Can Facing the Truth Improve Outcomes? Effects of Information in Consumer Finance*

Jessica Fong[†], Megan Hunter[‡]

January 7, 2020

Abstract

Information tracking services, such as Fitbit, MyFitnessPal, and Mint, are gaining in popularity as data becomes more easily available and accessible. These services often attract users through the notion that information can help them reach their goals. However, using data from a consumer finance company where individuals sign up to receive their credit report, we find that retention is the lowest for individuals who presumably would benefit from information the most - those who have low credit scores. This paper explores when individuals change their demand for information and the impact of information on financial health. Specifically, we first document a causal link between credit score trajectories and the demand for information. A decline in credit score decreases the likelihood the individual checks her credit score in the future. Second, we find this decrease in demand for information might be rational. We use variation induced by the firm's email campaigns and A/B tests to instrument for whether an individual checks her credit report in a given month. We find heterogeneous effects of information on credit score. Individuals with a declining credit score prior to checking her credit report experience a decline in credit score as a result of checking. However, there is no significant effect of information on credit score for those who did not have a declining credit score. This finding suggests that encouraging people to access information when they have a declining credit score, and thus a lower demand for information, may worsen their financial health.

*We thank Sam Antill, Andrey Fradkin, Wes Hartmann, Szu-Chi Huang, Sridhar Narayanan, and Navdeep Sahni, along with participants at Stanford Marketing WIP seminar, Advances in Field Experiments 2019 for helpful comments and suggestions. Also we are grateful to the firm for providing the data. All remaining errors are our own.

[†]Ross School of Business, University of Michigan, jyfong@umich.edu

[‡]Graduate School of Business, Stanford University, mkhunter@stanford.edu

1 Introduction

Information tracking services comprise a massive and fast-growing industry; every year tens of millions of people sign up for information tracking services, wearables, and apps. They track their caloric intake and weight on health apps such as MyFitnessPal, record their exercise on trackers like Fitbit or Apple Watch, and regulate their monetary spending and budget through websites like Mint. One of the main motivations for consumers to join one of these platforms is to improve their outcomes, whether it's their weight, fitness, or financial health. In fact, these platforms' advertisements often promote their ability to help consumers reach their goals. For example, the iPhone's activity tracker app ad states: "All you need is a well-built plan, a strong willpower, and the right fitness app to help you stay on track."

While the sign-up rate for these types of information tracking services is high, the retention rate is low. For instance, 42% of consumers who buy a fitness tracker stop using it within 6 months of purchase.[1] This pattern is also prevalent in our setting - a consumer finance platform. Individuals on this platform receive free updated credit reports when they log in each month. These reports include a user's credit score and a report card of the breakdown of their credit score components. The attrition rate for this platform is high - 38% of users do not return after receiving their first credit report. Additionally, the users who never return have significantly lower credit scores than those who do return.

A priori, we would expect that users with lower scores have the greatest room for improvement and may be the most motivated to improve.[2] Therefore, it may be more useful for them to track their progress in order to see which steps they need to take to improve their financial health. Given this, why is the retention rate so low for the individuals who would likely benefit the most from viewing future credit reports? To answer this question, we need to understand the effect of information on outcomes. We explore the following research

[1] https://bigthink.com/ideafeed/users-lose-interest-in-fitness-trackers-after-6-months

[2] The average credit score of users when they first register with this website is 599, which is lower than the national average of 673.

questions. First, what is the impact of the content of information (positive or negative information) on a user's demand for information in the future? Second, is information helpful to these low-demand users, or are they being rational in avoiding it? In other words, what is the impact of information on outcomes, which in this setting is credit scores, especially for those who are more likely to avoid information.

We draw on the information avoidance literature to understand how information impacts an individual's future demand for information. This literature documents that individuals have a tendency to avoid information that is potentially unwanted, such as declining stock markets and positive HIV tests (Golman, Hagmann, and Loewenstein, 2017; Karlsson, Loewenstein, and Seppi, 2009; Lyter et al., 1987). Extending this to the effect of learning the unwanted information on future demand for information, we hypothesize that low retention rates can be attributed in part to individuals lowering their demand for information when they receive "bad news" - a declining credit score.[3]

There are two main challenges in measuring this causal effect. The first is selection bias; individuals who have decreasing credit scores may be less financially responsible, and are less likely to check their financial information. Thus, we first show that, controlling for individual-level differences, when an individual experiences a decrease in her credit score, she is less likely to view her updated credit report in the next month, even for users with "good" credit scores (greater than 660). While this controls for selection bias, there may be shocks that are unobserved to the econometrician that affect both the user's financial status and her valuation of time, making it less likely that she logs in to her account. To overcome this identification challenge, we take advantage of the well-documented "left digit bias" phenomenon, in which people are more likely to focus on the left-most digits of a number, making numbers that end in 0 seem much greater than numbers that end in 9 (Thomas and Morwitz, 2005). When comparing users who have credit scores only one point

[3]While unwanted information could encompass more than just a declining credit score, such as a small increase or no change despite effort to improve one's score, we cannot observe these efforts and thus simplify to this definition as that is what is mostly easily observed in the data.

apart, we find that users who have credit scores that end in 99 are less likely to check their credit report in the following month than users who have credit scores that are only 1 point higher, but end in 00. Therefore, this suggests that a decrease in credit scores causes individuals to be less likely to seek out their financial information in the future.

Given that users are lowering their demand for information, particularly after receiving bad news, we then explore if consumers are being rational in doing so. How would information impact their future credit scores? The observed variation in whether an individual checks her score is not sufficient to answer this question due to the aforementioned selection bias. To obtain exogenous variation needed to measure the causal effect of information on financial outcomes, we utilize the firm's email campaign experiments. For many of their email campaigns, the firm runs A/B tests on different versions of emails to determine the most effective copy. Due to the variation in the effectiveness of emails used in A/B tests, we use the email copy the user receives as an instrument for whether she checks her information. While which email a user receives is generally endogenous to her credit reports and website usage, we select the subset of email campaigns that were involved in A/B tests for which we observe the firm's targeting rules. After controlling for the campaign's targeting rules, the email copy the user receives is exogenous to a her credit reports and website usage.

We find a heterogenous effect of information on credit scores. Users with either a flat or increasing trend in credit scores prior to checking their information experienced a positive, but not statistically significant, change in their credit score. However, within this group of users, those with higher credit scores have a statistically significant increase in their credit scores. Users who had a declining credit score prior to checking their credit report experience a 24 point decrease in credit score on average. This average effect seems to be mainly driven by users who had credit scores in the bottom tercile of the sample (lower than 581) prior to receiving an email, implying that a decrease in the demand for information may be rational. We find that the decrease in credit score for these users is driven by increased spending and more late or missing payments after receiving information.

Our first contribution is to document when individuals change their demand for information in the context of consumer finance. The demand for information relates closely to the information avoidance literature, and more specifically, passive information avoidance. Passive information avoidance, which occurs when individuals do not actively take steps to reveal information, has been documented in many contexts. For example, Lyter et al. (1987) surveyed men with high risk for HIV and found that 80% did not want to conduct a HIV test for fear of receiving a positive result. In finance, Karlsson, Loewenstein, and Seppi (2009) show that stock-holders avoid checking their portfolios when the stock market goes down, and 401(k) investors with larger holdings are more likely to avoid information about their holdings (Sicherman et al., 2015). In more recent work, Huang (2018) finds that people actively avoid information about their peers when they are pursuing a goal, and Masatlioglu, Orhun, and Raymond (2017) show that individual preferences for different types of information (positively or negatively skewed) can mitigate information avoidance.[4] We add to this literature by demonstrating the causal effect of an individual's trajectory and her demand for information in a field setting with financial health outcomes.

We further contribute to this branch of literature by demonstrating the implications of passive information avoidance for outcomes, as well as showing under what situations avoidance can have a positive versus negative impact. In some cases, passive information avoidance may be counterproductive, as information may be helpful in decision-making. Golman, Hagmann, and Loewenstein (2017) discuss potential negative implications of information avoidance on the spread of disease, groupthink, and media bias. However, this paper finds that for users with declining credit scores, information avoidance may improve their credit score and thus be the rational decision. A few papers consider the positive effects of information avoidance. Sicherman et al. (2015) suggest that when the stock market dips, information avoidance may help prevent investors from panicking and making rash decisions. In a lab study, Huck, Szech, and Wenner (2018) show that subjects who completed a task at a randomized hourly

[4]See Sweeny et al. (2010) and Golman, Hagmann, and Loewenstein (2017) for comprehensive reviews of information avoidance across many different contexts.

rate, strategically avoid information on their hourly rate because they may be demotivated by learning that they were selected to receive a low rate. This also relates to the growing literature on selective attention in consumer finance. Stango and Zinman (2009), Medina (2017), Karlan et al. (2016), and Liu, Montgomery, and Srinivasan (2018), among many others, show that reminders in the contexts of overdrafts and savings accounts can reduce overdrafts and increase savings, and have heterogeneous effects. We contribute to this literature by showing the causal, heterogeneous effect of information, where we use reminders to create exogenous variation in receiving the information, on credit scores.

The remainder of the paper is structured as follows. In Section 2, we provide background on the empirical setting of this paper and summary statistics of the types of users on this platform. Section 3 presents evidence for the varying demand for information. Section 4 explains the design for the causal analysis. In Section 5, we measure the causal effect of information, and Section 6 concludes.

2 Data

Empirical Setting

The empirical setting of this paper is a financial monitoring website with more than 15 million users in the United States. Through the website, users can check their credit scores for free. A secondary service that the firm provides is matching users to loans and credit cards; the website suggests credit cards, loans, and other credit lines to users, given their credit history and account information.[5] The site also provides fraud monitoring and alerts users if they suspect fraud on their financial accounts.

When a user creates an account, all of her credit lines are automatically added to the platform. This includes credit cards, mortgages, auto loans, student loans, etc. We can observe

[5]Unfortunately, we do not know how other actions on the site affect a user's score. We leave this to future work.

how many accounts a user has, their required monthly payments, whether the payments are on time, and her credit report. When the user logs in, she sees her credit score, which is updated on a monthly basis. The firm monetizes through a freemium model, where users can pay to upgrade to the premium account, in which they can update their credit scores weekly, rather than monthly.[6] The remainder of this paper focuses on free users.

Summary Statistics

To provide more background about the types of individuals who use this site, this section provides summary statistics on a random sample of 959,564 users who have logged in to their account between January 1, 2016 and June 12, 2018 (the maximum observed date in our data).[7] In this sample, the average credit score is 598.7 with a standard deviation of 90.8. Figure 1 displays a histogram of the average credit score for this sample of users. These

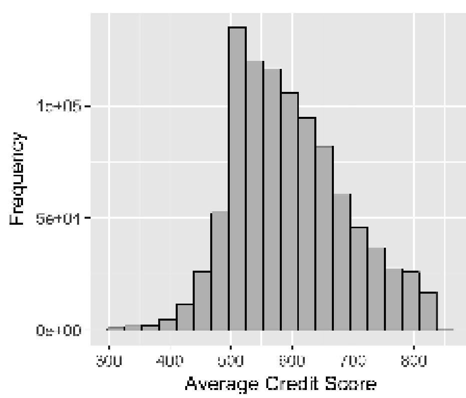

Figure 1: Histogram of the average credit score of the random sample of 1 million users.

individuals, on average, have a lower credit score than the 2018 national average of 673, as measured by Vantage.[8]

For free users, credit scores are updated once a month. More precisely, they are updated the first time that a user logs into her account in a calendar month. For example, if a user

[6] About 1% of the users are premium users.

[7] The original sample is 1 million users, but users who have invalid credit scores (below 300 or above 850) were removed from this sample.

[8] https://www.valuepenguin.com/average-credit-score

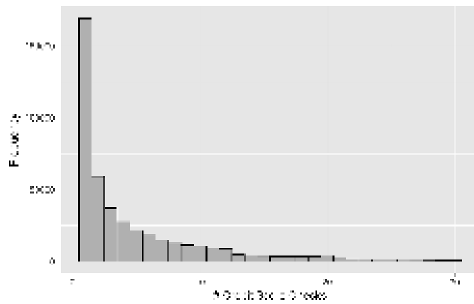

Figure 2: Histogram of the number of times a user views his or her credit score.

creates an account on June 30, logs in again on July 1 and July 15, her credit report will be updated on June 30, and July 1, but not July 15. Figure 2 displays the number of times a user's credit report is updated in the observed time period. A large proportion of users update their credit reports only once, but there is a long right tail.

[Table 1 about here.]

Table 1 provides more detailed information on how frequently a user updates her credit report and how her credit score changes over time. 62% of users pulls her updated credit report more than once. For these users, the average user logs in and updates her credit report every 1-2 months. Note that this includes individuals who log in for two months only and never return.

Since most individuals log in in consecutive months, changes in credit scores are small but on average, increasing, with an average of a 1 point increase in credit score between updated reports. While the average change is small, the variance is relatively high; 50% of all consecutive changes in credit scores between updated reports are between -6 and +10, and the other 50% experience even larger changes. Additionally, in a given three-month time period, about one in four people has a 20 point change in their credit score. These large changes in credit score may have resulted from actions that quickly impact scores, such as adding credit lines, increasing the amount of credit allowed, reducing late payments, and decreasing debt.[9]

[9] According to https://budgeting.thenest.com/fast-can-increase-fico-score-21042.html

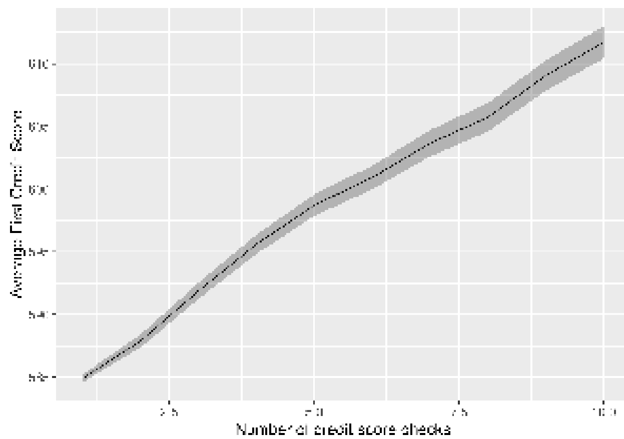

Figure 3: The first score pulled compared to how many times a user pulls their score.

Lastly, we explore what types of users have the lowest retention rate. As this firm advertises its services as a tool to help users improve their credit scores, users who have lower credit scores may have the largest incentive to continuing using this website and viewing their credit report. However, we find that these users are the ones who are most likely to drop off. Users who have lower credit scores are less likely to check their credit report again in the future, as shown by Figure 3, which displays the average first observed credit score, grouped by how many times they pulled an updated credit report during the observed time period. For instance, for users who have updated their credit report 10 times, their average credit score when first joining the website was 604, while users who viewed their credit report only once have an average credit score of 585.

Email Summary Statistics

[Table 2 about here.]

Our identification strategy relies on emails that the user receives, so we provide summary statistics about the frequency and user engagement with emails, which are shown in Table 2. There is a substantial group of users who receive no emails, as they have opted out of any email communication. On average, a user receives 28 emails per month, and opens 25% of them. Conditional on opening an email, a user clicks on 21%.

3 Evidence for Change in the Demand For Information with Credit Score

In the previous section, we document that users with low credit scores are the least likely to return to the website and view their updated credit score in the following month. While this pattern might occur because users with lower credit scores tend to be less financially responsible and thus less likely to continue monitoring their credit reports, users with low credit scores may also be avoiding information. In other words, the type of information, such as unwanted information ("bad news"), may affect their future demand for information. Information avoidance has been been defined as a behavior that is intended to prevent or delay the acquisition of available but *potentially* unwanted information (Sweeny et al., 2010), but in this paper, we measure the effects of *confirmed* unwanted information on the future demand for information. In this setting, we defined "unwanted information" as a decreasing credit score and measure its effect on whether the user views her updated credit report in the future.

Credit Score Trajectories and Information Avoidance

The correlation between credit score and whether the individual checks her updated report may be driven by selection bias. To for selection bias, we measure the effect of a decreasing credit score and whether the individual updates her credit report in the next month, controlling for individual-level effects. We estimate the following logistic regression.

$$\begin{aligned}\mathbb{1}\{CheckNextMonth\}_{it} = \delta_1 \mathbb{1}\{DecreasingScore\}_{it} + \delta_2 |\Delta_{it}^{cs}| + \\ \delta_3 \mathbb{1}\{DecreasingScore\}_{it} * |\Delta_{it}^{cs}| + \delta_4 NChecks_{it} + \nu_i + \epsilon_{it} \quad (1)\end{aligned}$$

Each observation is at a user-month level, where the month t is when the user checks her credit score. $\mathbb{1}\{CheckNextMonth\}_{it}$ is an indicator for whether i updates her credit report in the next consecutive month after t. ν_i is the individual random effect. $\mathbb{1}\{DecreasingScore\}_{it}$ is an indicator for whether the user's score has decreased between $t-1$ and t ($cs_{it} - cs_{i,t-1} < 0$), and $|\Delta_{it}^{cs}|$ is the absolute change in i's credit score between time $t-1$ and t ($|\Delta_{it}^{cs}| = |cs_{it} - cs_{i,t-1}|$). Note that t is the month when a user checks her credit score, rather than consecutive calendar months. If a user checks her credit score in January and then in March, the difference in her credit score is between her credit score in March and her credit score in January. Again, we control for how many times the user has checked her score in the past. The absolute change in credit score is divided by 100, so a 1 unit increase in $|\Delta_{it}^{cs}|$ is interpreted as a 100-point change in credit score.

Table 3 displays the logit results of Equation 1. First we note that users who are on a declining trend are less likely to check their scores the next month. The log odds of checking a score decreases by 0.022 if a user is on a declining trend. For users on a positive trend, the larger the increase, the more likely they are to log in the next month. A 100 point increase in credit score translates to a 0.013 absolute increase in log odds of checking next month.

[Table 3 about here.]

This approach, however, is still subject to endogeneity, as there may be unobserved shocks that simultaneously affect the individual's credit score and her time, reducing the likelihood she checks her information in the next month. Thus, in the following section, we provide another test for information avoidance by leveraging left-digit bias.

Leveraging Left Digit Bias

Left digit bias is the tendency for individuals to pay more attention to the leftmost digits and less attention to the rightmost digits when evaluating a number. This bias has been documented and tested in a variety of settings, from prices (Anderson and Simester, 2003; Shlain, 2018; Thomas and Morwitz, 2005) to car mileage (Lacetera, Pope, and Sydnor, 2012).

The general finding in this literature is there is a greater increase in demand between prices that end in \$9 than prices that end in \$0, compared other \$1 differences with other digits. The mechanism is that consumers perceive the difference between 99 and 100 to be much greater than the difference between 97 and 98. We leverage this bias to measure the causal effect of credit score on whether the individual avoids information, with the assumption that users who have 1 point difference in their credit scores are identical on unobservables.

The identification strategy is as follows. Due to the left digit bias, an individual would perceive a credit score of 399 to be much lower than 400, while the perceived difference between 397 and 398 is not as stark, even though both differ by 1 point. Thus, if an individual is more likely to avoid information when they are doing worse, individuals who have a credit score of 400 would be more likely to check their credit score in the next month, compared to those with a credit score of 399. Because credit scores 399 and 400 are close together, we expect the individuals to be very similar in other respects.[10]

To estimate this effect, we select observations (at the user, credit report level) where the credit score ends in either 99 or 00 for all thresholds, including credit scores of 399, 400, 499, 500,... up to 800. For this sample, we estimate the following logistic regression.

$$\mathbb{1}\{Check\}_{i,t+1} = \alpha_{RoundedCS} + \gamma_1 \mathbb{1}\{EndsIn00\}_{it} + \gamma_2 \mathbb{1}\{EndsIn98\}_{it} + \epsilon_{it} \quad (2)$$

where $\mathbb{1}\{EndsIn00\}_{it}$ is the indicator variable for whether i's credit score at t ends in 00. $\mathbb{1}\{EndsIn98\}_{it}$ is an indicator for ending in 98 and is included as a placebo test. Credit scores that end in 99 are the baseline for this regression. We also control for the threshold with a indicator variable for each credit score, rounded to the nearest 100, $\alpha_{RoundedCS}$. We

[10]Note that for some credit cards and loans there could be credit score cut-offs such that someone with a credit score of 599 may be rejected but someone with a credit score with 600 may be accepted. In this way, the consumers are different in an important aspect. However, since here we are just looking at whether or not a user logs-in based on their score, this issue should not play a large role. It is possible that someone with a credit score of 600 knows they are more likely to be accepted and therefore are more likely to log-in in order to apply for credit cards and loans, this should only be a very small percentage of users, most are logging-in to check their score. Only 0.28% of actions on the website are applications for credit tradelines.

expect that if there is left digit bias, the coefficient for comparing 100 to 99 should be larger in magnitude than the coefficient for 99 to 98, and if there is information avoidance, the coefficient should be positive, as individuals should be more (less) likely to view their updated credit report if their credit score is higher (lower).

[Table 4 about here.]

Table 4 displays the results of Equation 2. The intercept represents the logged odds of checking in the next month if the individual's credit score is 399. The magnitude of $endsin00$ is the difference in the logged odds of checking next month for users with a score that ends in 00 compared to users with a score that ends in 99. The coefficient for $endsin00$ is positive and statistically significant, and larger in magnitude than the coefficient for $endsin98$, meaning the larger the perceived increase in credit score, the more likely the individual is to check her score in the next month.[11] We also find that this effect holds across credit score groups of 100s as shown in Table 10 in the Appendix.

4 Causal Impact of Receiving Information

In this section, we explore whether users are rational in decreasing their demand for information after receiving unwanted information. Specifically, we measure the heterogeneous effect of information on their future credit score for users who receive positive (a non-decreasing credit score) or negative (decreasing credit score) information. In other words, how does receiving information change their outcomes, especially for users who were avoiding it? To measure the causal impact of information on her future credit score, we need exogenous variation in whether a user logs in and receives an updated credit report. One of the ways that

[11] One concern with this identification strategy is that users may be more likely to log in and use the platform to find loans if their credit score is greater than a loan-qualifying threshold. However, this is unlikely in this case, as most lenders use the FICO credit score, which is slightly different from the Vantage credit score. Also, the only cutoff for mortgages that ends in 00 is 700. The results are similar after excluding 700 from the analysis.

the firm can influence the user to log into her account is through the emails that they send. This firm often conducts A/B tests with different email versions to find the most effective copy for an email campaign. These email copies generate variation in whether a user logs into her account and checks her credit score. Thus, we use the email copies as an instrument for whether a user logs into her account. The following section describes the emails in more detail and the assumptions needed to use them as a valid instrument.

Overview of Email Experiments

Campaign Selection

Between 2014 and 2017, the firm tested thousands of email campaigns. These email campaigns are mainly reminders for users to log in to the platform. Examples of common messages are "Your credit score has been updated", "We've updated your credit usage grade", "Have you seen your updated approval odds?". To find the most effective email copy to use for the campaign, the firm sends out multiple variants of the message to test groups in the following way. The firm first selects which users will receive a campaign based on a set of targeting rules. From within those targeted users, they randomize a proportion of users to be in the email experiment. Then they randomize these users into groups and each group receives a different email copy. For example, if the email campaign's message is to let users know their credit score has been updated, email variants within the campaign may have subject lines like "We have updated your credit score", or "Check your updated credit score", or the variants may be sent out at different times of day. After the email experiment, the firm then selects the email copy that has the highest open or click rate to send to the remainder of the users who were selected for the campaign but not the experiment.

To maximize the variation in website logins resulting from the emails, we use data from the top 25 campaigns that had the highest variation in click rates across email copies. We received targeting rules from the firm for 18 of the 25 campaigns. Additionally, there are some emails with text may violate the exclusion restriction, such as subject lines contain information about how the user is doing. For example, one of the subject lines is "your credit usage may be affecting your score". Another email's subject line directly tells the user how many points her credit score has increased. Since it is possible that the text of these subject lines may directly impact a user's credit score outside of encouraging them to check their score, we remove these types of emails from the data set.[12] The final data set for this analysis has 16 campaigns and Table 5 displays the subject lines for all email variants in these campaigns. There is a total of 70 different email copies, and 531,232 users on the platform received at least one of these email copies.

[Table 5 about here.]

Summary Statistics of Selected Sample

Users who received these email campaigns are those who did not opt out of receiving emails. The difference between this sample of users from the average user is reflected in the following summary statistics. The average credit score of the users who received these email campaigns have an average credit score of 631.4, with a standard deviation of 91.1, which is about 33 points greater than the overall average credit score of the randomly selected sample of all users. The histogram of the average credit score for each of these users is reported in Figure 9 in the Appendix. This sample of users also check their credit score more frequently than the random sample; there is a larger mass of users who check their updated credit score each month. In addition, the average user in this sample receives 10 more emails than the average user from the entire population, due to the fact that these users have not turned off their email subscriptions. More detailed summary statistics are included in the Appendix.

Although the users in this sample are different from the average user, they exhibit similar information avoidance behavior. Similar to the total population, these users are also less

[12] Specifically, we remove the email with the subject line "Your credit increased X points since you signed up" and all emails with "affecting" or "points increased" in the subject line.

likely to check their credit score in the future if they currently have a decreasing credit score, and more likely to check their score when it's on an increasing trend. The magnitude of this effect is very similar to the random sample. Table 6 displays the estimates of Equation 1 for the sample of users who receive the selected email campaigns.

[Table 6 about here.]

However, it is important to keep in mind that that these selected users are not representative of the entire population, and this is a limitation of this dataset.

Another limitation is that we do not observe credit scores of individuals who do not log in that month. We can only observe changes in credit score of those users who log in after receiving the email at some point in the time window of our data. We choose emails that occurred at least a year before the end of our time window to allow time for users to return. Since those users who do log in may be very different from those who do not log in, the implications of our results are limited to the self-selected sample of individuals who check their updated credit reports at least twice.

While we cannot extend our results to users who do not return after receiving an email, we can check users who receive positive or negative information have different retention rates, regardless of whether they check their information. While the users with decreasing and non-decreasing trends may be inherently different, we do not find that users in one group are disproportionately more likely to drop out than the other. The drop-out rates for these two groups are similar, with the positive information group having a 20.8% drop out rate and the negative information group having a 19.4% drop out rate.[13]

[13]Note that these drop out rates are statistically significantly different, but not meaningfully so. The 95% confidence interval of the difference between the two proportions is (0.004, 0.007) which is not very economically meaningful. In addition, since the positive pre-trend group has a larger drop-out rate than the negative pre-trend group, we are not concerned that we are losing a disproportionate number of users due to traditional information avoidance.

Email Campaigns as Instruments

Variable Creation and Timeline of User Events

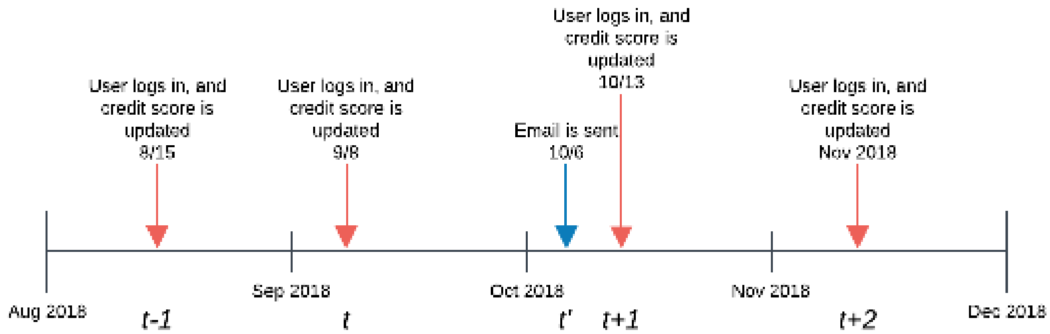

Figure 4: Timeline of user behavior.

To measure the effect of a user logging in and checking her credit report on how her credit score changes after checking, we construct "pre-trend" and "post-trend" variables. Figure 4 depicts a sample user's history to better illustrate how these variables are created. In terms of this figure, we want to measure how checking her credit score at $t+1$ affects how her credit score changes at $t+2$, and how that effect differs for users who are most likely to engage in information avoidance (users who have declining credit scores). We construct $pre\_trend_{i,t+1}$ to be the trend in i's credit score before checking her score at $t+1$.

$$pre\_trend_{i,t+1} = cs_{i,t} - cs_{i,t-1}$$

To measure how checking at $t+1$ affects how her credit score changes at $t+2$, we create the following variable:

$$post\_trend_{i,t+1} = cs_{i,t+2} - cs_{i,t}$$

where this is the change in credit score before and after receiving the email. In the absence of endogeneity bias, we would estimate the following regression using OLS to measure the effect of checking the credit score on future changes in credit score.

$$post\_trend_{i,t+1} = \alpha + \beta_1 \mathbb{1}\{Check_{i,t+1}\} + \beta_2 \mathbb{1}\{Check_{i,t+1}\} \times \mathbb{1}\{pre\_trend_{i,t+1} < 0\} + \beta_3 \mathbb{1}\{pre\_trend_{i,t+1} < 0\} + \epsilon_{it} \tag{3}$$

$\mathbb{1}\{Check_{i,t+1}\}$ is an indicator for whether user i checks her credit score at time $t+1$, and $\mathbb{1}\{pre\_trend_{i,t+1} < 0\}$ is an indicator for whether i's credit score was declining before checking her score. The coefficient for the interaction variable of $\mathbb{1}\{Check_{i,t+1}\}$ and $\mathbb{1}\{pre\_trend_{i,t+1} < 0\}$ is the effect of checking for users had declining credit scores prior to checking the email, and thus more likely to avoid checking their report.

However, whether a user checks her credit score in a given month is endogenous to how her credit score changes. For example, users may be more likely to check their report when they believe their credit score is increasing, which would result in a selection bias. Therefore, we use the emails as instruments that shift the likelihood that users check their score. In this analysis, we refer to a user "checking" her credit information as the user logging in and viewing an updated credit report within 1 month (31) days after the email was sent.

Two Stage Least Squares

In an ideal world, the exogenous variation in whether a user checks her credit score in a given month comes solely from which email test group that a user is in, with in a campaign. Conditional on receiving a campaign, whether a user receives email variant 'A' or variant 'B' is randomly assigned. While there is variation in click rates across email variants within a campaign, this variation is not large enough to be a strong instrument on its own.[14] The alternative we adopt is to use each email copy across all campaigns as an instrument, as there is significant variation in click rates in email copies across campaigns. However, a concern is a potential correlation between email campaign click rates and users' credit scores, as email

[14]The average standard deviation in click rates across all email variants within a campaign, averaged over campaigns, is 4.7%

campaign assignment is non-random. To control these correlations, we obtained the email targeting rules the firm used for each campaign in this data set. Conditional on the targeting rules, whether the user receives an email is not correlated with how her credit score changes in the future.[15] With this in mind, we estimate the instrumental variables regression in the following way.

We estimate the regressions separately for users with decreasing and non-decreasing credit scores prior to checking.[16] The first stage of the IV is the following:

$$\mathbb{1}\{Check_{ie}\} = \alpha_e + \gamma X_{ie} + \eta_{ie}$$

where α_e is the fixed effect for email copy e. X_{ie} is a matrix of i's characteristics at time that e was sent. These characteristics are the values used to determine campaign targeting. Since we observe all the characteristics that are used to determine which user receives which campaign, we are able to control for correlations between α_e and η_{ie} through X_{ie}. The first stage regression results are included Table 14 in the Appendix.

The second stage of the regression is

$$CS\_post\_trend_{ie} = \beta_0 + \beta_1 \mathbb{1}\{\hat{Check}_{ie}\} + \Delta X_{ie} + \epsilon_{ie} \tag{4}$$

where $\mathbb{1}\{\hat{Check}_{ie}\}$ is the fitted values from the first stage. Before presenting the estimates, we first provide evidence that email copies are valid and strong instruments.

Instrument Validity and Strength

In order for the email e to be a valid instrument, e must be correlated with whether i logs in and must satisfy the exclusion restriction. Figure 5 shows the histogram of the click rates

[15]Section 7 in the Appendix shows that 1) within each campaign, the email variants are not correlated with the users' credit scores and their trend in credit score, and 2) controlling for campaign targeting rules, campaign fixed effects do not explain additional variation in a user's credit score and trend in credit score.

[16]Estimating the regressions together with an interaction variable $\mathbb{1}\{Check_{i,t+1}\} \times \mathbb{1}\{pre\_trend_{i,t+1} < 0\}$, like in Equation 3 would require additional instruments, as the instrumental variable is also endogenous.

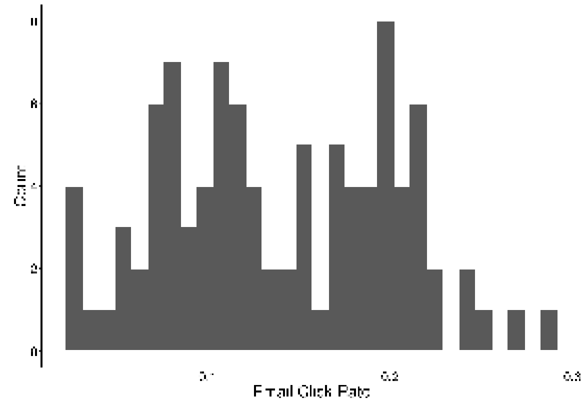

Figure 5: Histogram of the click rates for each email that was sent. This figure shows that there is a lot of variation in how effective different emails and campaigns are, in terms of clicks.

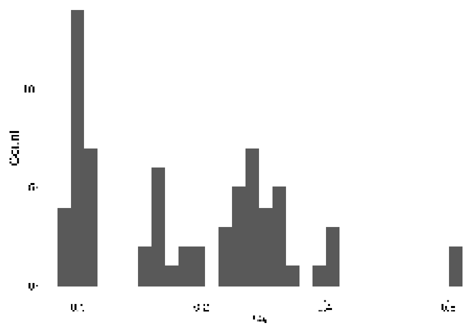

Figure 6: Histogram of the fixed effects of emails on whether the user updates her credit report in the month after receiving the email.

for all the emails in these 16 campaigns. The large variation in the click rates suggests that emails have differing levels of effectiveness. However, a portion of this variation may be driven by the targeting rules; some email campaigns may be sent to users who are more likely to respond to the email. To check whether the email itself induces variation, we regress email fixed effects on whether a user checks her score, while controlling for the targeting rules.

$$\mathbb{1}\{Check_{ie}\} = \alpha_e + \gamma X_{ie} + \epsilon_{it} \tag{5}$$

Figure 6 displays a histogram of α_e in Equation 5. While there are five emails that seem to not be correlated with checking, the majority of campaigns have fairly large correlations on whether a user views her updated credit score in the month after receiving the email.

We believe the instrument satisfies the exclusion restriction due to the text in the email and the frequency at which the firm sends emails. One way the instrument can violate the exclusion restriction is if the email text causes the user to take actions to change her credit score, regardless of whether she checks her report. We believe that this is unlikely in this context, as the emails do not contain prescriptive advice on how to change their credit scores. Another way that emails could violate the exclusion restriction is if they serve as a reminder that their credit score exists. If this were true, then the emails may affect the user's behavior even when they do not check their credit score. While this may happen in other platforms, this firm sends a large number of emails to users who have not unsubscribed. On average, users who have not unsubscribed from the email list receives 31 emails per month, with a median of 21.6 emails per month, as shown in Table 7.

[Table 7 about here.]

Therefore, the focal emails in these campaigns are unlikely to serve as reminders about the existence of their credit score, as the users are also receiving many other emails from the firm.

We test for whether the emails serve as strong instruments by conducting an F-test on the instruments to see if they are jointly significant for each sample (users with decreasing and non-decreasing credit scores prior to checking). The unrestricted and restricted regressions are below, and the regression estimates are reported in Table 13 in the Appendix.

$$\mathbb{1}\{Check_{ie}\} = \alpha_e + \gamma X_{ie} + \eta_{ie}$$

$$\mathbb{1}\{Check_{ie}\} = \gamma X_{ie} + \eta_{ie}$$

The F-statistic on the joint test is 27 for users with decreasing credit scores, 107 for users with non-decreasing credit scores, indicating the emails are not weak instruments for whether the user views her credit information.

5 Results

Effect of checking information on credit score

[Table 8 about here.]

Table 8 displays the two-stage least squares estimates. The first column reports the effect of checking information for users who had declining credit scores prior to checking, and the second column is for users who had either constant or increasing credit scores prior to checking. Standard errors are clustered at the campaign level for all regressions. The coefficients for the campaign control variables are omitted for better readability. The full table, with the campaign control coefficients is in Table 15 in the Appendix. For users with non-decreasing credit score trends before checking, the coefficient for *check_ind* is positive, indicating that she increases her credit score by almost 6 points if she checks her credit score. However, this effect is not statistically significant. On the other hand, checking credit information decreases credit score by 25 points for users who had declining credit scores.

Effects for users with poor, average, and good financial health

The above effects show that on average, credit score information does not help increase credit scores for users who are experiencing a downward trajectory in their credit score. However, is this effect mainly driven by users who have low credit scores to begin with? To test this, we estimate the 2SLS regression for users who had poor, medium, or good credit scores before receiving the email. We classify users into these bins based on the tercile of their credit score. In this data, the 33 and 66 percentile credit scores are 582 and 664, respectively. Users in the "poor" credit score bin have a credit score of below 582 the last time their credit report is pulled prior to receiving the email, "medium" is between 582 and 664, and "good" is greater than 664.

[Table 9 about here.]

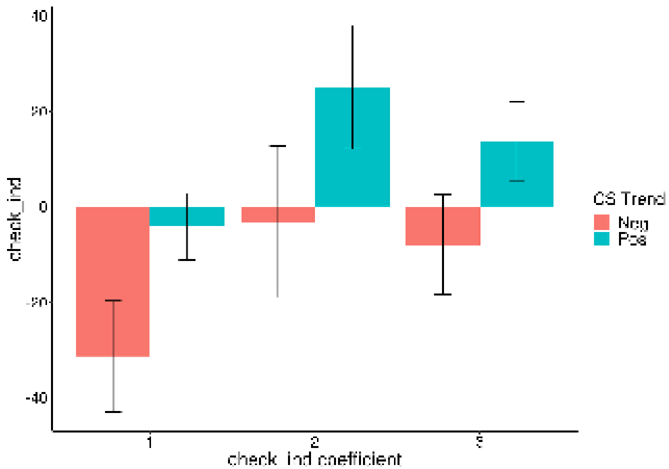

Figure 7: Plot of the coefficients for *check_ind* for users with different credit score trends and in different credit score terciles. The cutoffs for the terciles are 582 and 644. The bars are the 95% confidence intervals.

Figure 7 shows the coefficients of the *check_ind* variable split by tercile and pre-trend status. For all terciles, those with a negative pre-trend have a further drop in score upon checking, however the results are largely driven by the bottom tercile which is the only tercile with a statistically significant drop of over 31 points. We further notice that for users who were in the top two terciles of credit scores, users with a positive pre-trend have a statistically significant boost in their scores upon checking, by 25 and 13 points for the middle and top tercile respectively. Users who are doing badly in general, and were on a declining trend most recently, continue to do worse when they learn their information. However, users who are doing well in general, and most recently were on an increasing trend, benefit from learning their information. We hypothesize that these users might be encouraged by the information and continue to keep up the work they are doing to improve their score.

How Users Change Their Behavior After Checking

Why are credit scores decreasing after information-avoidance-prone individuals check their credit score? In this section, we We use the components of the credit report, such as the total monthly payments and balances and "credit grades", to provide insight on what actions a person takes to change her credit score. A user's total monthly payments includes credit card and loan payments, and her total monthly balance is how much money she owes. "Credit grades", which are only included in some credit reports, are letter grades (A, B, C, D, or F) assigned to different areas of credit that may affect the overall credit score. The different areas of credit considered are payment history, credit usage, credit age, account mix, credit inquiries, and debt to income ratio. We estimate Equation 4 on these outcome variables separately for users with decreasing and non-decreasing credit scores prior to checking. We convert the letter grades to numbers ($A = 5$ and $F = 1$) such that a positive coefficient for $\mathbb{1}\{\hat{Check_{ie}}\}$ indicates a better grade.

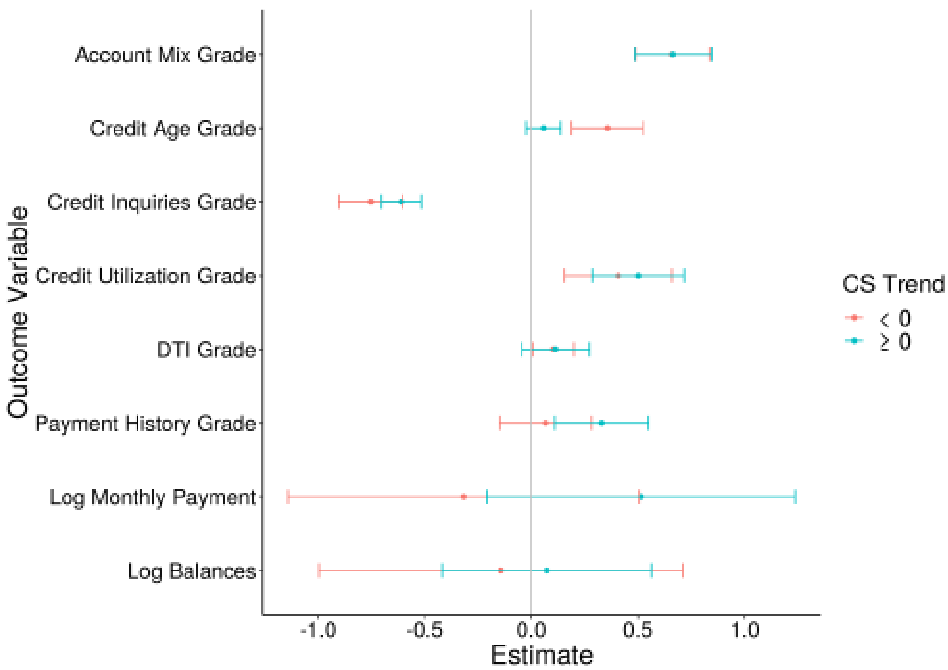

Figure 8: Estimates of the coefficient for $check\_ind_{ie}$ on the listed outcome variables. Standard errors are clustered by the email campaign. Error bars represent the 95% confidence intervals. Users who have monthly payments and balances in the 98+ percentiles are excluded from this analysis. For regressions where the outcome variable is a grade, $N = 24,182$ for declining credit score trends, and $N = 67,922$ for non-decreasing trends. For the monthly payments and balances regression, $N = 194,738$ for declining trends, and $N = 339,964$ for non-decreasing trends.

Figure 8 displays the estimated effect of $\mathbb{1}\{\hat{Check_{ie}}\}$ on the various outcome variables.[17]

[17]Note that many credit reports missing grades and balances. So to maximize the number of observations, the outcome variables in this section are the absolute grades and balances, rather than the change in grades and balances before and after checking.

The top 6 outcome variables are the user's grades. Users who had declining credit scores prior to checking tend to have a better credit age grade than users with non-declining credit scores. An increase in credit age means that users are less likely to open new accounts, thus the average age of their credit accounts increases. However, users with declining scores experience slight decreases in their credit inquiries, credit utilization, and payment history grades, meaning they have more credit inquiries, are using more available credit (getting closer to the credit limits), and have more missing/late payments.

Lastly, users with non-decreasing scores are more likely to increase their total balance, indicating they spend more. However, they increase their monthly payments more than their total balance, meaning they are repaying debt at a faster rate. On the other hand, the total balance is growing at a faster rate than monthly payments for users with declining credit scores. This suggests that the negative effect on credit score from checking information for users most likely to avoid information is due to the increased spending and more late/missing payments.

6 Conclusion

This paper studies the demand for information in the context of consumer finance and estimates the heterogeneous causal effect of information on individuals who are more or less likely to avoid information. Using novel data from a website that provides its users with free monthly credit reports, we observe that individuals who are more likely to churn from the website have lower credit scores. Extending prior literature on information avoidance which focuses on avoiding unwanted information, we study the causal effect of receiving unwanted information in the propensity to seek information. We leverage left-digit bias, with the assumption that people with credit scores that end in 99 and 00 are very similar, in order to measure the causal effect of credit score on checking their information in the next month. We find that individuals with a higher credit score, which ends in 00, are more likely to check their credit report in the next month, compared to individuals with a one point lower credit score, which ends in 99.

To measure the causal effect of checking information on future credit scores, we instrument whether an individual checks her credit score with email experiments that the firm has implemented. The randomization from the email experiments, combined with the targeting rules the firm uses to target campaigns, introduces exogenous variation in an individual's propensity to log in to their account and check their information. We find that individuals who are most likely to avoid information (had a decreasing credit score prior to checking their score) have more late/missing payments and increase their monthly debt and thus experience a significant decrease in their credit score after checking. This suggests that information tracking firms should adjust their targeting rules to incorporate their customers' variable demand for information.

A limitation of this paper is that we measure this effect only on a subset of users due to the way that the firm retrieves the credit reports. Since we can only observe credit reports for users who log in, interpretation of our results is limited to the returning users. In addition, our instrumental variable analysis can only allow us to measure the local average treatment effect. In other words, our estimated effect of checking information is the treatment effect for compliers - people who are able to be persuaded by the email copy to check their information. Lastly, the types of email content used in this paper to nudge individuals is limited - the emails are purely reminders that the individual's credit score has been updated. Thus, we are not able to explore how the content of the emails can affect the effect of checking information. Perhaps including information about peers or placing social pressure can moderate the effect of information on future financial outcomes for information-avoiders. Understanding how to improve financial health, especially for individuals with worsening financial health, is an important topic for future research.

References

Anderson, Eric T and Duncan I Simester (2003). "Effects of 9$priceendingsonretailsales : Evidencefromfieldexperiments$". In: *Quantitative marketing and Economics* 1.1, pp. 93–110.

Golman, Russell, David Hagmann, and George Loewenstein (2017). "Information Avoidance". In: *Journal of Economic Literature* 55, pp. 96–135.

Huang, Szu-Chi (2018). "Social information avoidance: when, why, and how it is costly in goal pursuit". In: *Journal of Marketing Research* 55.3, pp. 382–395.

Huck, S, N Szech, and LM Wenner (2018). *More effort with less pay: On information avoidance, optimistic beliefs*. Tech. rep. and performance. working paper.

Karlan, Dean et al. (2016). "Getting to the top of mind: How reminders increase saving". In: *Management Science* 62.12, pp. 3393–3411.

Karlsson, Niklas, George Loewenstein, and Duane Seppi (2009). "The ostrich effect: Selective attention to information". In: *Journal of Risk and uncertainty* 38.2, pp. 95–115.

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor (2012). "Heuristic thinking and limited attention in the car market". In: *American Economic Review* 102.5, pp. 2206–36.

Liu, Xiao, Alan Montgomery, and Kannan Srinivasan (2018). "Analyzing Bank Overdraft Fees with Big Data". In: *Marketing Science* 37.6, pp. 855–882.

Lyter, David W et al. (1987). "The HIV antibody test: why gay and bisexual men want or do not want to know their results." In: *Public Health Reports* 102.5, p. 468.

Masatlioglu, Yusufcan, A Yesim Orhun, and Collin Raymond (2017). "Intrinsic information preferences and skewness". In: *Working Paper.*

Medina, Paolina C (2017). *Selective attention in consumer finance: Evidence from a randomized intervention in the credit card market*. Tech. rep. Working Paper, Northwestern University.

Shlain, Avner S (2018). "More than a penny's worth: Left-digit bias and firm pricing". In: *Working paper.*

Sicherman, Nachum et al. (2015). "Financial attention". In: *The Review of Financial Studies* 29.4, pp. 863–897.

Stango, Victor and Jonathan Zinman (2009). "What Do Consumers Really Pay on Their Checking and Credit Card Accounts ? Explicit , Implicit , and Avoidable Costs". In: *The American Economic Review* 99.2.

Sweeny, Kate et al. (2010). "Information Avoidance: Who, What, When, and Why". In: *Review of General Psychology* 14.4, pp. 340–353. ISSN: 10892680. DOI: 10.1037/a0021288.

Thomas, Manoj and Vicki Morwitz (2005). "Penny wise and pound foolish: the left-digit effect in price cognition". In: *Journal of Consumer Research* 32.1, pp. 54–64.

7 Appendix

Left Digit Bias Additional Tables

[Table 10 about here.]

Email Campaign User Summary Statistics

Figures 9 and 10 show histograms of the average credit scores and number of times they view their credit score for users who receive at least one of the email campaigns. Table 11 shows summary statistics for their email engagement.

[Table 11 about here.]

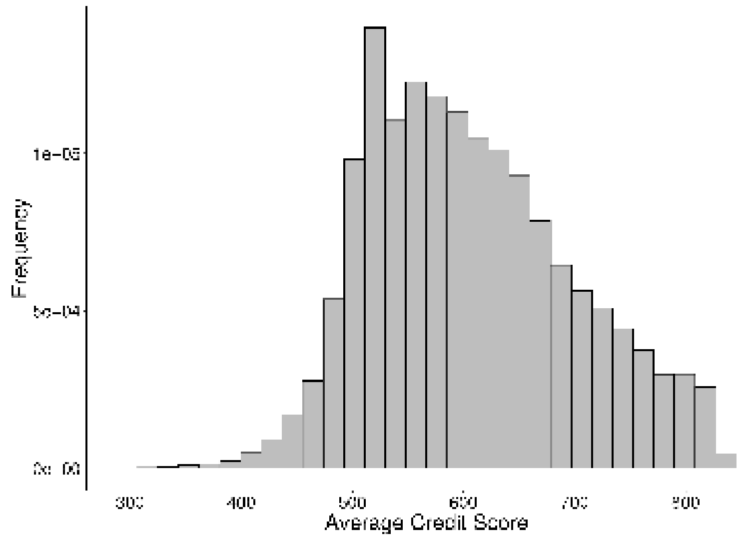

Figure 9: Histogram of the average credit score of the users in the email campaigns.

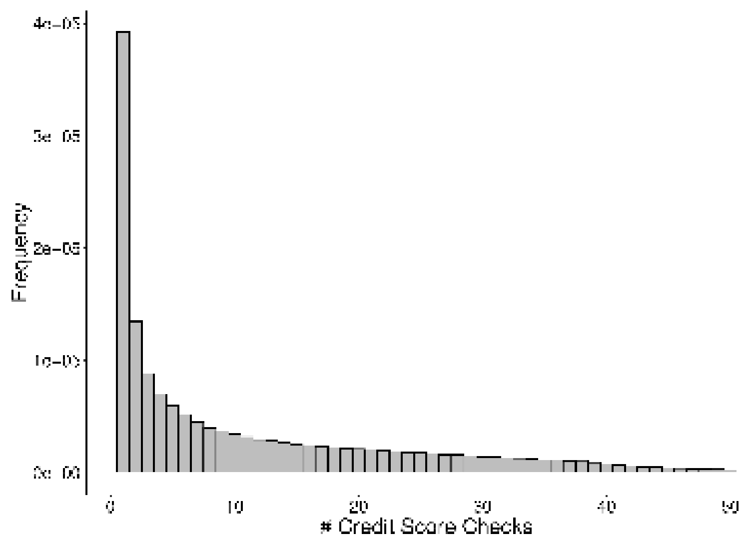

Figure 10: Histogram of the number of times a user views his or her credit score in the first year, for all users who have received at least one of the select email campaigns.

Email Randomization Checks

The identification strategies requires 1) exogenous variation introduced through email copies A/B tests within each campaign and 2) variation in click rates across campaigns that is not correlated with pre-treatment variables, conditional on campaign target rules. This section presents evidence of these conditions are fulfilled.

First, we test the randomization of the email copies/variants within a single campaign.

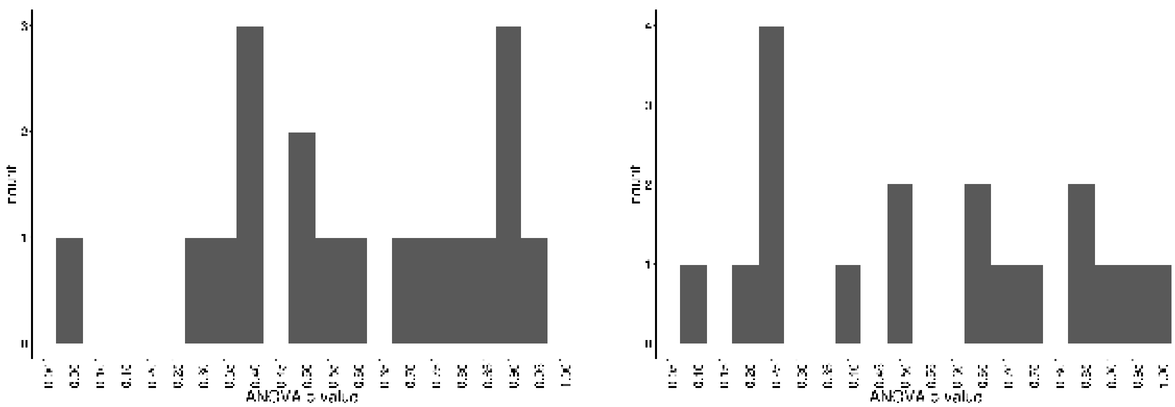

Figure 11: Histogram of p-values from ANOVA tests to detect if the user has a negative trend in their credit score (left plot) and credit score (right plot) are significantly different across email variants within each campaign. Each observation is one campaign.

Specifically, we test whether the credit scores and the trends in credit score are significantly different for users who receive different email variants within the same campaign. Figure 11 displays a histogram of the p-values from ANOVA tests, where the null hypothesis is that there is no significant difference in means across email variants. No campaigns had significantly different credit scores across users in different variants, while only one campaign had different rates of users with negative credit score trends (p-value = 0.06) across variants. This suggests that within campaigns, the variants are randomized across relevant observables.

Second, we show that after controlling for the campaign targeting rules, there is no correlation between which campaign the user received and her credit score and trend in credit score. In other words, we show that the targeting rules explain the vast majority of the variation in the credit score and the trend in credit score. To do this, for each campaign, we plot each outcome variable demeaned by the variables used in the targeting rules, compared to the outcome variable demeaned by the mean across all users in the sample, regardless of campaign.[18] Figure 12 displays these plots for whether a user has a declining credit score (left plot) and her credit score (right plot) prior to receiving an email in each campaign. We can visually see that much of the variation across campaigns in *negative_pre_trend*

[18] We demean by the targeting rules by estimating $y_i = \alpha + \beta X_i + \epsilon_i$, where X_i is the matrix of variables used in the targeting rules. These variables are listed in the regressors in Table 14, excluding *negative_pre_trend*. The demeaned variables are then the residuals of this regression, $y_i - \hat{y}_i$.

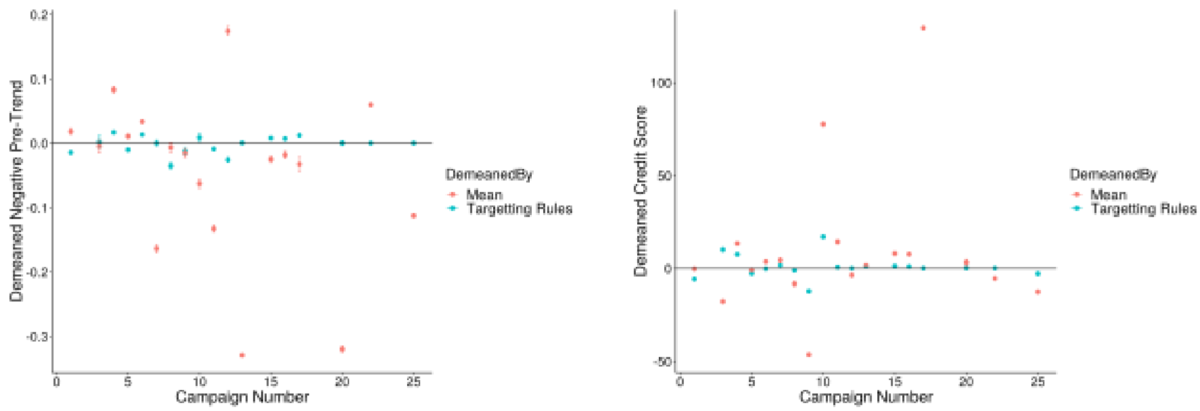

Figure 12: This figure shows the variation in *negative_pre_trend* (left) and credit score (right) before and after controlling for the campaign targeting rules. The red points represent the average variation in the outcome variables, where the outcome variables are demeaned by the average value across all users. The blue points are the average outcome variables after they are demeaned by the campaign targeting rules.

and credit score is eliminated after controlling for the campaign targeting rules. To compare this more quantitatively, we measure how much incremental variation campaign fixed effects explain. We compare the adjusted R^2 between the following regressions.

$$y_i = \alpha + \beta X_i + \epsilon_i$$

$$y_i = \alpha_c + \beta X_i + \eta_i$$

where α_c are campaign fixed effects. We find that adding campaign fixed effects explains only 0.37% more variation in credit scores and 0.04% more variation in whether the user has a declining credit score.[19]

Data Imputation and Bootstrap Steps

We predict whether a user's credit score at the time of checking their information is less or greater than their last observed credit score. However, this predicted variable is predicted with noise, which must be accounted for in regressions where predicted variable is on the

[19] For the regression on *negative_pre_trend*, adding campaign fixed effects increases the adjusted R^2 from 0.7919 to 0.7923, and 0.6646 to 0.6670 for the regression on credit score.

right hand side. Therefore, we bootstrap and then impute $negative\_cstrend\_check_{ie}$ to account for this noise. We use observations where users did check their credit information to predict the trend for users who did not check.

1. Randomly select 18 campaigns with replacement and get all observations (at the user level) for these randomly selected campaigns.

2. Estimate the following logit regression

$$negative\_cstrend\_check_{ie} = \alpha + \beta X_{it} + \epsilon_{it}$$

   where $X_{it} = \{negative\_pre\_trend_{i,t}, cs_{it}, negative\_pre\_trend_{i,t-1}, cs_{i,t-1}, cs^2_{i,t-1}, cs_{i,t-2}, cs^2_{i,t-2}, cs_{i1}, month, year, year^2, accountage_{it}, MonthsUntilCheck_{it}\}$. cs_{i1} is i's first observed credit score, and $MonthsuntilCheck_{it}$ is the number of months between when the user checked her score and the last time she checked $(t - (t-1))$.

3. For all users, predict $negative\_cst\hat{r}end\_check_{ie}$ given $\hat{\alpha}, \hat{\beta}$.

4. Estimate $CS\_post\_trend_{ie} = \alpha + \gamma check\_ind_{ie} + \epsilon_{ie}$ for users with negative pre-trends and $negative\_cstrend\_check_{ie} = 1$; negative pre-trends and $negative\_cstrend\_check_{ie} = 1$; positive/flat pre-trends and $negative\_cstrend\_check_{ie} = 1$; and positive/flat pre-trends and $negative\_cstrend\_check_{ie} = 0$.

Repeat Steps 1-4 500 times. Figure **??** plots the coefficient γ for each group, averaged over all bootstrap samples, along with the 2.5 and 97.5 percentiles to create the 5% confidence intervals.

What factors mainly affect credit score?

This platform offers a credit analysis, where it assigns a letter grade (A, B, C, D, or F) on different areas of credit that may affect the overall credit score. The factors are payment history, credit usage, credit age, account mix, credit inquiries, and debt to income ratio. Thus, when an individual's credit score changes, we can observe what areas that the individual improved upon, which can provide insight on the actions that a person takes to improve her score. This section provides a sense of how the average individual improves her credit score.

To quantify how the individual changes her credit score, we first map the letter grades to numbers, where a letter grade of "A" is mapped to 5, "B" to 4, etc. Then for individuals who experienced an increase in credit score between time t and $t-1$, we measure how each factor changed between t and $t-1$.

[Table 12 about here.]

[Table 13 about here.]

Full Regression Output

Here we show the full regression output of the first and second stage of the IV analysis.

[Table 14 about here.]

[Table 15 about here.]

[Table 16 about here.]

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	N
Change in CS	-386.00	-6.00	1.00	1.02	10.00	409.00	593123
Mon Between Checking	0.03	0.87	1.07	1.72	1.53	206.50	593123

Table 1: Summary statistics of the change in credit score between updates.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Emails Per Month	0.00	11.31	25.07	26.45	41.87	84.00
Percent Opened	0.18	6.52	16.76	25.29	36.98	100.00
Percent Clicked	0.00	3.11	12.50	21.06	30.56	100.00

Table 2: Summary statistics for the number of emails a user receives monthly, the percent of emails that he or she opens, and clicks on.

	Dependent variable:
	$\mathbb{1}\{CheckNextMonth\}$
$DecreasingScore$	-0.022^{***}
	(0.0004)
$\lvert\Delta_{it}^{cs}\rvert$	0.013^{***}
	(0.0005)
$DecreasingScore \times \lvert\Delta_{it}^{cs}\rvert$	-0.012^{***}
	(0.001)
NCheck	-0.019^{***}
	(0.0004)
Constant	0.726^{***}
	(0.0004)
Person FE	Y
Observations	4,348,553
Log Likelihood	−2,319,477.000
Akaike Inf. Crit.	4,638,967.000
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3: Logit results for Equation 1. Includes user random effects. Covariates have been scaled for model convergence.

	Dependent variable:
	$\mathbf{1}\{CheckNextMonth\}$
endsin00	0.024***
	(0.009)
endsin98	0.017*
	(0.009)
rounded_cs400	0.234***
	(0.083)
rounded_cs500	0.464***
	(0.077)
rounded_cs600	0.694***
	(0.076)
rounded_cs700	0.793***
	(0.076)
rounded_cs800	0.777***
	(0.077)
Constant	−0.087
	(0.076)
Observations	324,731
Log Likelihood	−209,976.800
Akaike Inf. Crit.	419,969.700
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4: Logit results of Equation 2.

subject_line
##[F][S][P][\first]##, Your Score Has Been Updated For August.
Don't forget to check your score change before changing your clocks this weekend
Your Year in Review
##[F][S][P][\first]##, get all 3 scores today!
##[F][S][P][\first]##, find out how to get all 3 scores today!
Do you know your credit age?
##[F][S][P][\first]##, we've organized your recommendations based on your great credit!
Your debt has changed.
Congrats! Your credit improved in 2016.
Find the right card for the holidays!
Start off 2017 with a free extra score update.
##[F][S][P][\first]##, we updated your Credit History Grade.
Your credit profile has changed since last month.
##[F][S][P][\first]##, find out how your credit usage affects your score.
##[F][S][P][\first]##, we made it easier for you to compare cards!
##[F][S][P][\first]##, we've customized your recommended actions to your amazing credit.
##[F][S][P][\first]##, weï¿œï¿œï¿œve updated your credit usage grade.
##[F][S][P][\first]##, get all 3 scores and your full credit profile today!
See your new score!
##[F][S][P][\first]##, have you seen your updated approval odds?
##[F][S][P][\first]##, your credit usage grade is ready for review.
Congrats! We're giving you an extra free score update this month.
Need a credit card? We got you.
##[F][S][P][\first]##, your updated credit usage grade is ready for review.
##[F][S][P][\first]##, have you seen your updated payment history grade?
##[F][S][P][\first]##, take advantage of your great credit usage grade today!
##[F][S][P][\first]##, find out if your other 2 scores are as awesome as your TransUnion score.
##[F][S][P][\first]##, are you spending too much? See your updated credit usage grade.
We've calculated how much you need to pay off to get an A.

Table 5: Subject lines for each of the emails sent out over all 16 campaigns. Each campaign has at least 2 email variants. "##[F][S][P][\first]##" is the user's first name.

	Dependent variable:	
	$\mathbb{1}\{CheckNextMonth\}$	
	(1)	(2)
abs(Change in CS)	-0.007^{***}	0.030^{***}
	(0.001)	(0.0004)
NCheck	-0.011^{***}	-0.032^{***}
	(0.001)	(0.001)
Constant	0.684^{***}	0.684^{***}
	(0.0003)	(0.0003)
Person FE	Y	Y
Trend in CS	Decreasing	Increasing
Observations	3,567,939	6,348,662
Log Likelihood	−2,245,097.000	−3,941,291.000
Akaike Inf. Crit.	4,490,204.000	7,882,593.000
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6: Logit results for Equation 1 for users in the email campaigns. Includes user random effects. Covariates have been scaled for model convergence.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Emails Per Month	0.15	9.41	21.59	31.05	39.81	431.65
Percent Opened	0.18	6.87	16.83	24.65	35.78	100.00
Percent Clicked	0.00	3.83	11.90	19.56	27.45	100.00

Table 7: Summary statistics for all emails sent to the selected sample of users. Percent Clicked is the percentage of emails that were clicked, conditional on opening.

	Dependent variable:	
	CS_{post}	
	(1)	(2)
check_ind	-25.163^{***}	5.827
	(5.524)	(3.947)
Campaign Controls	Y	Y
CS Trend	< 0	≥ 0
Observations	221,491	395,716
R^2	-0.088	0.019
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 8: 2SLS results. Controls for email targeting are omitted from output for better readability. Full regression output is in Table 15 in the Appendix. Standard errors are clustered at the campaign level.

	Dependent variable:					
	CS_{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
check_ind	-31.352^{***}	-3.154	-7.969	-4.136	25.004^{***}	13.731^{***}
	(5.944)	(8.078)	(5.355)	(3.550)	(6.569)	(4.186)
Campaign Controls	Y	Y	Y	Y	Y	Y
CS Trend	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
Tercile	First	Second	Third	First	Second	Third
Observations	80,448	74,369	72,845	123,997	135,655	152,359
R^2	-0.107	0.017	-0.012	0.006	-0.007	0.0003
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 9: 2SLS results by credit score tercile. The first three columns are for individuals with declining credit scores prior to checking their information, and the last three columns are for individuals who had either constant or increasing credit scores prior to checking their information. Controls for email targeting are omitted from output for better readability. Full regression output is in the Appendix.

	Dependent variable:				
	$\mathbf{1}\{CheckNextMonth\}$				
	(1)	(2)	(3)	(4)	(5)
endsin00	0.148	0.126***	0.012	0.039*	0.070**
	(0.097)	(0.024)	(0.018)	(0.023)	(0.033)
endsin98	0.141	0.096***	−0.021	−0.022	−0.001
	(0.099)	(0.025)	(0.017)	(0.023)	(0.033)
Constant	−0.324***	−0.166***	0.247***	0.311***	0.327***
	(0.069)	(0.017)	(0.012)	(0.016)	(0.023)
Rounded CS	400	500	600	700	800
Observations	2,551	39,700	79,402	44,757	23,377
Log Likelihood	−1,750.345	−27,457.870	−54,450.760	−30,467.100	−15,846.700
Akaike Inf. Crit.	3,506.691	54,921.730	108,907.500	60,940.200	31,699.390
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 10: Logit regression results of whether the user checks her score in the next month on the last 2 digits of her credit score for each 100 threshold. The observations in Column 1 are for users with credit scores equal to 398, 399, and 400. Column 2 is users with credit scores equal to 498, 499, 500, and so on.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Emails Per Month	0.75	26.55	36.29	39.35	49.52	252.30
Percent Opened	0.20	8.71	19.51	27.10	39.42	100.00
Percent Clicked	0.00	3.65	10.64	17.41	24.44	100.00

Table 11: Sample of users who receive email campaign: Summary statistics for the number of emails a user receives monthly, the percent of emails that he or she opens, and clicks on.

	Mean	SD
Payment History Change	0.08	0.53
Credit Utilization Change	0.36	1.07
Credit Age Change	0.01	0.46
Account Mix Change	0.03	0.40
Credit Inquiries Change	0.00	0.54
Debt:Income Change	0.36	1.18

Table 12: Summary statistics for how credit score "grades" change between credit report updates. The mean is the average probability that the corresponding grade has changed between credit report updates.

	Dependent variable:	
	1{*Check*}	
	(1)	(2)
negative_pre_trend	−0.052***	−0.054***
	(0.003)	(0.003)
Homeowner	0.010***	0.011***
	(0.001)	(0.001)
ccclicks	0.033***	0.026***
	(0.001)	(0.001)
active	0.041***	0.032***
	(0.002)	(0.002)
refreshSameMonth	0.265***	0.174***
	(0.001)	(0.003)
CreditUtilizationAB	−0.061***	−0.215***
	(0.004)	(0.005)
CreditUtilizationDF	−0.139***	−0.152***
	(0.005)	(0.005)
CSgreater700	0.008***	0.011***
	(0.001)	(0.002)
ChangeinDebt	0.093***	0.085***
	(0.002)	(0.002)
Collections	−0.033***	−0.026***
	(0.001)	(0.001)
PayHistoryGradeCDF	0.063***	−0.013***
	(0.004)	(0.004)
CreditAgeGradeB	0.036***	0.038***
	(0.014)	(0.014)
CreditAgeGradeC	0.035**	0.038***
	(0.015)	(0.014)
CreditAgeGradeD	0.059***	0.057***
	(0.014)	(0.014)
CreditAgeGradeF	0.098***	0.087***
	(0.016)	(0.016)
CreditAgeGrademissing	−0.141***	−0.141***
	(0.011)	(0.023)
CreditUtilizationNotA	−0.021***	−0.178***
	(0.005)	(0.005)
ChangeinCreditUtilizationGrade	−0.255***	−0.167***
	(0.003)	(0.004)
DebtIncrease	−0.015***	−0.012***
	(0.001)	(0.001)
CSIncrease	−0.016***	−0.031***
	(0.003)	(0.003)
ChangeinCS	0.050***	0.055***
	(0.003)	(0.003)
CSIncrease2016	0.017***	0.025***
	(0.002)	(0.003)
Constant	0.530***	
	(0.011)	
Email FE	N	Y
Observations	639,673	639,673
R^2	0.122	0.137
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 13: OLS results for the restricted (first column) and unrestricted (second column) first stage regression.

	Dependent variable:	
	$1\{Check\}$	
	(1)	(2)
Homeowner	0.016***	0.008***
	(0.002)	(0.002)
ccclicks	0.036***	0.029***
	(0.002)	(0.001)
active	0.030***	0.044***
	(0.004)	(0.003)
refreshSameMonth	0.227***	0.272***
	(0.002)	(0.002)
CreditUtilizationAB	−0.130***	−0.048***
	(0.015)	(0.005)
CreditUtilizationDF	−0.097***	−0.128***
	(0.015)	(0.005)
CSgreater700	0.022***	−0.005**
	(0.003)	(0.002)
ChangeinDebt	0.078***	0.091***
	(0.005)	(0.002)
Collections	−0.028***	−0.035***
	(0.002)	(0.002)
PayHistoryGradeCDF	0.031***	0.058***
	(0.012)	(0.005)
CreditAgeGradeB	0.055	0.020
	(0.104)	(0.017)
CreditAgeGradeC	0.021	0.026
	(0.109)	(0.018)
CreditAgeGradeD	−0.044	0.049***
	(0.108)	(0.018)
CreditAgeGradeF	−0.029	0.080***
	(0.120)	(0.023)
CreditAgeGrademissing	−0.338***	−0.113***
	(0.081)	(0.013)
CreditUtilizationNotA	−0.111***	−0.003
	(0.015)	(0.005)
ChangeinCreditUtilizationGrade	−0.347***	−0.220***
	(0.014)	(0.004)
DebtIncrease	−0.006***	−0.020***
	(0.002)	(0.002)
CSIncrease	0.118***	−0.075***
	(0.005)	(0.004)
ChangeinCS	0.120***	0.097***
	(0.007)	(0.004)
CSIncrease2016	0.007**	0.016***
	(0.004)	(0.002)
Constant	0.637***	0.504***
	(0.082)	(0.014)
CS Trend	< 0	≥ 0
Observations	221,491	395,716
R^2	0.121	0.121
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 14: OLS results for the first stage in the IV estimation.

	Dependent variable:	
	CS_{post}	
	(1)	(2)
check_ind	−25.163***	5.827
	(5.524)	(3.947)
active	−0.763***	−0.267
	(0.279)	(0.403)
Homeowner	2.361***	1.452***
	(0.220)	(0.188)
ccclicks	−1.466***	−2.699***
	(0.391)	(0.234)
refreshSameMonth	9.446***	3.032***
	(1.358)	(1.132)
CreditUtilizationAB	−4.216**	−0.350
	(2.145)	(0.446)
CreditUtilizationDF	7.917**	4.815***
	(3.375)	(0.774)
CSgreater700	−2.999***	−3.716***
	(0.536)	(0.392)
ChangeinDebt	2.224***	−3.277***
	(0.699)	(0.449)
Collections	2.494***	2.379***
	(0.474)	(0.421)
PayHistoryGradeCDF	13.605***	1.312
	(2.536)	(1.009)
CreditAgeGradeB	0.511	1.401***
	(4.199)	(0.247)
CreditAgeGradeC	−3.088	−2.167***
	(3.932)	(0.193)
CreditAgeGradeD	1.164	−0.625
	(3.085)	(0.819)
CreditAgeGradeF	−6.454	−3.437***
	(8.107)	(1.017)
CreditAgeGrademissing	−7.357***	−1.438
	(2.165)	(0.931)
CreditUtilizationNotA	−11.152***	2.210*
	(2.387)	(1.313)
ChangeinCreditUtilizationGrade	−15.968***	−2.656**
	(3.842)	(1.164)
DebtIncrease	−5.087***	−5.011***
	(0.209)	(0.192)
CSIncrease	−0.717	−6.310***
	(0.812)	(0.630)
ChangeinCS	−3.305***	2.926***
	(0.961)	(0.701)
CSIncrease2016	−2.120***	−0.977***
	(0.649)	(0.272)
Constant	22.564***	−1.643
	(3.588)	(2.370)
Campaign Controls	Y	Y
CS Trend	< 0	≥ 0
Observations	221,491	395,716
R^2	−0.088	0.019
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 15: 2SLS results. Output includes all campaign targeting controls.

	Dependent variable:					
	CS_{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
check_ind	−31.352***	−3.154	−7.969	−4.136	25.004***	13.731***
	(5.944)	(8.078)	(5.355)	(3.550)	(6.569)	(4.186)
active	−3.027***	−2.448***	0.378	−1.643*	−2.184***	−0.445
	(0.539)	(0.708)	(0.515)	(0.909)	(0.827)	(0.321)
Homeowner	2.732***	4.348***	2.742***	1.464***	3.029***	2.531***
	(0.411)	(0.501)	(0.288)	(0.333)	(0.399)	(0.230)
ccclicks	−0.803**	−2.463***	−3.190***	−2.115***	−3.263***	−3.597***
	(0.374)	(0.566)	(0.639)	(0.165)	(0.420)	(0.526)
refreshSameMonth	10.041***	8.939***	4.935***	4.161***	1.715	1.394
	(1.563)	(2.003)	(1.194)	(1.000)	(2.067)	(1.095)
CreditUtilizationAB	0.368	3.666	−0.197	3.827***	2.340***	3.072***
	(4.527)	(4.063)	(3.022)	(1.235)	(0.606)	(0.765)
CreditUtilizationDF	−4.888	−4.842	−2.405	2.040**	0.428	1.070
	(3.452)	(3.812)	(4.784)	(0.904)	(0.809)	(1.077)
CSgreater700			0.330			−0.468
			(0.449)			(0.387)
ChangeinDebt	2.897***	5.641***	7.408***	−3.053***	0.427	3.266***
	(0.682)	(1.159)	(0.997)	(0.442)	(0.653)	(0.591)
Collections	−0.193	−3.141***	−2.186***	−1.715***	−2.524***	−2.019***
	(0.556)	(0.294)	(0.781)	(0.499)	(0.232)	(0.283)
PayHistoryGradeCDF	12.522***	6.586*	−7.605**	0.357	−1.242*	−0.560
	(4.488)	(3.484)	(3.448)	(1.899)	(0.637)	(0.632)
CreditAgeGradeB	5.241	2.942	−12.789	4.644**	−0.112	2.283***
	(1.658)	(4.705)	(15.462)	(1.811)	(0.917)	(0.457)
CreditAgeGradeC	−0.253	−4.329***	−18.343**	−0.531	−1.053	1.152
	(8.690)	(1.433)	(8.224)	(1.604)	(2.040)	(2.007)
CreditAgeGradeD	1.019	2.379	−1.056	4.468***	−0.282	−0.027
	(3.501)	(2.488)	(14.611)	(1.363)	(1.437)	(1.611)
CreditAgeGradeF	−11.066***	−6.670***	−9.982	0.392	−3.699***	0.520
	(18.716)	(1.134)	(6.550)	(1.649)	(0.657)	(2.116)
CreditAgeGrademissing	5.070	−8.764***	−18.802**	2.425***	0.001	−1.114
	(3.708)	(3.311)	(7.424)	(0.652)	(1.489)	(0.990)
CreditUtilizationNotA	4.419	−6.678**	−7.932***	5.038***	6.575***	1.394
	(5.254)	(3.010)	(2.434)	(1.790)	(0.473)	(1.135)
ChangeinCreditUtilizationGrade	−13.228***	−12.472***	−6.595	0.244	−2.038*	−4.922***
	(2.754)	(4.090)	(4.418)	(1.301)	(1.236)	(1.235)
DebtIncrease	−6.862***	−5.589***	−1.455***	−4.323***	−6.413***	−4.118***
	(0.408)	(0.349)	(0.268)	(0.280)	(0.308)	(0.245)
CSIncrease	0.877	−1.313	−1.474**	−5.899***	−5.838***	−5.401***
	(1.398)	(0.904)	(0.720)	(0.773)	(0.977)	(0.661)
ChangeinCS	−2.060**	−8.734***	−5.583***	4.713***	0.937	1.039
	(1.357)	(1.306)	(1.376)	(0.872)	(1.025)	(0.702)
CSIncrease2016	−2.756	−1.454**	−0.659	−1.707***	−0.963**	−0.089
	(0.825)	(0.625)	(0.817)	(0.346)	(0.447)	(0.379)
Constant	21.608***	10.983*	15.524*	8.250***	−17.643***	−15.408***
	(4.788)	(5.983)	(8.172)	(2.064)	(3.918)	(2.601)
Campaign Controls	Y	Y	Y	Y	Y	Y
CS Trend	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
Tercile	First	Second	Third	First	Second	Third
Observations	80,448	74,369	72,845	123,997	135,655	152,359
R^2	−0.107	0.017	−0.012	0.006	−0.007	0.0003
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 16: Full 2SLS regression outputs for users separated by their credit score tercile.