CONSUMER FINANCIAL PROTECTION BUREAU | SEPTEMBER 2019

Credit Characteristics, Credit Engagement Tools, and Financial Well-Being

Innovation Insights

Consumer Financial Protection Bureau

- Éva Nagypál, Ph.D.
- Jeremy Tobacman, Ph.D.

Executive Summary

This report presents results from a joint research study between the Consumer Financial Protection Bureau (CFPB) and Credit Karma. Credit Karma describes itself as "a personal finance technology company" that "offers a suite of products for members to monitor and improve credit health."[1,2] The purpose of the study is to examine how consumers' subjective financial well-being relates to objective measures of consumers' financial health, specifically, consumers' credit report characteristics. The study also seeks to relate consumers' subjective financial well-being to consumers' engagement with financial information through educational tools. A better understanding of these relationships helps uncover the factors that work together to determine consumers' financial well-being and helps inform the CFPB's long-term strategy for improving financial capability.

Financial well-being is measured using the CFPB's Financial Well-Being (FWB) Scale which results in a FWB score. The scale was administered through a voluntary survey that resulted in close to 3,000 de-identified observations on respondents' FWB score matched with background, credit report, and website usage data. The main takeaways are as follows:

- A consumer's credit score is very strongly positively correlated with the FWB score, with a correlation coefficient of 0.44.
- There is a positive correlation between age and the FWB score, but this correlation all but disappears when controlling for credit score.
- In addition to credit score and age, the study identifies seven credit report variables and three engagement variables that correlate strongly with a consumer's FWB score.
    - Credit Report Variables: Credit card limits, holding a credit card, and the number of accounts recently opened with a balance all correlate positively with a consumer's FWB score. Credit card utilization, the number of revolving accounts, the number of collections in the past two years, and having a student loan all correlate negatively with a consumer's FWB score.
    - Engagement with Credit Karma Platform Variables: A consumer's FWB score correlates positively with the number of times the credit simulator was used and the number of times credit factors were reviewed. Finally, FWB score correlates negatively with the number of emails from Credit Karma opened in the last sixty days.

[1] Based on press releases at https://www.creditkarma.com/about/releases. Latest from Jan 24, 2019.

[2] Credit Karma is not an agent or affiliate of the CFPB. The CFPB does not endorse Credit Karma, its views, or the products or services it offers.

This is the first study of its size to study the relationship between financial well-being and credit score and other credit report variables, on the one hand, and engagement with financial information, on the other. The observed correlations might be causal (with changes in the credit and engagement characteristics in question leading to higher FWB scores) or they might be explained by reverse causality or omitted variables like the propensity to plan. Either way, the results are intriguing and warrant further study of these relationships as the CFPB develops its strategy for improving financial capability using the concept of financial well-being.

Table of contents

1. Introduction

An essential part of the mission of the Consumer Financial Protection Bureau (CFPB or the Bureau) is empowering consumers to take control over their financial lives. Numerous provisions of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 task the Bureau with researching, developing, promoting, and implementing financial literacy programs and activities. In order to measure the success of these efforts, the Bureau undertook rigorous research to understand and formally define financial well-being in ways that allow it to be measured and that allow meaningful comparisons between approaches to achieving it.[3] To measure the defined concept, the Bureau developed the Financial Well-Being (FWB) Scale, a set of questions used to measure an individual's subjective level of financial well-being.[4] This scale was fielded for the first time in 2016 in a nationwide survey of adults in the United States, the National Financial Well-Being Survey.[5] The survey collected a variety of self-reported correlates of financial well-being, such as individual and family characteristics; income; savings; financial experiences, behaviors, skills, and attitudes. To inform the CFPB's long-term strategy for improving financial capability, the CFPB is now interested in better understanding how its measure of financial well-being relates to objective financial characteristics and to consumers' engagement with financial information and skill-building opportunities through educational tools.

Credit Karma is a personal finance technology company that provides consumers with free access to credit scores, reports, and credit and identity monitoring through web and mobile interfaces.[6] The company also gives members access to free tools, including the Credit Score Simulator, a variety of loan calculators, and a direct-dispute process to challenge erroneous credit-report entries.[7] It uses data modeling to "identify financial products that are a good fit for

---

3 CFPB defines financial well-being as a state of being wherein a person can fully meet current and ongoing financial obligations, can feel secure in their financial future, and is able to make choices that allow enjoyment of life. For more information on this definition, see "Financial Well-Being: The goal of financial education," available at https://www.consumerfinance.gov/data-research/research-reports/financial-well-being/.

4 See "Financial Well-Being Scale Development Technical Report," available at https://www.consumerfinance.gov/documents/4800/201705_cfpb_financial-well-being-scale-technical-report.pdf.

5 See "Financial Well-Being in America," available at https://www.consumerfinance.gov/documents/5606/201709_cfpb_financial-well-being-in-America.pdf.

6 See https://www.creditkarma.com/free-credit-score, https://www.creditkarma.com/credit-monitoring, and http://www.creditkarma.com/id-monitoring.

7 See https://www.creditkarma.com/tools/credit-score-simulator, https://www.creditkarma.com/calculators/loan, and https://www.creditkarma.com/advice/i/credit-karma-direct-dispute.

its members”[8] and makes money from partners offering these products.[9] Because of the nature of its services, its engagement with a large membership base of more than 100 million members,[10] and its online platform, Credit Karma is well-positioned to provide subjects and an outlet to study the Financial Well-Being Scale and its correlates among a sample of their members.

This study took place in the context of the CFPB’s work to encourage consumer-friendly innovation in markets for consumer financial products and services. As part of the study, Credit Karma agreed to share insights with the CFPB from a survey on financial well-being and around peoples’ financial characteristics and engagement with the educational tools on its platform.[11] The study was facilitated by the Office of Research, which provided subject matter expertise in data and regression analysis. The Office of Financial Education provided advice regarding financial well-being.

This study has two main goals: (1) to understand how consumers’ subjective assessments of financial well-being relate to objective financial characteristics measured in credit reports, and (2) to explore how these assessments of financial well-being relate to engagement on the Credit Karma platform.[12] Financial well-being is measured by surveying Credit Karma members using the Bureau’s FWB Scale. By linking these survey results to administrative credit report data (i.e., credit report data in the possession of Credit Karma) for a significant number of consumers, correlations between objective financial factors and a consumer’s own sense of financial well-being can be measured. This is also a novel opportunity to study how well-being is related to a consumer’s propensity to review educational materials, access and monitor credit reports and scores, and use other tools available on the Credit Karma website. This report covers the research methodology, implementation details, and analytical results.

[8] *Supra* note 1.

[9] See https://www.creditkarma.com/about.

[10] See https://www.bloomberg.com/news/articles/2019-06-21/credit-karma-wants-to-become-your-trusted-credit-adviser.

[11] See https://www.consumerfinance.gov/about-us/blog/project-catalyst-collaboration-improve-understanding-financial-well-being/.

[12] In other work (see *supra* note 4), the Bureau showed a connection between savings and the habit of savings and financial well-being. This relationship is not tested for here because savings patterns and balances do not show up in credit reports.

2. Methods and Data

For this project, Credit Karma administered a voluntary survey to some of its members between September 7, 2017, and October 31, 2017. During the field period, Credit Karma invited members to participate in the survey when they logged out of their Credit Karma session by asking, "Before you go, as part of a joint project between Credit Karma and the CFPB, would you be willing to take a short survey?" Invitations were presented using a quota sampling scheme[13] targeting 5,000 total responses. No member was invited to participate more than once. Credit Karma users who agreed to take the survey were directed to a Credit Karma/CFPB-branded page. The survey consisted of the full 10-question version of the CFPB's Financial Well-Being Scale. The 10 questions are included as Appendix A. The survey was administered to 4,559 consumers who accepted the invitation, of whom 4,067 provided complete answers. The survey data also contain the start and end time and date of the survey.

Each respondent's survey data were then merged by Credit Karma with their background, credit report, and website usage data. As part of standard business practice, Credit Karma collects background information on its members, including when they signed up for Credit Karma services, their state of residence, and their age. In addition, Credit Karma receives credit report information at every login, conditional on there being at least a one week delay since a prior credit report refresh. As a result, Credit Karma received up-to-date credit report information, including credit score, at the time the survey was administered and could match this information to the survey responses. VantageScore 3.0 credit scores were available at the time of the survey and each month going back to January 2017 depending upon the length of time that an individual respondent had been a registered user of Credit Karma. An additional 252 credit report variables provide information on utilization, performance, and inquiries covering all major consumer financial product categories (mortgages, credit cards, student loans, auto loans, and installment loans).[14] In addition, several variables provide summary measures across several categories (e.g., total number of accounts, number of accounts opened in last 6 months).

[13] In this quota sampling scheme, Credit Karma members were segmented into mutually exclusive subgroups defined by age and credit score and a specified number of responses were sought from each subgroup. The number of responses sought was set to reflect the distribution of Credit Karma members over the age and credit score groups. For example, 2 percent of Credit Karma members are between the age of 18 and 24 and have a credit score below 700, so the quota for this group was set at 100 (2 percent of 5000). After the quota for a subgroup was reached during the data collection phase, no more invitations were extended to members of that subgroup.

[14] The credit report information provided is at most seven days old at the time the survey is taken. A consumer's credit report is updated upon log-in unless the existing credit report information is less than seven days old. See https://help.creditkarma.com/hc/en-ca/articles/115004683866-How-often-does-my-credit-report-information-update-

Engagement metrics are also collected in the normal course of Credit Karma's operations. Relevant variables include the number of educational articles viewed by a member, the number of times they viewed their TransUnion or Equifax credit report information, the number of times they disputed credit report information, and whether they used the Credit Score Simulator tool over the previous 30, 60, and 90 days. Overall, Credit Karma provided 75 variables regarding member engagement and use patterns. These variables were merged by Credit Karma with the survey and credit report data listed above.

The data thus collected and matched by Credit Karma were provided to the CFPB after removing any information that could directly identify, or reasonably be linked to, an individual consumer. Of the 4,067 completed survey responses, Credit Karma provided matched credit report and engagement data for 2,966 respondents. Appendix B compares the matched and unmatched observations in terms of their age and Financial Well-Being Score and finds that they are very similar. All of the analysis that follows is based on the 2,966 matched observations.

3. Descriptive Statistics

The FWB Score is an individual's score on the CFPB Financial Well-Being Scale and is a standardized number between 0 and 100 that quantifies that person's underlying level of financial well-being. The questions that make up the scale were presented in the Credit Karma survey, which allowed Credit Karma then to calculate a FWB score for each respondent.

Figure 1: Distribution of Financial Well-Being Score in the Credit Karma survey and in the CFPB's National Financial Well-Being survey

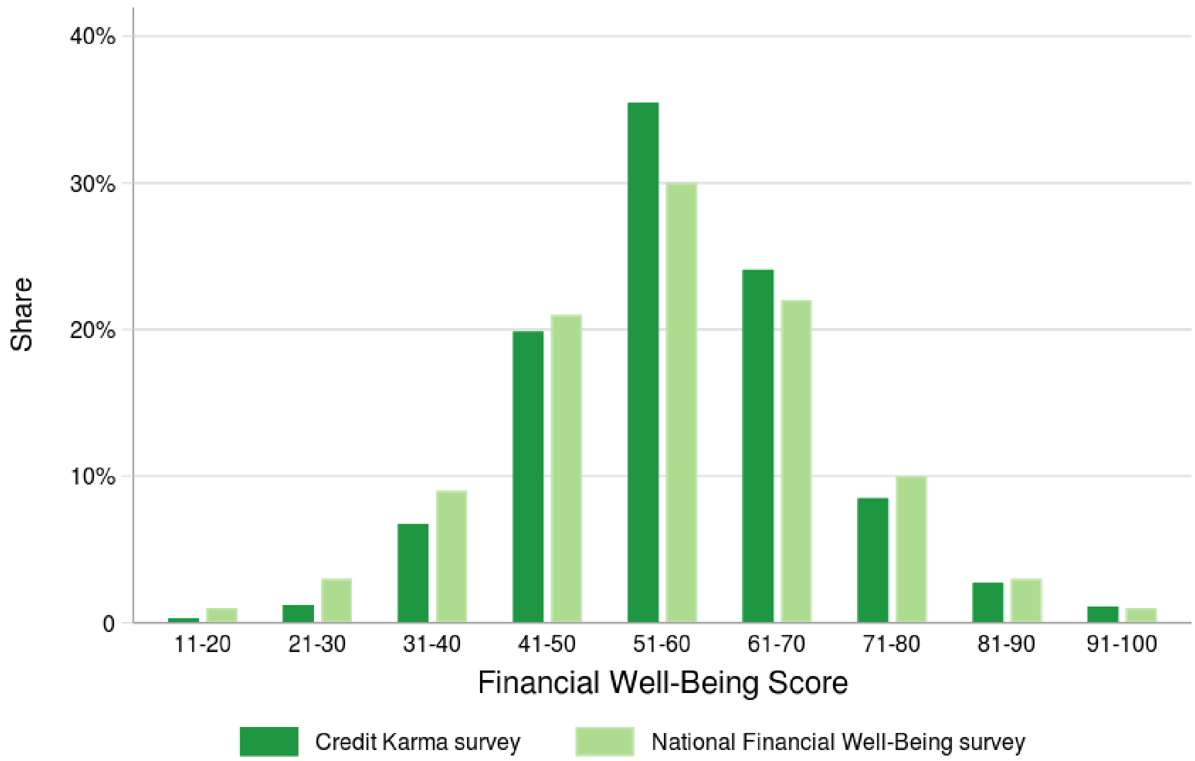
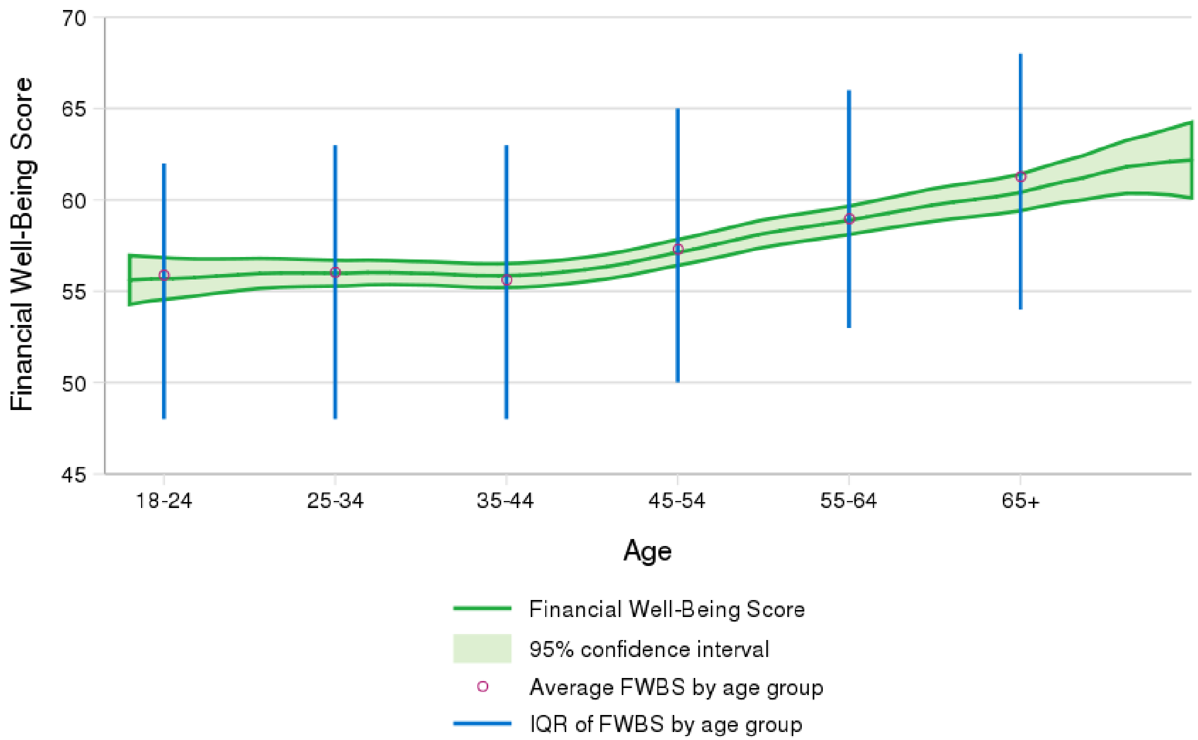


Figure 1 shows the distribution of FWB score among the nationally representative sample of U.S. adults included in the CFPB's National Financial Well-Being Survey and among the Credit Karma respondents.[15] As can be seen, the distributions are quite similar, with the notable difference that Credit Karma respondents are significantly less likely to have low FWB scores

[15] The micro data used here from the CFPB's nationally representative survey about financial well-being have also been publicly released. For more information including data access, see "National Financial Well-Being Survey Public Use File User's Guide," available at https://files.consumerfinance.gov/f/documents/cfpb_nfwbs-puf-user-guide.pdf.

(scores of 40 or below) and are somewhat less likely to have very high FWB scores (scores above 70). Instead, the scores of Credit Karma members are more concentrated towards the middle.

Figure 2: Financial Well-Being Score as a function of age: functional approximation with confidence intervals and average and interquartile range (IQR) by age group

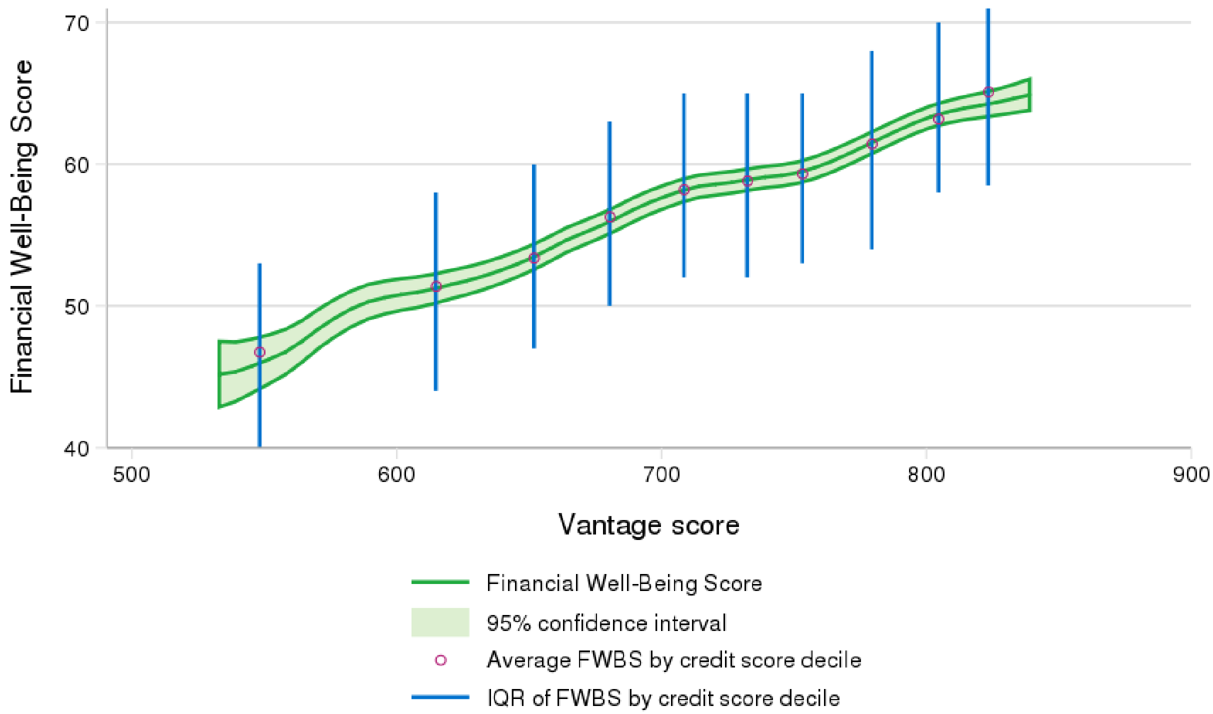


Note: FWB Score is plotted as a local polynomial function of age with 95% confidence intervals and with average and IQR plotted by age group.

Figure 2 shows the relationship between FWB score and age using two methods: first, displaying the relationship between the two using a flexible technique that allows one to uncover a smooth relationship between two variables (known as "local polynomial smoothing") and second showing the average FWB score and the range between the 25^{th} and 75^{th} percentile (the interquartile range) by age group. Both methods clearly show that the FWB score is constant on average at a score of around 56 until age 40, after which it starts to rise, reaching over 60 by age 70. This is again in line with the findings of the National Financial Well-Being Survey, which also found increases with age, especially after 40, albeit from a somewhat lower base of 51 for those aged 34 and younger.[16] The interquartile ranges are between 13 and 15 points which demonstrates that there is substantial variation in FWB score within each age group.

[16] See "Financial Well-Being in America," *supra* note 4, Figure 7, p. 31.

Figure 3: Financial Well-Being Score as a function of VantageScore 3.0 credit score: functional approximation with confidence intervals and average and interquartile range by credit score decile



Note: FWB Score is plotted as a local polynomial function of credit score with 95% confidence intervals and with average and IQR plotted by credit score deciles. Bottom 3 percent of credit score values are excluded due to noise in data.

Next, Figure 3 shows the relationship between FWB score and credit score[17] using the same two methods as above: first, displaying the relationship between the two using a local polynomial smoothing and second showing the average FWB score and its interquartile range for each credit score decile. A credit score measures a consumer's creditworthiness based on data from the consumer's credit report that generally reflects credit use and repayment history. Since credit scores are based on algorithms derived from third-party data from a consumer's financial life, they are considered objective measures of a consumer's credit risk.[18]

As is clear from the figure, the FWB score (which is designed to measure the inherently subjective concept of financial well-being) is strongly positively related to a consumer's credit score (which is designed to be an objective measure of creditworthiness). In fact, the correlation

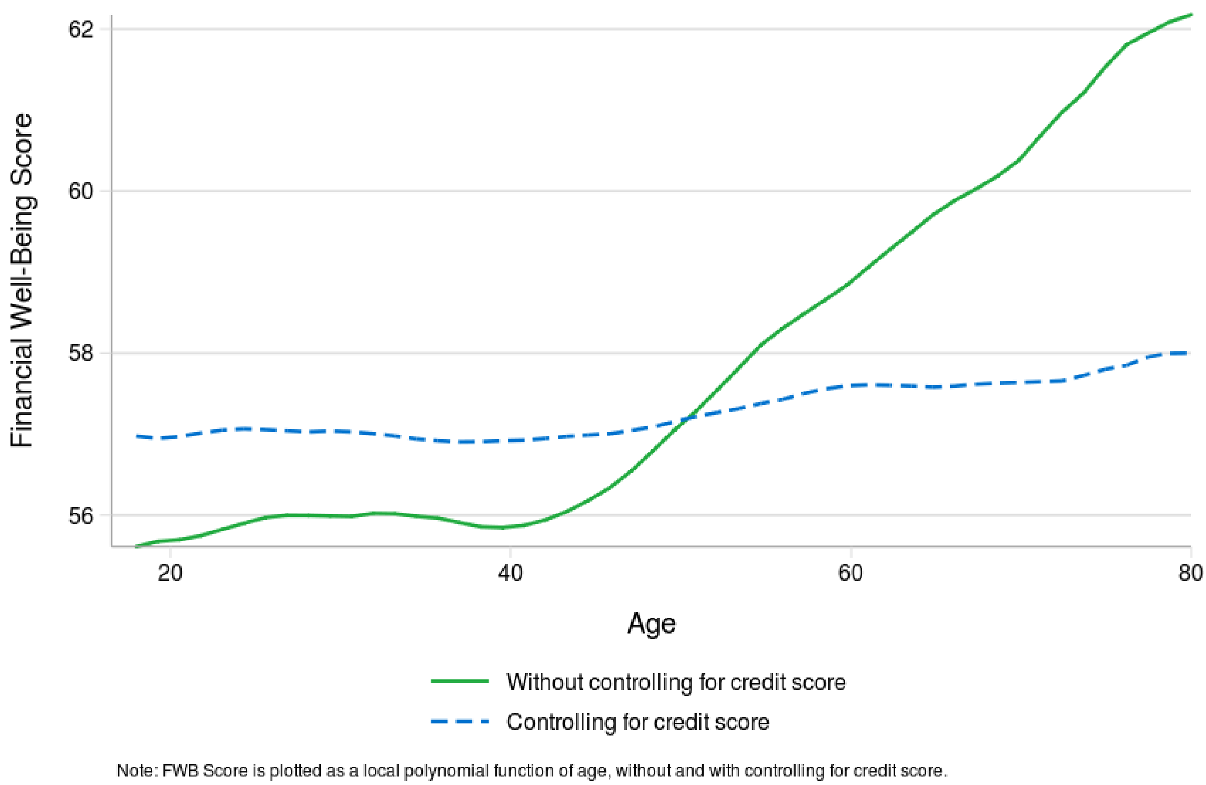
[17] The credit score contained in the data is the VantageScore 3.0 score.

[18] These objective credit report measures complement the objective financial situation measures (income retirement status, etc.) that were studied in "Pathways to financial well-being: Research brief," available at https://www.consumerfinance.gov/data-research/research-reports/pathways-financial-well-being/.

between the two variables is 0.44 and credit score by itself explains 19.7 percent of the variation in FWB scores across the respondents in the data. The average FWB score among consumers in the bottom 10 percent of the credit score distribution is 46.7 while the average among those in the top 10 percent is 65.0. While it is not surprising that the correlation between the FWB score and the credit score is positive and strong, this report is the first one to document the relationship between individuals' assessment of their financial well-being and their creditworthiness.

It is worth noting that while the documented relationship is strong, over 80 percent of the variation in FWB score cannot be explained by one's credit score, possibly indicating that the FWB score is a more complex measure than credit score and is influenced by factors beyond creditworthiness and access to credit. The interquartile range within each credit score decile is 12 to 14 points, implying that there is substantial variation in FWB score within these deciles. The next section explores further other factors in the data that are correlated with the FWB score.

Figure 4: Financial Well-Being Score as a function of age: the effect of controlling for credit score



Before turning to those analyses, Figure 4 documents that most of the observed variation in FWB score with age is related to the fact that, on average, a consumer's credit scores increase with age. The figure shows the relationship between the average FWB score and age with and without controlling for credit score. The strong positive relationship that exists between the FWB score and age after age 40 (shown by the steep line in Figure 4) all but disappears when controlling for credit score.

4. Correlating Credit Characteristics and Credit Engagement with Financial Well-Being

As outlined in the Methods and Data section, for each Credit Karma member in the study, 252 credit report variables and 75 engagement variables are observed. To identify variables that are important correlates of the FWB score in addition to credit score and age,[19] a variable selection method known as supervised block-wise forward selection is performed. The details of this method are described in Appendix C.

This method selected seven credit report variables and three engagement variables that are statistically significant correlates of a consumer's FWB score in a multivariate regression model. The graphs below present 1) the relationship between each of these variables and the FWB score without any controls, and 2) the marginal effect of the variable on the FWB score controlling in each case for credit score, age, and the other nine selected variables. The marginal effect results from running the multivariate regression model with credit score, age, and all the selected variables included, and then plotting the effect of each variable separately holding the other explanatory variables constant at their means.

It is imperative to keep in mind that the relationships found do not imply that selected variables cause the FWB score to increase. Such a conclusion cannot be drawn from a cross-sectional research design such as this one. To demonstrate what the results do mean, consider the example of credit simulator use. The credit simulator use results say that consumers who use the credit simulator tend to a have higher FWB score in the cross-section, even after controlling for many other observable characteristics of the consumer. The relationship could be a result of credit simulator use causing the FWB score to increase (causality), consumers with a higher FWB score choosing to use the credit simulator more often (reverse causality), or a third, omitted variable (such as, for example, a propensity to plan), affecting both outcomes (omitted variable bias).

It is also worth noting that several of the selected credit report variables are strongly correlated with credit score. This is either because they are directly included in credit scoring models (e.g.,

[19] In the regression model, the coefficient on age is allowed to change at age 40.

utilization) or are determined in part by credit score (e.g., credit limit). Given that both credit score and the other selected variables are included in the regression model, the additional effect of the selected variables above and beyond the effect coming through the credit score can be parsed out. This additional effect is what the marginal effects reported below capture.

4.1 Credit Report Variables

In this section, important credit report correlates of the FWB score are identified. Table 1 shows the credit report variables selected by method outlined above, the sign of their conditional correlation with the FWB score, and the marginal effect as each selected variable moves across interquartile range of its distribution (or as it moves from 0 to 1 in the case of variables that take on only these two values).

Table 1: Summary of effect of selected credit report variables on FWB score

Selected variable	Conditional correlation	Interquartile effect*
Credit score	+	4.99
Credit card limit	+	2.23
Utilization of bank card	-	1.97
Number of revolving accounts	-	1.83
Number of collections in past two years	-	1.45
Has a bank card	+	2.48†
Number of accounts opened w/ balance	+	0.80
Has student loan	-	0.98†

*Effects reported are interquartile effects except for 0/1 variables marked with † for which the effect of moving from 0 to 1 is reported.

Figure 5: Financial Well-Being Score as a function of total credit card limit: raw data by deciles and marginal effect based on regression model

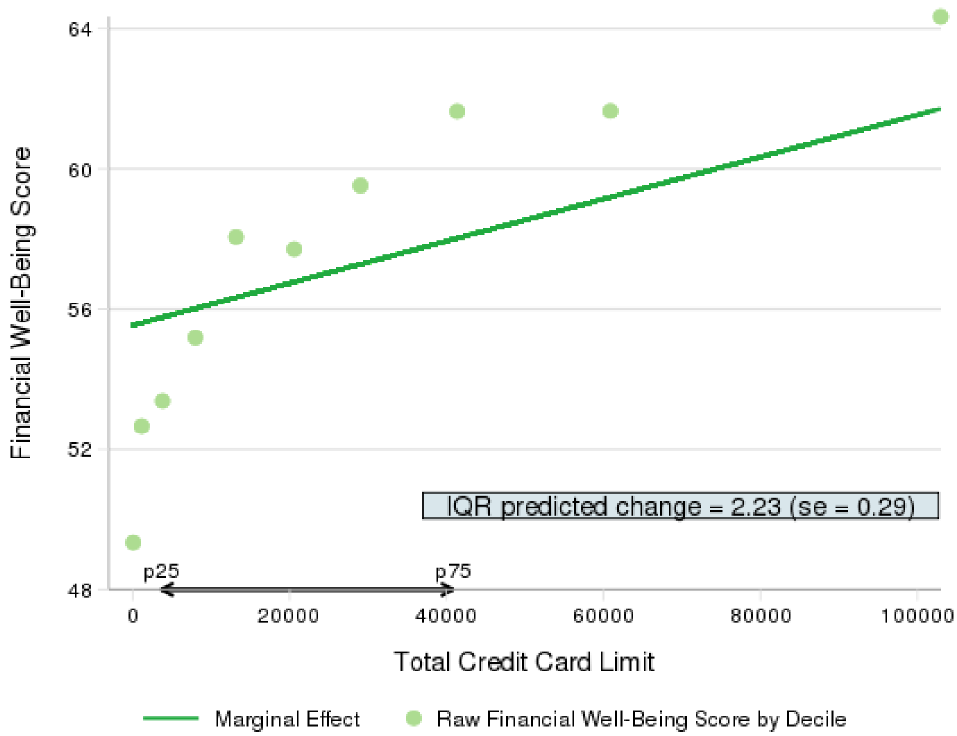
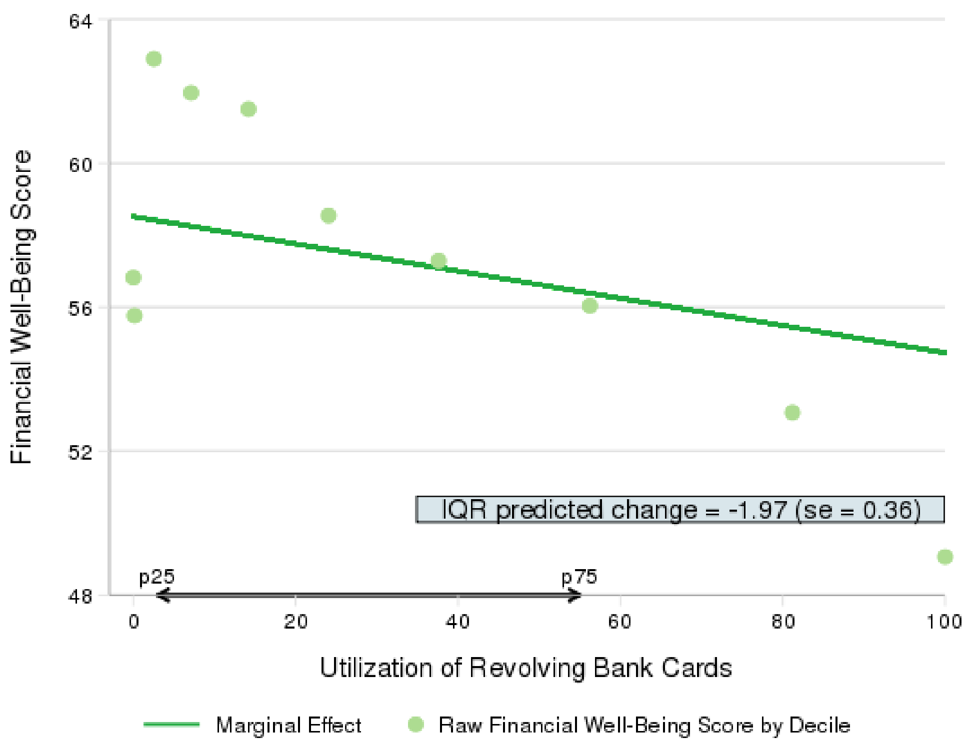


Figure 5 shows the relationship between a consumer's total credit card limit and FWB score. The small circles represent average FWB score in the data within each credit card limit decile. The solid line represents the regression model-implied relationship between the FWB score and credit card limit once credit score and the other control variables are taken into account. As can be seen, there is a strong positive relationship between FWB score and credit card limit without any further controls (i.e., in the raw data). Of course, much of this is due to the fact that both credit card limit and the FWB score increase with a consumer's credit score. The solid line implies, however, that there is a positive relationship between credit card limit and the FWB score even after controlling for credit score and the other control variables. This implies that given two consumers with identical credit scores and identical values on the other control variables, the one with the higher credit card limit is expected to have a higher FWB score. The marginal effect is reported as the predicted change in FWB score as the credit card limit moves across the interquartile range of its distribution. The reported marginal effect of 2.23 in Figure 5 then means that as one moves from the 25th percentile of the credit card limit distribution ($3,700, marked as "p25" in the figures) to the 75th percentile ($40,800, marked as "p75" in the figures) holding all other model variables (credit score, age, and other selected variables) constant, the FWB score is expected to increase by 2.23 points.[20] As a comparison, the regression model implies that moving across the interquartile range of the credit score distribution (from 652 to 780) increases the FWB score by 4.99 points. The 0.29 reported in parentheses is the corresponding standard error.

The fact that this conditional correlation is positive could be due to the fact that total credit limit increases with income or education, for example, which themselves have a positive relationship with the FWB score. It could also be due to the direct impact of more access to credit on the FWB score. Since a direct measure of income or education is not available, it is not possible to assign appropriate weights to these alternative explanations.

Figure 6: Financial Well-Being Score as a function of the utilization of general purpose credit cards: raw data by deciles and marginal effect based on regression model



[20] In other words, the interquartile range is defined as moving from the 25th to the 75th percentile. These percentiles in turn are defined based on the observation that out of every 100 representative consumers, 25 have credit limits below $3,700, the 25th percentile, and 25 have credit limits above $40,800, the 75th percentile.

Figure 6 shows the second most predictive of the selected variables: utilization of general purpose credit cards.[21] The graph is structured similarly to above: the small circles represent the raw relationship between utilization and the FWB score, while the solid line shows the marginal effect of utilization keeping credit score, age, and the other selected variables constant. There is a negative relationship between utilization and the FWB score both in the raw data and in the regression analysis. Again, the marginal effect from the regression shows a more muted relationship than the raw data, which is due to the fact that utilization is significantly correlated with credit score. This correlation implies that consumers with a low credit score and high utilization have a lower FWB score. The negative marginal effect means that utilization is negatively related to the FWB score above and beyond its effect through the credit score. Quantitatively, the predicted change as utilization moves across the interquartile range of the distribution from 3 percent (25^{th} percentile) to 55 percent (75^{th} percentile) is a decrease of 1.97 points on the FWB score, with a standard error of 0.36.

[21] General purpose credit cards (called "revolving bank cards" in credit reporting terminology) comprise the majority of credit cards and are general-use cards branded by a financial institution that allow the consumer to revolve a balance. (The card is considered revolving if it allows for balances to be revolved, regardless of whether consumers take advantage of this feature. We do not directly observe in our data whether balances are revolved or not by consumers.) Other large categories of credit cards are retail credit cards branded by and useable at department stores and other retailers and charge cards branded by financial institutions that do not allow the consumer to revolve a balance.

Figure 7: Financial Well-Being Score as a function of the number of revolving accounts: raw data by deciles and marginal effect based on regression model

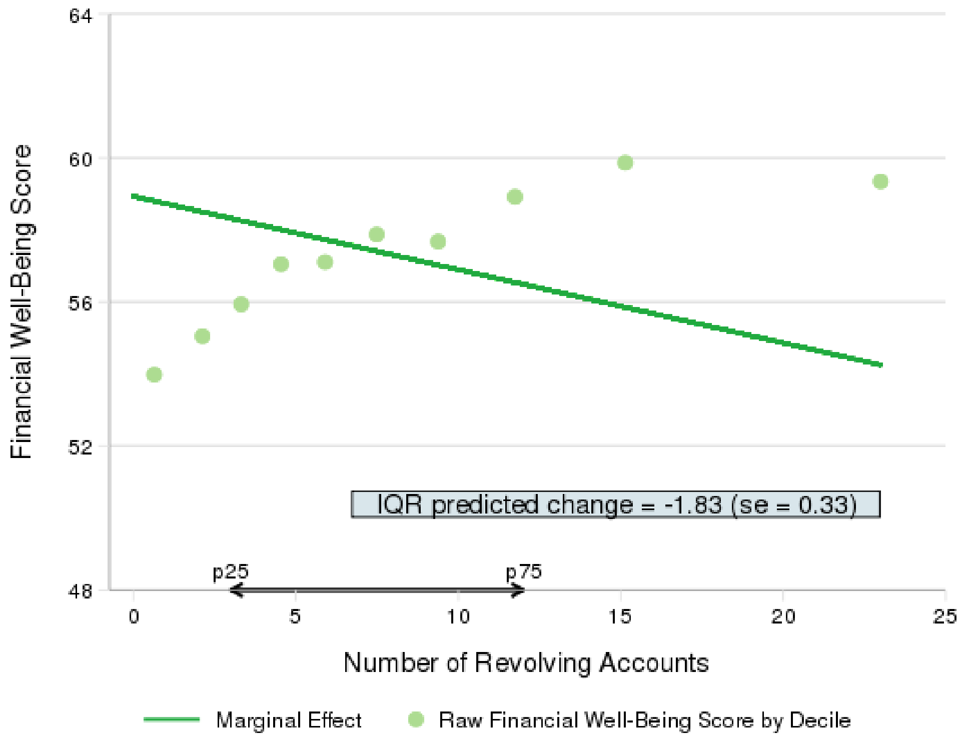


Figure 7 shows the third most predictive of the selected variables: number of revolving accounts.[22] Here the raw relationship between FWB score and the number of revolving accounts is positive: people with more revolving accounts have higher FWB scores. This can once again be explained by the correlation with credit score: consumers with higher credit scores tend to have more revolving accounts. Once credit score is controlled for, however, considering two consumers with the same credit score (and the same values for the other model variables), the one with more revolving accounts is likely have a lower FWB score. In terms of size, the predicted change here is slightly lower: as the number of revolving accounts moves across the interquartile range of the distribution from three accounts (25th percentile) to 12 accounts (75th percentile), the FWB score is expected to decrease by 1.83, with a standard error of 0.33.

[22] Here revolving accounts include both revolving bank card and revolving retail card accounts.

Figure 8: Financial Well-Being Score as a function of the number of collections first reported in the last 24 months: raw data by deciles and marginal effect based on regression model

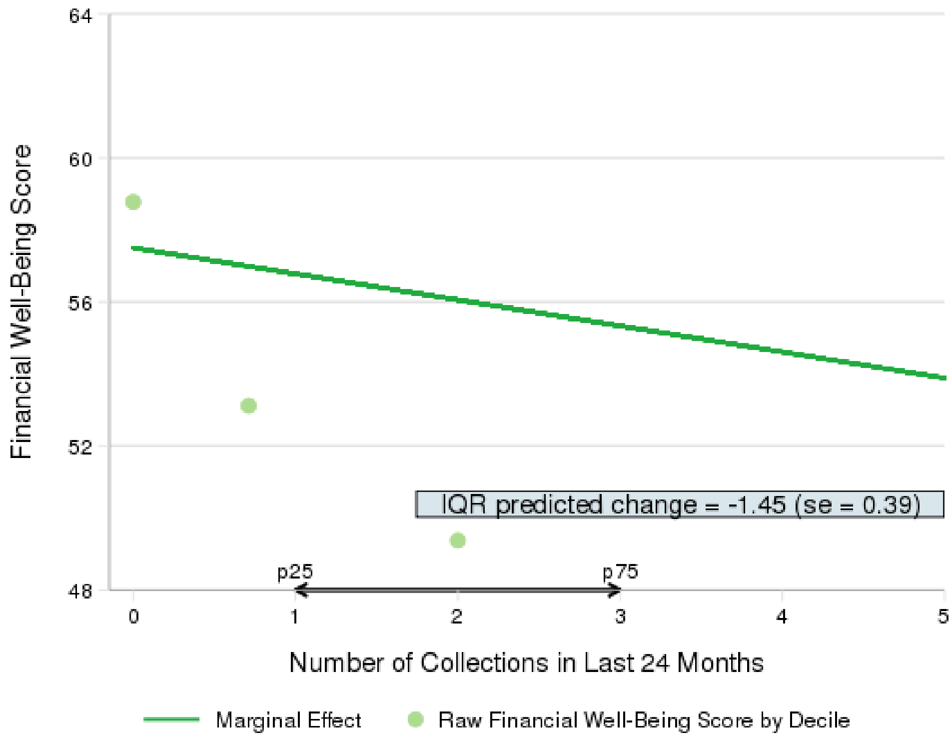


Figure 8 shows the next most predictive of the selected variables: number of collections first reported in the last 24 months. While close to 83 percent of the consumers do not have any collections reported in the preceding two years, 9 percent do have one and 8.5 percent have more than one. Having more collections is strongly negatively associated with a consumer's FWB score even after controlling for credit score and the other variables in the model. As the number of collections moves across the interquartile range of the distribution from one to three collections (among those with a collection), the FWB score is expected to decrease by 1.45, with a standard error of 0.39.

Figure 9 shows a variable that captures whether a consumer has a general purpose credit card. This variable can take on only the value 0 (corresponding to not having a general purpose credit card) or 1 (corresponding to having a general purpose credit card). This means that there are only two small circles on the graph, corresponding to the average FWB score for those who do not and those who do have a general purpose credit card (51.3 and 58.2, respectively). Again, some of the difference is muted when controlling for credit score and the other model variables. The two diamonds on the graph correspond to these marginal effects. Going from not having a general purpose credit card to having a general purpose credit card increases the FWB score by 2.48, with a standard error of 0.72. Notice that this is not directly comparable to the earlier marginal effects for continuous variables where a change equivalent to the interquartile range was considered.

Figure 9: Financial Well-Being Score as a function of whether consumer has general purpose credit card: raw data by category and marginal effect based on regression model

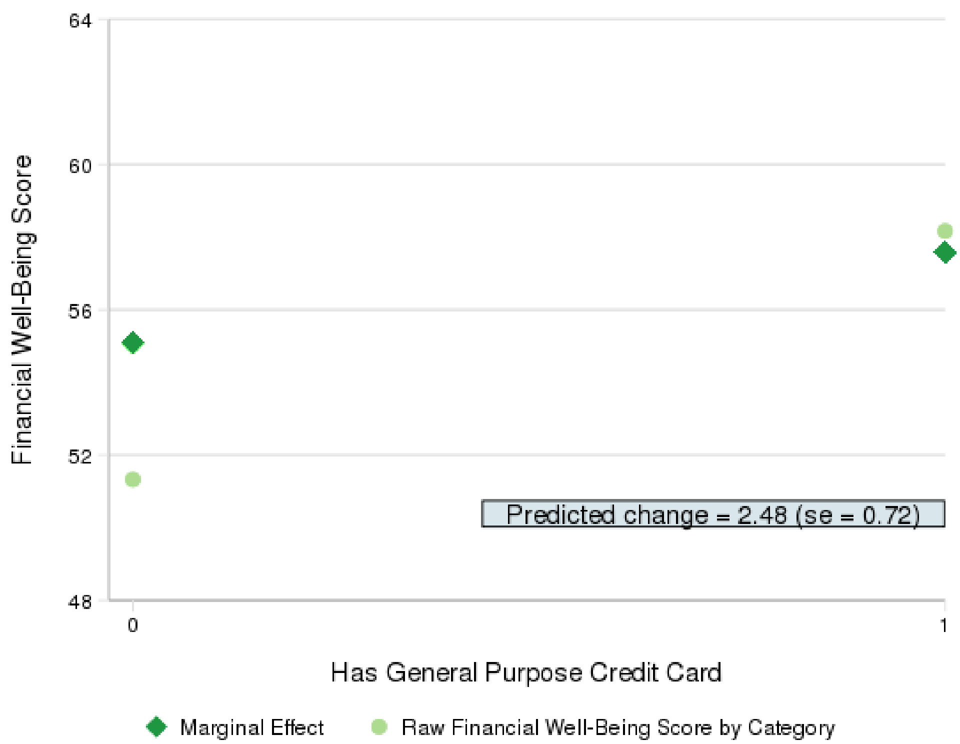


Figure 10: Financial Well-Being Score as a function of number of accounts opened in last 12 months with a balance: raw data by decile and marginal effect based on regression model

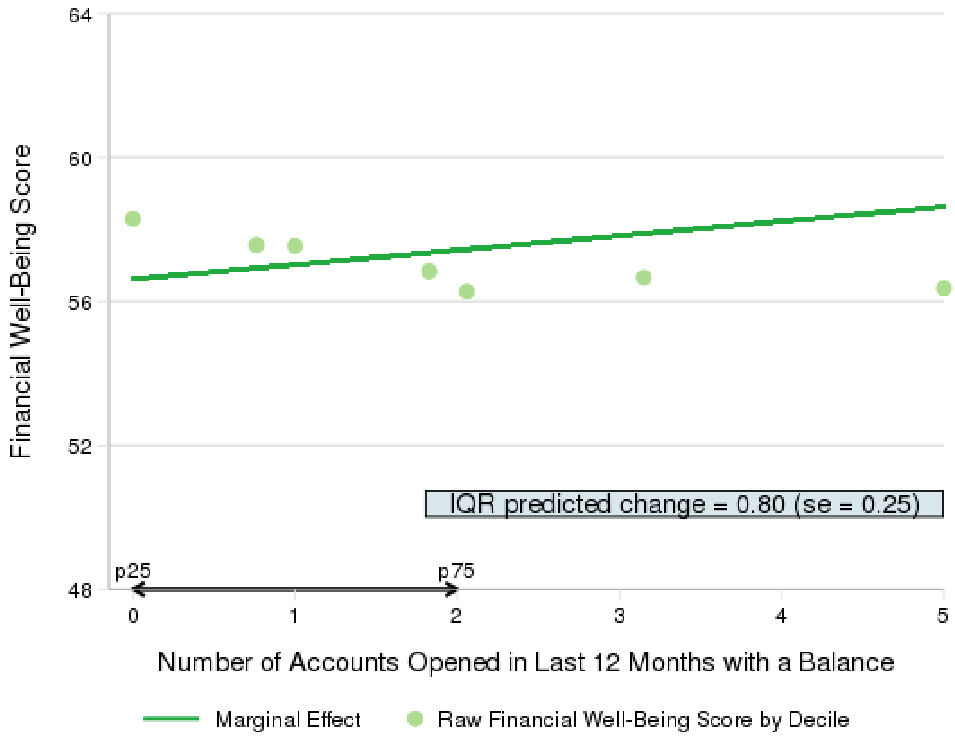


Figure 10 shows the relationship between the FWB score and the number of accounts that the consumer opened in the last 12 months that carry a balance.[23] Here the raw relationship between the FWB score and the number of such accounts is not very strong, but the regression analysis uncovers a somewhat positive conditional correlation: as the number of accounts moves across the interquartile range of the distribution, the FWB score is expected to increase by 0.80, with a standard error of 0.25. As before, this need not imply a causal relationship, and could be due to selection on variables (such as income) that are not observed in our study.

Finally, Figure 11 depicts the relationship between having a student loan and the FWB score. This again is a variable that can take on only the value 0 (corresponding to not having any student loans) or 1 (corresponding to having some student loans). The corresponding two small circles on the graph show that the average FWB score for those who do not have a student loan is 58.3 and 54.6 for those who do have student loans. Of course, consumers with student loans are different from consumers without such loans (notably, they are younger on average). The regression model uncovers that the marginal effect of having a student loan versus not having

[23] In some figures there are fewer than ten points representing the deciles. This happens when multiple deciles take on the same value due to the concentration of observations on certain values.

one (shown by the diamonds in Figure 11) is a decrease in the FWB score of 0.98 (with a standard error of 0.46), where this effect controls for the other model variables, including age. Again, this is not directly comparable to the earlier marginal effects for continuous variables.

Figure 11: Financial Well-Being Score as a function of whether consumer has student loans: raw data by category and marginal effect based on regression model.

4.2 Engagement Variables

In this section, important engagement correlates of the FWB score are identified. Table 2 shows the engagement variables selected by method outlined above, the sign of their conditional correlation with the FWB score, and the marginal effect as each selected variable moves across interquartile range of its distribution.

Table 2: Summary of effect of selected variables on FWB score

Selected variable	Conditional correlation	Interquartile effect
Number of times credit simulator used	+	0.42
Number of times credit factors reviewed	+	0.50
Number of emails opened in last 60 days	-	0.53

The block-wise forward selection procedure highlighted three engagement variables as significant correlates of the FWB score.[24] These correlations are reported in the next three figures using the same methodology as above, i.e., showing the raw correlations and the marginal effect controlling for all other selected variables.

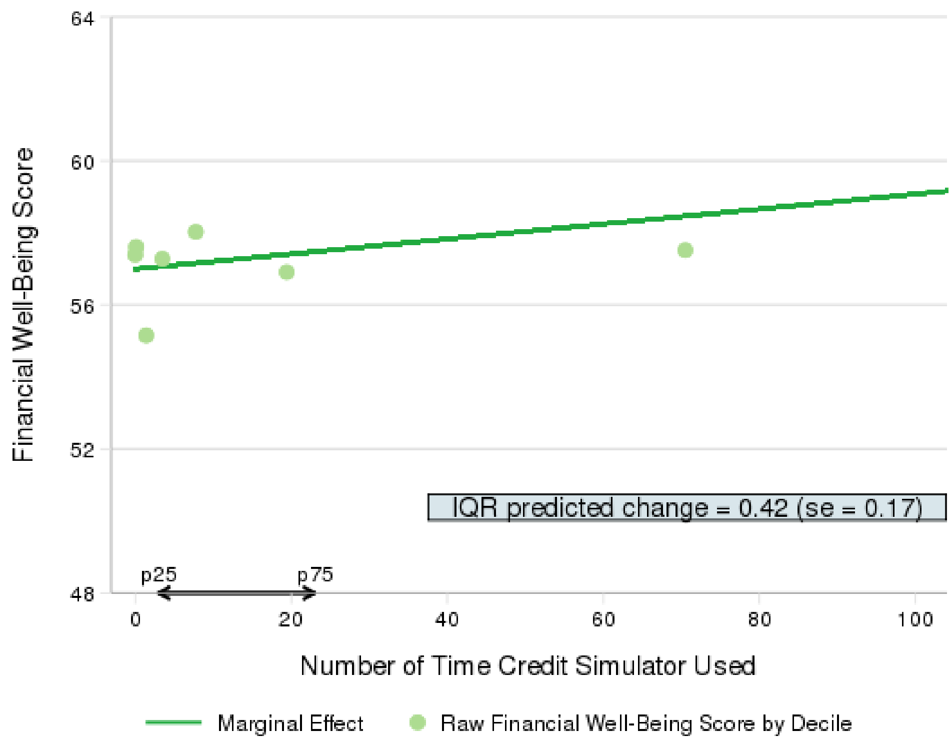
Credit Karma offers a Credit Score Simulator tool to its members. This is an educational tool that starts with the information in the consumer's credit report and explores how changing that information could affect the consumer's score. Figure 12 shows the raw relationship between the number of times a consumer used the credit simulator tool and the FWB score and the marginal effect of using the simulator.[25] There is not a strong relationship between the two measures in the raw data, but the regression analysis uncovers a somewhat positive and statistically significant conditional correlation: as simulator use moves across the interquartile range of the distribution,[26] the FWB score is expected to increase by 0.42 (with a standard error of 0.17). While this change is smaller than that induced by the credit report variables, according to the regression model, it is equivalent to the change associated with a credit limit increase of $7,000, a sizable increase.

[24] See Appendix C for details.

[25] Note that the number of times the simulator was used is top-coded at the top three percentile. There are some very large counts for this variable in the data, most likely accounted for by how website usage and reload is recorded as opposed to actual use.

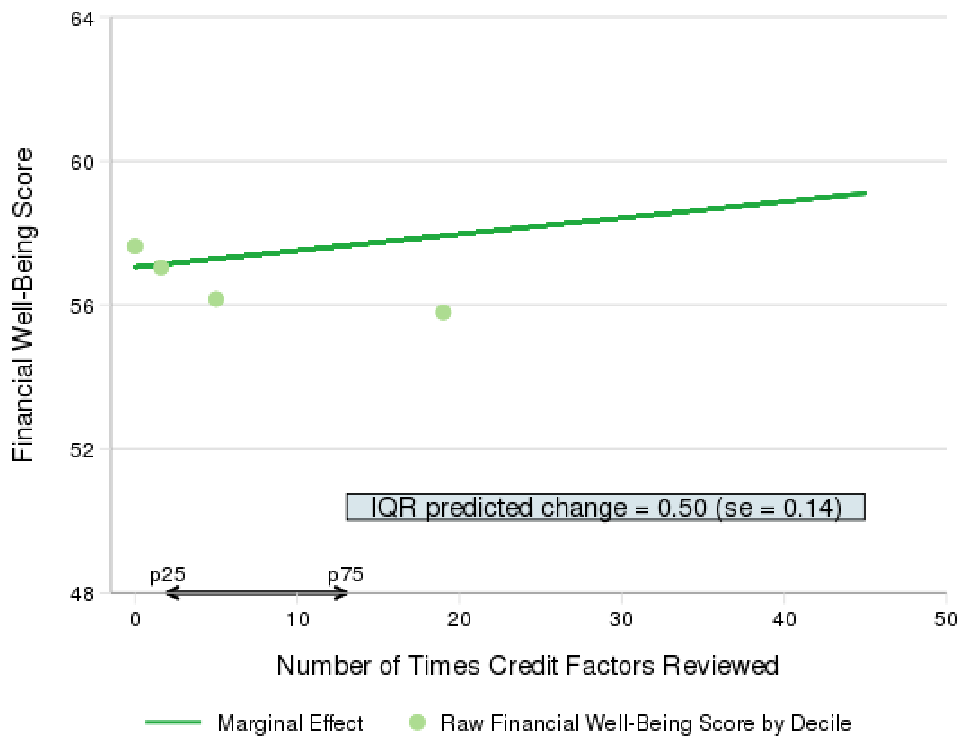
[26] Note that due to the large share of members who do not engage with Credit Karma's tools (47.1 percent for the Credit Score Simulator, 68.6 percent for the credit factors, and 69.8 percent for the email within 60 days), for the engagement variables the interquartile range is defined among members who engage with the respective tool at least once.

Figure 12: Financial Well-Being Score as a function of the number of times the credit simulator is used: raw data by decile and marginal effect based on regression model



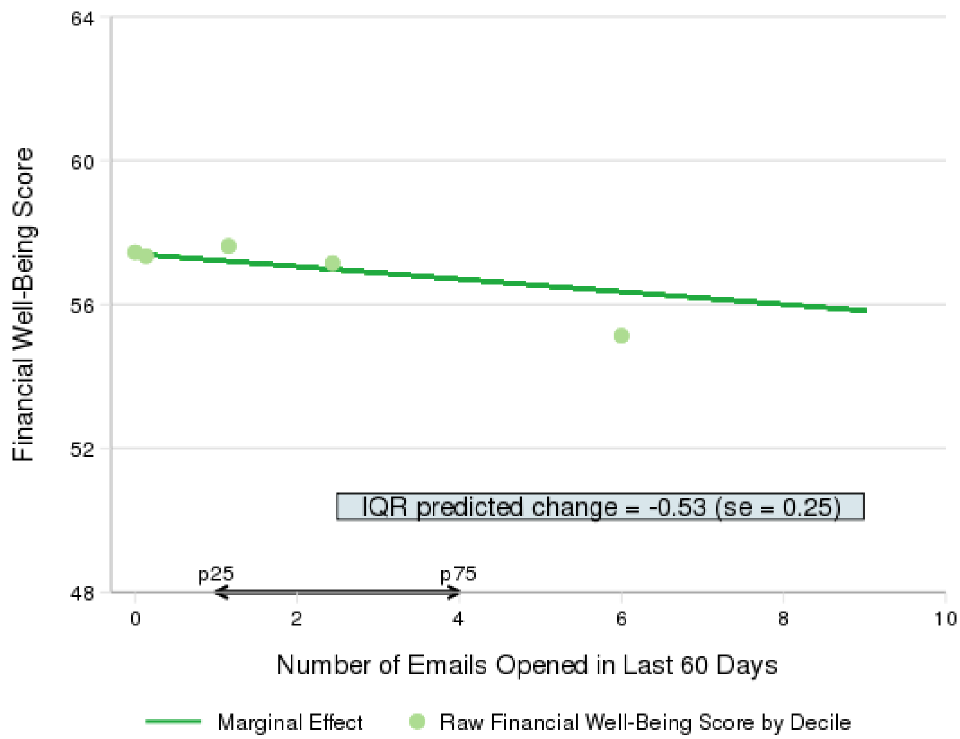
Credit Karma offers another educational tool to its members, one that allows members to better understand the factors that affect credit scores in general and that member's credit score in particular. Figure 13 shows the raw relationship between the number of times a consumer reviewed the credit factors and the FWB score and the marginal effect of reviewing the credit factors. Again, there is not a strong relationship between the two measures in the raw data, but the regression analysis uncovers a statistically significant positive conditional correlation: as credit factor use moves across the interquartile range of the distribution, the FWB score is predicted to increase by 0.50 (with a standard error of 0.14). According to the regression model, it is equivalent to a change associated with a credit limit increase of $8,300, a sizable increase. Again, this is a relationship that deserves further study to determine any possible causality.

Figure 13: Financial Well-Being Score as a function of the number of times credit factors are reviewed: raw data by decile and marginal effect based on regression model



Credit Karma routinely emails their members with information, suggestions, and credit monitoring alerts. Members can opt in to some categories of email and opt out of most categories. As Figure 14 demonstrates, the number of emails opened within the past 60 days is negatively correlated with the FWB score both in the raw data and in the regression model. Choosing to open many emails could be an indication of concern about one's financial situation, an increase in changes on credit reports, or it could imply information overload. Or the relationship could be due to omitted variable bias. Again, the data do not allow researchers to determine the mechanism. Moving across the interquartile range of the opened email count distribution, the FWB score decreases in expectation by 0.53 (with a standard error of 0.25), which is equivalent to the change associated with a credit limit decrease of $8,800.

Figure 14: Financial Well-Being Score as a function of the number of emails opened in last 60 days: raw data by quintile and marginal effect based on regression model



5. Summary and Open Questions

This collaboration has brought together Credit Karma, a personal finance technology company providing free credit scores and reports and credit-related educational tools, and the CFPB to study new survey data to understand the relationship between Financial Well-Being as measured by the FWB score, credit report variables, and variables summarizing engagement with Credit Karma's platforms.

Ten variables were identified– in addition to age and credit score – that correlate strongly with a consumer's Financial Well-Being score. Among variables found on credit reports, credit card limits, holding a credit card, and the number of accounts recently opened with a balance all correlate positively with a consumer's FWB score. Credit card utilization, the number of revolving accounts, the number of collections in the past two years, and having a student loan all correlate negatively with a consumer's FWB score.

Among variables characterizing engagement with Credit Karma's platform, a consumer's Financial Well-Being score correlates positively with the number of times the credit simulator was used and the number of times credit factors were reviewed. FWB score correlates negatively with the number of emails opened in the last sixty days.

This report's findings are intriguing and novel. No prior work has studied the relationship between financial well-being and engagement with financial information using administrative engagement data. Only one other exploratory project linked FWB score with credit report data to date, but with a much smaller sample. The observed correlations might be causal, or they might be explained by reverse causality or omitted variables like the propensity to plan. Either way, the results are intriguing and warrant further study of these relationships as the CFPB develops its strategy for improving financial capability using the concept of financial well-being.

APPENDIX A:

Financial Well-Being Scale Questions

Table 3: Ten-question version of Financial Well-Being Scale

Questions	Statement	Response Options
How well does this statement describe you or your situation?	1. I could handle a major unexpected expense. 2. I am securing my financial future. 3. Because of my money situation, I feel like I will never have the things I want in life.* 4. I can enjoy life because of the way I'm managing my money. 5. I am just getting by financially.* 6. I am concerned that the money I have or will save won't last.*	• Describes me completely • Describes me very well • Describes me somewhat • Describes me very little • Does not describe me at all
How often does this statement apply to you?	7. Giving a gift for a wedding, birthday or other occasion would put a strain on my finances for the month.* 8. I have money left over at the end of the month. 9. I am behind with my finances.* 10. My finances control my life.*	• Always • Often • Sometimes • Rarely • Never

* More affirmative responses indicate lower levels of financial well-being

APPENDIX B:

Matched versus Unmatched Observations

As reported in the main text, 4,067 Credit Karma customers responded to the financial well-being survey. 2,966 of these customers had matched data on credit reports and engagement. In this appendix, it is examined whether the distribution of observable variables (FWB score and age) are different across the matched and unmatched observations. The kernel density estimates in Figure 15 and 16 indicate that the distributions are very similar. This suggests that any bias from considering only the matched observations in the study is likely small.

Figure 15: Kernel density estimate of distribution of Financial Well-Being Score among unmatched and matched observations

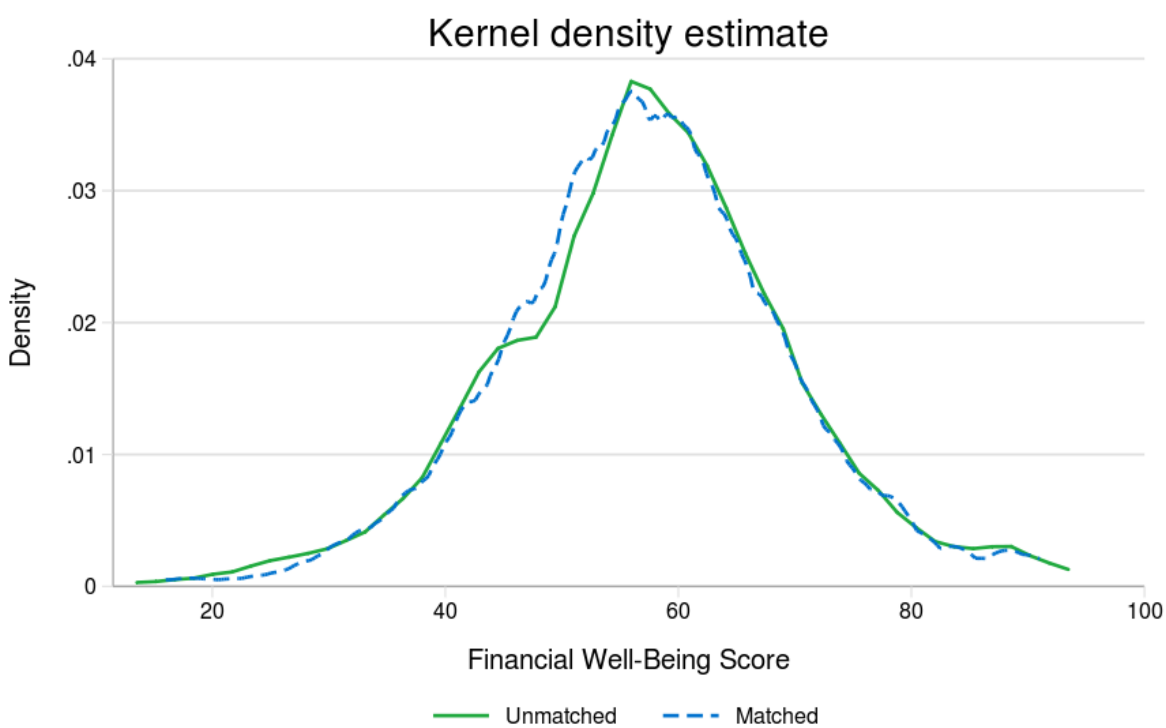
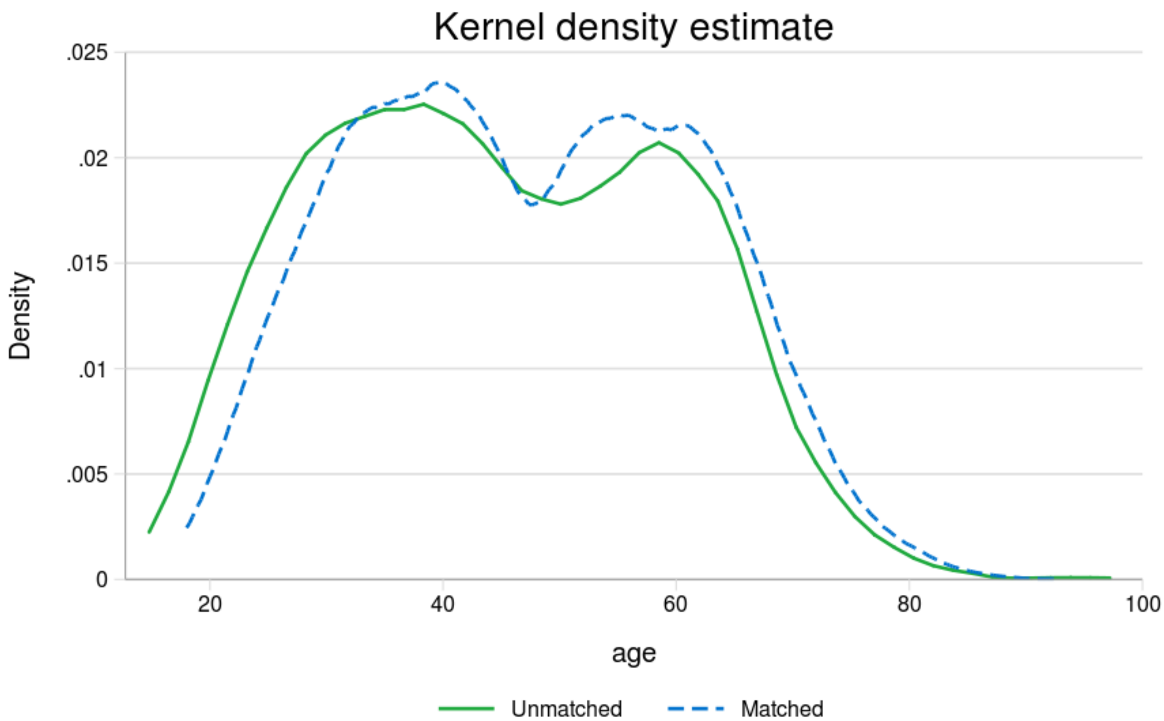


Figure 16: Kernel density estimate of distribution of age among unmatched and matched observations



APPENDIX C:

Supervised Block-Wise Forward Selection Technical Details

Block-wise forward selection with supervision is a transparent method for choosing explanatory variables for exploratory statistical analysis. The objective is not to discern causal relationships, but rather to focus attention on variables that correlate strongly with the main variable of interest (Financial Well-Being score, in this case).

The variables are first preprocessed as follows. Four variables without any within-sample variation and 13 variables with fewer than 1% of observations having non-zero values are dropped. Two variables are dropped due to being too aggregated and one variable due to being very highly correlated with a similar variable. Next, each variable that has a “Not applicable” value is coded as follows: for the observations with a “Not applicable” value originally, a value of 0 is assigned. Then an indicator value is defined which takes on the value of 1 for observations with a “Not applicable” value originally, and 0 otherwise. This allows one to keep all observations in the dataset, but allow for “Not applicable” values to enter flexibly.

First, one credit score among several that were provided is chosen, the VantageScore 3.0 score. Next, credit report variables are considered. From this set of variables, first all performance related variables (various measures of being past due or delinquent) are excluded from consideration. The credit score is a summary variable reflecting expected loan performance, so this exclusion assures that the data are not overfitted (overfitting occurs when a model unintentionally extracts some of the noise as if that variation represented underlying model structure). Then the rest of the credit report variables are partitioned into the following mutually exclusive and collectively exhaustive blocks: mortgage loan and home equity line of credit variables, student loan variables, auto loan variables, installment loan variables, credit card utilization variables, other credit card variables, summary utilization variables, and other summary variables. Two final blocks consist of collection and inquiry related variables, which are excluded from the loan type blocks.

Within each block, the algorithm selects variables using the following process. The algorithm sets the set of provisionally selected variables to the null set. This set of provisionally selected variables is updated by adding explanatory variables as follows, at most one per round. In each round, the algorithm passes through all unselected variables, one by one, and regresses the Financial Well-Being score on age, allowing for a different slope over age 40, the credit score, the existing set of provisionally selected variables, and the candidate explanatory variable. In each regression only the variation between the 25th and 75th percentiles of the candidate variables' distribution is used, conditional on the candidate variables being positive. This is done because it is not desirable to have the observed relationship to be driven only by outliers. A candidate explanatory variable is added to the set of provisionally selected variables if two conditions are met: 1) it is statistically significant at the 90 percent level of confidence and 2) the R^2 from the regression with this variable is higher any other R^2 observed during the current round. If a round does not produce any candidate explanatory variable that satisfies these conditions, then the algorithm stops and the set of provisionally selected variables is finalized. After each round that produces an additional selected variable, any variable from the set of provisionally selected variables that fails to continue to be statistically significant at the 90 percent level of confidence is dropped. After implementing this algorithm within each block, the same algorithm is run, pooling all the block-level selectees using appropriate supervision.

After selecting credit report variables, the engagement variables are selected using a very similar methodology. These variables are partitioned into four blocks, covering recent logon frequencies; responsiveness to offers measured by views and clicks; the duration since joining Credit Karma; and all the remaining engagement measures. The engagement variables are treated differently in two ways. First, while selecting engagement variables both within and between blocks, the selected credit report variables are controlled for. Second, when selecting the engagement variables, attention is not restricted to observations between the 25th and the 75th percentile since there is less inherent variation in these variables, as many consumers do not engage with the available tools. To limit the influence of outliers, here topcoding is used.

Note that the process never selects between credit report and engagement variables simultaneously. Exploration of both these domains is important, novel, and therefore merits positive attention.