow Metrics	.	.	.	.
Cash Flow Metric #7	--	1.35	0.636	1.37	0.548
Missing Cash Flow Metric #7	Not Missing Cash Flow Metric #7	1.32	0.763	2.44	0.213
Cash Flow Metric #12	--	0.27	0.709	134.13	0.004
Missing Cash Flow Metric #12	Not Missing Cash Flow Metric #12	.	.	.	.
Cash Flow Metric #15	--	1.23	0.747	1.29	0.628
Missing Cash Flow Metric #15	Not Missing Cash Flow Metric #15	0.00	0.024	.	.
Cash Flow Data Quality Bucket A	Cash Flow Data Quality Bucket C	1.30	0.158	1.49	0.016
Cash Flow Data Quality Bucket B		0.68	0.781	.	.

Constant	--	0.11	0.000	0.12	0.000
Pseudo R Squared		0.085		0.087	
AUC		0.711		0.702	
Sample Size		1,231		1,660	

APPENDIX G: Technical Glossary

AUC Statistics: The Area Under the Receiver Operating Characteristic ("ROC") curve, or "AUC" statistic, is a standard measure of model fit or performance used by developers of credit models and other risk models. Intuitively, it measures how well a scoring model performs in distinguishing accounts that perform from those that do not. A scoring model that does no better than random chance would have an AUC statistic of 0.5, and a scoring model that perfectly predicts loan performance would have an AUC of 1.0.

Difference in Means Test: A difference in means test is used to determine whether two sample groups (e.g. applicants or borrowers) have mean values for a given attribute that are, statistically speaking, different from one another and not likely the result of random chance.

Odds Ratios: We use logistic models to estimate the effect of an explanatory variable on a binary outcome variable, i.e., an indicator of whether or not a borrower charged off. These estimates are expressed as "Odds Ratios" in the tables. For example, an odds ratio estimated for a demographic group indicator variable is a measure of the relative likelihood that one group of applicants will charge off as compared to another group. An estimated odds ratio of 1.0 indicates equality in the likelihood of charge-off between the groups being compared; a value between zero and 1.0 indicates that the likelihood of charge-off is lower for the target group than for the comparison group. An odds ratio greater than 1.0 indicates that the likelihood of charge-off is greater for the target group than for the comparison group.

Marginal Effects: Logistic model estimates of prohibited basis differences in charge-off rates can also be expressed as "average marginal effects." An average marginal effect represents the estimated difference in charge-off rates (measured in percentage points) between a target group and its comparison group, after controlling for the effects of the other explanatory variables in the model. Marginal effects can provide a more intuitive interpretation to model estimates than odds ratios in certain contexts.

p-Value: The statistical significance is indicated by the p-value statistic. Intuitively, the p-value represents the probability that the differences observed between groups has occurred only by chance.[80] The lower the number, the more confident one can be that the difference observed

[80] More technically, it represents the probability of observing a difference as large or larger than observed under the null hypothesis of a difference of zero.

between groups is not a result of random chance. For purposes of this analysis, the threshold for statistical significance is five percent, or a p-value equal to 0.05 or less.[81] The level of statistical significance is often referred to as a "confidence level" in terms of a percentage. The confidence level is equal to one minus the significance level, and represents the probability that the observed difference between the groups has not occurred by chance. For example, a 95% confidence level corresponds to a five percent significance level. We use the expression "statistically significant" in this report to mean significant at the 95% confidence level unless specifically stated otherwise.

81 In our experience, the federal financial regulatory and enforcement agencies typically use the 95-percent confidence level (five-percent significance level) as the threshold to determine statistical significance.

Digital version available at finreglab.org

Published by FinRegLab, Inc.

805 15th Street NW #1100
Washington, DC 20005
United States