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

Jessica Fong† Megan Hunter‡

January 24, 2021

Abstract

This paper explores the impact of information avoidance in the context of consumer finance. Specifically, under what circumstances do individuals avoid information about their credit, and how does avoiding this information affect their future credit scores? Using data from a consumer finance platform, we find that a decline in credit score decreases the likelihood that an individual views her credit report in the future. We then measure the impact of receiving information on future credit scores, especially for those likely to avoid information. To obtain a causal local average treatment effect, we use variation in whether an individual views her credit report induced by email campaign A/B tests on a subsample of users who do not opt out of email communication. We find heterogeneous effects of information on credit scores. For individuals who were more likely to avoid information (users whose credit scores were decreasing), viewing their credit reports further decreases credit scores, while information increases credit scores for individuals less likely to avoid information. This finding suggests that encouraging individuals to access information when they are more likely to avoid information may worsen their financial health. We discuss the implications for firms' targeting strategies in retention efforts.

Keywords— information avoidance, consumer finance, instrumental variables, economics of information

*We are thankful for helpful comments and suggestions from Sam Antill, Katherine Carman, Andrey Fradkin, Wes Hartmann, Szu-Chi Huang, Avi Goldfarb, Puneet Manchanda, Sridhar Narayanan, Yesim Orhun, Navdeep Sahni, and Chunhua Wu, along with an anonymous Associate Editor, two anonymous referees, and participants at the Stanford Marketing WIP seminar, Advances in Field Experiments 2019, Bass FORMS Conference 2020, Cornell SC Johnson College of Business Marketing Workshop, Marketing Science 2020, and the RAND Behavioral Finance Forum. We are also grateful to the firm for providing the data. All remaining errors are our own.

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

‡Carroll School of Management, Boston College, megan.hunter@bc.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 such as Fitbit or Apple Watch, and regulate their monetary spending and budget using websites like Mint. A major motivation for consumers to join these platforms is to improve their outcomes, whether it is their weight, fitness, or financial health. Advertising by these platforms often promises 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."

Although sign up rates for information-tracking services are high, retention rates are low. For example, 42% of consumers who buy a fitness tracker stop using it within six months of purchase.[1] This pattern is also prevalent in our setting, which is a consumer-financial-tracking website through which consumers receive free credit reports when they log on each month. We find that 38% of users do not return after receiving their first credit report. This paper explores a potential reason for high attrition - information avoidance, and its impact on outcomes for individuals.

Specifically, the research questions are two-fold. First, what drives an individual's demand for information? To form our hypothesis about what causes individuals to decrease their demand for information, we draw on the literature on information avoidance. Information avoidance, defined by Sweeny et al. (2010) as "any behavior intended to prevent or delay the acquisition of available but *potentially* unwanted information" has been documented in many settings (Golman et al., 2017; Karlsson et al., 2009; Lyter et al., 1987). We extend these findings and hypothesize that *receiving* unwanted information causes further information avoidance. In this paper's setting, we define unwanted information as a declining credit score.

There are two main challenges in measuring the causal effect of the trajectory of credit score on demand for information. The first is selection bias. Individuals who have declining credit scores may be less financially responsible, and are less likely to check their financial information. Second, there may be shocks that are unobserved by the econometrician that affect both the user's financial status and her valuation of time, making it less likely that she logs on to her account and views her credit report. For example, sudden financial stress not only impacts a user's credit score but also her value of time, making it less likely for her to log on to her website in a given month.

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

To address these identification challenges, we take advantage of the well-documented "left-digit bias" phenomenon, whereby 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). For example, the difference between 500 and 499 seems relatively larger than the difference between 499 and 498, even though the absolute difference is the same. Since the difference is the same and the levels are similar, a user with these scores is likely to fare similarly on unobservables. We use a regression discontinuity approach to measure how a change in users' hundreds digit of their credit score affects their demand for information. We find that the larger the user's perceived drop in credit score, the less likely she is to view her updated credit report in the future.

Given that users avoid information after receiving bad news, the second research question focuses on the impacts of information and information avoidance: how does information affect outcomes? For instance, can avoiding information about one's credit improve one's future credit scores? With this question, we add to the literature on the consequences of information avoidance, which have discussed and/or found positive (Sicherman et al., 2015; Huck et al., 2018; Goulas and Megalokonomou, 2015) and negative implications (Golman et al., 2017) of information avoidance.[2] Since whether a user views her credit report is endogenous to her credit score trajectory, we utilize multiple A/B tests that the firm had run on email campaigns. Email copies vary in the rate at which they get users to log on to the website, which generates variation in whether a user views her report in a given month. While the email a user receives is generally correlated with her credit and other unobservable characteristics, we select a subset of email copies that were involved in A/B tests for which we observe the targeting rules used by the firm to decide which users to send the campaign to. After controlling for the emails' targeting rules, the particular email that the user receives is not correlated with her credit report and website usage due to the randomization introduced by the A/B tests. In addition to being uncorrelated with unobservable shocks, the emails must also satisfy the exclusion restriction. In this context, this means we assume that the emails themselves do not directly cause the users to change their behavior to change their credit score and do not differentially impact whether users' return to the platform after receiving the email, which would lead to selection bias.[3]

Given the setting and available data, in order to implement this analysis, we focus on users who have not opted out of email communication from the firm (since the emails provide the exogenous variation in checking) and who have checked their credit report at least three times (since we must observe the users' trends in credit score before receiving

[2]Refer to Section 2 for a more detailed discussion of this literature.

[3]We provide justifications and robustness checks for why these assumptions are reasonable in Sections 5.3.3 and 6.4, respectively.

the email and her credit score after receiving the email).[4] We find a *negative* intent-to-treat effect of the emails for users with declining credit scores, and a *positive* intent-to-treat effect for users with non-declining scores; receiving a more effective email (which are emails with higher click-rates) leads to a decrease (increase) in credit score if a user's previous credit score trend was decreasing (non-decreasing). Using the emails as an instrument for whether the user checks her report, we are able to obtain a causal local average treatment effect (LATE) of checking one's credit report on one's future credit score. Again, we find a heterogeneous LATE of information on credit scores. On average, users who had declining credit score trends prior to checking their credit report experience a 23-point decrease in credit score after viewing their updated report, while users who had non-declining credit scores experienced a nine-point increase in their credit score after viewing their updated report. This negative effect for users with declining credit scores is largely driven by users who had credit scores in the bottom tercile of the sample (lower than 581) prior to receiving an email.

Although this effect applies only to compliers, who in this setting are users who would not have checked their score unless they received an email that had a high rate of encouraging users to view their credit report, the LATE may inform the firm's email targeting policies, as firms can change behavior only for those users who engage with their emails.

The results suggest that avoiding information may be helpful to some users in increasing their credit scores. This has implications not only for firms' customer retention strategies, in that nudging users to return to the website may not be effective for all users, but also for their strategies in helping their users improve their credit scores. Through simulations of credit scores and retention, given the effects information and the avoidance of information, we show that sending retention emails to those who are the most likely to leave the platform, which are users with declining credit scores, leads to the largest reduction in retention and credit scores, compared to policies that target all users or only users with increasing scores. Therefore, accounting for user trajectories in targeting criteria for retention campaigns may benefit both the firm in terms of retention and the users in terms of financial health.

The remainder of the paper is structured as follows. Section 2 describes the related literature. In Section 3, we provide background on the empirical setting of this paper and summary statistics of the types of users on this platform. Section 4 presents evidence for the impact of credit score trajectories on demand for information. Section 5 explains the instrumental variable design for the causal analysis of the effect of information on outcomes.

[4]Note that this sample is not representative of the firm's general consumer population. Of the general population, 38% fit the criteria for this sample. Although the sample differs from the general population, we show that they are similar in several important dimensions, such as their propensity to avoid information and their monthly changes in credit score.

In Section 6, we present the results of the IV analysis, and Section 7 concludes.

2 Literature Review

The findings in this paper contribute mainly to the literature on information avoidance and selective attention. The information avoidance literature has largely focused on the finding that individuals avoid information when they feel uncertain about the outcome (Kőszegi, 2003; Masatlioglu et al., 2017; Karlsson et al., 2009). This tendency to avoid confirming bad news has been documented in a variety of settings. People who undergo HIV tests fail to return to receive their results (Centers for Disease Control and Prevention, 1997; Lyter et al., 1987; Thornton, 2008), investors are less likely to view their holdings if the stock market goes down (Sicherman et al., 2015), and during goal pursuit activities, people avoid information about the performance of others to avoid potentially negative comparisons (Huang, 2018). We contribute to this area of the literature by showing that people may continue to avoid information, even after the bad news (a decreasing credit score) has been confirmed.

We further contribute to the literature on information avoidance and selective attention by demonstrating the implications of information avoidance on future outcomes. Existing literature has suggested that in some cases, information avoidance may be detrimental, since information may be helpful in decision-making. Golman et al. (2017) discuss potential negative implications of information avoidance on the spread of disease, groupthink, and media bias. A few papers consider the positive effects of information avoidance. Sicherman et al. (2015) hypothesize that when the stock market dips, information avoidance may help prevent investors from panicking and making rash decisions. In a lab study, Huck et al. (2018) show that subjects who completed a task at a randomized hourly rate perform worse when they discover that their wage is lower relative to their peers. Using field data, Goulas and Megalokonomou (2015) find that students who learned that they were low achieving on an exam, compared to their peers, did worse on the next exam than students who did not learn they were low achieving. Theoretically, this paper's finding that people avoid more information after receiving bad news adds support to the motivation maintenance theory, which posits that individuals avoid potentially negative information because they fear that the information may be demotivating (Golman et al., 2017).

This work also relates to the growing literature on selective attention in consumer finance. Stango and Zinman (2009); Medina (2017); Karlan et al. (2016); Liu et al. (2018), among many others, show that reminders in the contexts of overdrafts and savings accounts can reduce overdrafts and increase savings as well as have heterogeneous effects on different types of users.

3 Setting

The empirical setting of this paper is a financial monitoring firm with more than 15 million users in the United States.[5] The main service that this firm provides is that it allows users to check their credit reports for free on a monthly basis when users log on to their website. Other services include matching users to loans and credit cards[6] and providing fraud monitoring alerts. This paper focuses on the credit report aspect of the website.

When a user creates an account, she inputs her social security number, and all of her credit lines are automatically added to her account. This includes credit cards, mortgages, auto loans, student loans, etc. The firm pulls the user's TransUnion credit profile when the user logs on to the firm's website for the first time in a calendar month, which includes the TransUnion VantageScore 3.0 credit score.[7] The credit score, along with other credit information, is displayed on the homepage. Figure 1 provides an example of what this information looks like to the user. In addition to the credit score, the homepage includes the user's credit score tier ("Excellent", "Good", "Fair", "Poor", or "Very Poor"); her credit grades, which are letter grades between A and F for aspects of a user's financial health that make up the total credit score; and how much her credit score has changed since her last report.

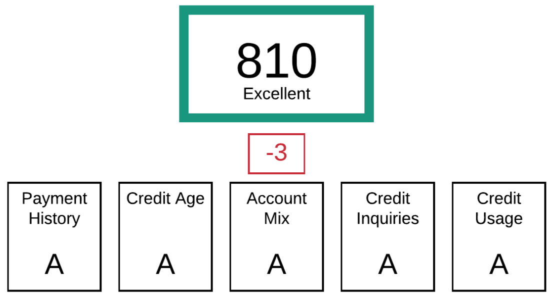
The timing of when credit reports are updated is important for the terminology used in this paper. A credit report is pulled (updated) the first time a user logs on during a calendar month. If the user does not log on in a calendar month, the website does not pull her credit report, and neither we nor the user can observe her credit report through the website. To illustrate, consider a user who creates an account on June 30, and logs on again on July 1, July 15, and September 10. Her credit report will be updated on June 30, July 1, and September 10 but not July 15. Additionally, there will be no updated credit report for August. On July 15, when she logs on, she sees her credit report that was updated on July 1. For the remainder of the paper, we refer to a user "checking her report," and "checking her updated report" as the user logging on to her account for the first time in the calendar

[5]As of 2018.

[6]The website suggests credit cards, loans, and other credit lines to users, given their credit history and account information

[7]TransUnion VantageScore 3.0 credit scores are similar to FICO credit scores. The main difference is the weights that they place on different credit areas in creating the credit scores. See `https://www.experian.com/blogs/ask-experian/the-difference-between-vantage-scores-and-fico-scores/` for a more detailed explanation of the differences.

Figure 1: Mock Design of Credit Information on the Homepage



Notes: This figure displays the credit report information that the user sees immediately upon logging on to her account. Her credit score is displayed along with her credit tier (Excellent, Good, Fair, Poor, Very Poor) and associated color. The "-3" is how much her credit score has changed since her last observed credit report. The bottom five panels are the letter grades for each credit grade.

month and viewing her updated credit report on the homepage. We classify this user as having checked her report on June 30, July 1, and September 10.

Finally, the firm monetizes through a freemium model, in which users can pay to upgrade to the premium account, in which they can update their credit reports weekly, rather than monthly. Around 1% of users are premium users. The remainder of this paper focuses on free users.

3.1 Data

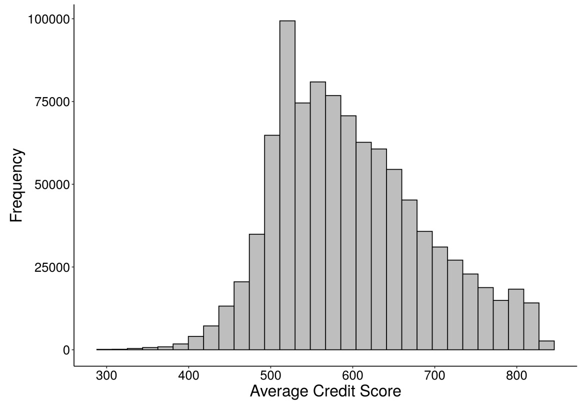
In this section, we provide background on the types of individuals who use this platform, how their credit scores change, and how often they check their credit reports. We select a random sample of 969,254 users who have logged on to their account between January 1, 2016 and June 12, 2018 (the maximum observed date in the data).[8] We then pulled all the credit reports for these users. The final data set consists of 5,725,491 credit reports, ranging from April 23, 2015 to June 12, 2018.

Figure 2 displays a histogram of the average credit score for each user. The average user's credit score is 598 with a standard deviation of 90, which is significantly lower than the 2018 national average of 673, as measured by VantageScore.[9] Thus, these users generally have lower financial health than the average American, suggesting that they may have joined this website hoping to improve their credit.

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

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

Figure 2: Histogram of the Users' Average Credit Score

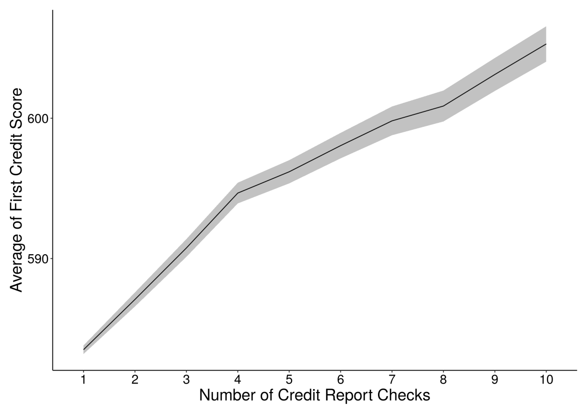


Notes: The sample consists of the random sample of 969,254 users. Each observation is at the individual level. A user's average credit score is averaged over all her observed credit reports.

Using this random sample, we explore what types of users have the lowest retention rates. Since 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 continue to use the website and check their credit reports. However, we find that these users are the ones who are most likely to attrit. Users who have lower credit scores are less likely to check their credit report in the future, as shown by Figure 3, which displays the average first observed credit score for users grouped by how many times they checked their updated credit report. For instance, for users who have updated their credit report 10 times, their average credit score when they first joined the website was 605, while users who checked their credit report only once had an average first credit score of 585. Note that this positive trend is not due to the possibility that checking scores causes an increase in credit score, since this figure plots the average of the *first* observed credit score.

Figure 4 displays the empirical CDF of the number of times a user checks her credit report. Sixty-two percent of users check their reports more than once, and 31% of users check their score more than five times. To explore changes in score, we focus on users in this random sample who return to the website after signing up (they have checked their updated credit reports at least twice). Table 1 provides more detailed summary statistics on how frequently a user updates her credit report and how her credit score changes over time. For these users, conditional on returning, the average user checks her updated credit report every one to two months. Since most individuals log on in consecutive months, changes in credit scores are small but on average increasing, with an average of a one-point increase in credit scores between updated reports. While the average change is small, the variance

Figure 3: The Average First Observed Credit Score for Users Grouped by the Number of Times that they Check their Credit Report



Notes: Each observation is a user's credit score from the first credit report she checked through this platform. The horizontal axis is the number of times the user has checked her updated credit report cut at 10 checks. Shaded areas depict 95% confidence intervals.

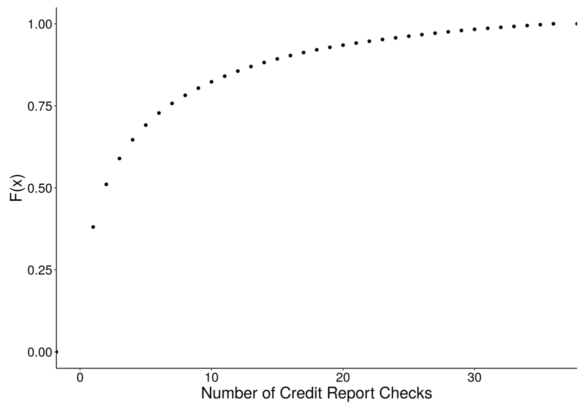
is relatively high; 50% of all consecutive changes in credit scores between updated reports are between -6 and +10, and the other 50% experience even larger changes. Additionally, in a given three-month period, about one in four users have credit score changes of at least 20 points. These large changes in credit scores may have resulted from actions that quickly impact scores, such as adding credit lines, increasing the amount of credit allowed, having more or fewer late payments, and changing the amount of debt. To provide a sense of magnitude of these changes in credit scores, a 20-point decrease in credit score can increase mortgage interest rates by 0.2%, leading to $8,640 more in payments on a $200,000 30-year fixed-rate mortgage.[10]

3.1.1 Email Data

The website also sends emails to users who do not opt out of email communication. The most common types of emails notify the user that they can check their new score this month, or suggest loans or credit card offers. Our identification strategy relies on emails that the user receives, so we provide summary statistics about the frequency of emails and user engagement with emails, which are shown in Table 2. Note that Table 2 includes users who have opted out of any email communication, hence the users who receive zero emails per month. On average, a user receives 26 emails per month, and opens 25% of them. She

[10]This example is based on interest rates in May 2020, calculated using `https://www.myfico.com/loan-center/home-mortgage-rate-comparison/default.aspx`.

Figure 4: CDF of Number of Credit Score Checks



Notes: This plots the empirical cumulative distribution function of the number of times a user checks her credit report. Each observation is at the individual level and sample is 969,254 users. For readability, the data is truncated to 36 or fewer checks, however in the data the maximum number of checks is 81. 0.5% of users have more than 36 credit report checks.

Table 1: Summary Statistics for Changes in Credit Score and Frequency of Checking Between Consecutive Reports

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Change in CS	-386.00	-6.00	1.00	0.93	10.00	326.00
Months Between Checking CS	0.00	0.90	1.07	1.73	1.57	36.53

Notes: The sample is 561,864 users who have checked their score at least twice. "Change in CS" is the difference in credit score between the consecutive reports, and "Months Between CHecking CS" is the number of months between when the user checks her updated reports.

clicks on 21% of the emails that she opens.

4 Evidence for Change in the Demand For Information

Theory and empirical evidence on information avoidance have shown that individuals are more likely to avoid information that is "potentially unwanted" (Sweeny et al. (2010); Golman et al. (2017); Kőszegi (2003), among others). In this setting, we define "unwanted information" to be a decrease in a user's credit score. We chose this metric, as opposed to a

Table 2: Email Engagement Summary Statistics

	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

Notes: This table presents summary statistics for the average number of emails a user receives each month, the percent of emails that she opens, and percentage of emails that she clicks on, conditional on opening. Each observation is at the individual level.

low credit score, because not all users may classify a low score as "unwanted." For example, a user with a score that decreases from 500 to 400 considers 400 to be negative information, but a user whose score increased from 300 to 400 may consider 400 to be positive.[11]

We measure how a decrease in credit score impacts whether a user checks her updated credit report in the next month. This not only answers the question of whether users are avoiding future information after receiving negative information, but also provides a way to identify which users are more likely to avoid information. Identifying which users avoid information will be relevant in Section 5 in which we measure the heterogeneous effects of information on credit score.

4.1 Leveraging Left Digit Bias

We measure how a user's current credit score trend affects the likelihood of whether the user checks her report the next month. In other words, if a user checks her credit report today and sees that her score has decreased, how likely is she to check her updated report in the next calendar month? A challenge in measuring the causal effect of a change in credit score is the presence of unobserved shocks that simultaneously affect both the individual's credit score and the availability of her time. For instance, a sudden financial stress may decrease the user's credit score and also reduce the time she has available to log on to her account and view her credit report. To control for these unobserved shocks, we leverage the left-digit bias phenomenon in a regression discontinuity framework.

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

[11] Additionally, the direction of the changes in credit score seem to follow a random walk, as verified by a Wald-Wolfowitz runs test, indicating changes in credit score tend to not be persistent in one direction, and thus, are informative to the users.

documented in a variety of settings, from prices (Anderson and Simester, 2003; Thomas and Morwitz, 2005; Shlain, 2018) to car mileage (Lacetera et al., 2012). The general finding in this literature is that there is a greater perceived difference between numbers that end in \$9 and \$0, compared to other one-unit differences with other digits. By assuming that users with similar credit scores and credit score trends are similar in unobservables, this phenomenon allows us to causally measure whether a larger perceived drop in credit score causes the individual to avoid information.

We outline the intuition before presenting the regression. Consider an individual A who observes that her score drops two points from 502 to 500. Another individual B observes her score drops two points as well from 501 to 499. While these two individuals have the same magnitude drop in credit score and similar levels, B perceives a greater drop in score because her left-most digit changed. If information avoidance exists, then B will be less likely to check her score next month than A due to this larger perceived drop.

Following this intuition, we create the following regression discontinuity framework (Lee and Lemieux, 2010). We select credit reports in which the credit score is within 10 points to the nearest hundred. This includes credit scores from 390 to 409, 490 to 509, 590 to 609, etc. We create an indicator variable $\mathbb{1}\{LeftDigitChanged\}$, which is equal to 1 if at t, the user's left-most digit has changed since her last observed credit score at $t-1$. To illustrate, if the user's score drops from 505 to 499, then $\mathbb{1}\{LeftDigitChanged\} = 1$. To measure the difference in effect between users whose score dropped to below the hundred threshold and users whose score increased to above the hundred threshold, we interact $\mathbb{1}\{LeftDigitChanged\}$ with $\mathbb{1}\{\Delta CS_{it} < 0\}$, resulting in the regression below.[12][13]

$$\begin{aligned}\mathbb{1}\{CheckNextMonth\}_{it} = \alpha_m + \alpha_w + \beta_1 \mathbb{1}\{LeftDigitChanged\}_{it} + \beta_2 \mathbb{1}\{\Delta CS_{it} < 0\} \\ + \beta_3(\mathbb{1}\{LeftDigitChanged\}_{it} \times \mathbb{1}\{\Delta CS_{it} < 0\}) \\ + \beta_4 RoundedCS_{it} + \beta_5 |\Delta CS_{it}| + \beta_6 AccountAge_{it} + \epsilon_{it} \quad (1)\end{aligned}$$

The coefficient of interest is β_3, the interaction between $\mathbb{1}\{LeftDigitChanged\}_{it}$ and $\mathbb{1}\{\Delta CS_{it} < 0\}$. A negative β_3 means that users whose score drops from above the hundred threshold to below the threshold are less likely to check their report in the next month, compared to a user whose score decreased but did not cross the hundred threshold. We control for the absolute change in score, since users whose scores drop by 10 points may be

[12] This method is similar to Bhattacharya et al. (2012), which utilizes left-digit bias and different price paths in stock prices to analyze buy-sell imbalances.

[13] Recall that the exact change in score is salient to users, as user's change in score since their last credit report is shown on the home page, as seen by the red -3 in Figure 1. If the score has increased it will have a "+" and be in green.

Table 3: OLS Estimates of Equation 1

	Dependent variable:
	$\mathbb{1}\{CheckNextMonth\}$
$\mathbb{1}\{LeftDigitChanged\}$	-0.008^{***}
	(0.001)
$\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$	-0.009^{***}
	(0.002)
$\mathbb{1}\{\Delta CS < 0\}$	-0.020^{***}
	(0.001)
$\lvert\Delta CS_{it}\rvert$	0.0001^{***}
	(0.00002)
$AccountAge$	0.001^{***}
	(0.0001)
Day of Week FE	Y
Month FE	Y
Rounded CS FE	Y
Observations	854,329
R^2	0.024
Note:	*p<0.1; **p<0.05; ***p<0.01

Notes: The sample consists of users with credit scores within 10 points to the nearest hundred (scores between X90 and Y09). Both specifications include controls for the rounded credit score to the nearest hundred, month, and day of week.

different than those whose scores drop by 50 points. Additional controls include account age, time fixed effects for when a user checks her score (month fixed effects (α_m), weekday fixed effects (α_w)), and the credit score rounded to the nearest hundred ($RoundedCS_{it}$).

Table 3 displays the OLS estimates of Equation 1. The coefficient for $\mathbb{1}\{\Delta CS_{it} < 0\}$ is negative, meaning that those with declining credit score trends in this sample are less likely to check their updated score in the next calendar month. Individuals with declining credit scores are less likely to check their credit score next month if their score passes the 00 threshold compared to individuals with non-declining scores, as evidenced by the statistically significant negative coefficient for $\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS_{it} < 0\}$. In other words, a user is less likely to check her report next month if her score drops to 499 than if it drops to 500, which means that a larger perceived decline in credit score leads to a lower demand for information. The results are robust to seasonality controls and different bandwidths, as shown in Tables 30 and 31 in Appendix G.1.[14]

To provide a sense of the magnitude of these effects, the average value of $\mathbb{1}\{CheckNextMonth\}$ for users in this sample is 65.9%. If a user's credit score decreases, the probability that she checks her report in the next calendar month decreases by 3%, or by two percentage points (coefficient of $\mathbb{1}\{\Delta CS < 0\}$). However, if her score decreases *and* her score crosses the hundred threshold, she is 5.6%, or 3.7 percentage points, less likely to check her score next month compared to a user whose score is not decreasing. In other words, the incremental difference of having the hundreds digit change is 2.6%. As follows, the marginal effect of crossing the hundred threshold in addition to a score decreasing, is the user is 45% less likely to check her score, compared to a user whose score decreased but did not cross the hundred threshold.[15]

Table 4 displays the heterogeneity of the effect of the credit score trend, in which each column is the estimates of Equation 1 for a single hundred cutoff. Table 4 shows that incremental effect of a left digit change for users with decreasing trends is largest for users with credit scores close to 500 (Column 2), and the incremental effect decreases as credit score increases, but remains statistically negative even for those with credit scores around 700. Therefore, this suggests that the effect of the credit score change is heterogeneous - receiving "bad news" causes users with low credit scores to avoid information the most.[16] Table 4 also includes the average rate of checking by each credit score group, the percentage decrease on checking when the score decreases and crosses a threshold, and the marginal effect of crossing the hundred threshold in addition to a score decreasing, as described in the text for Table 3.[17].

A potential concern with this identification strategy is the possibility that individuals have more options on this website, such as applying for loans or credit cards, if they have

[14]Section Appendix A presents an alternative method to measuring the effect of credit score trend on information avoidance using panel data. The direction of the effects are the same as the effects from the regression discontinuity method.

[15]This calculation is done by dividing the coefficient of $\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$ by $\mathbb{1}\{\Delta CS < 0\}$ (0.009/0.02).

[16]Note that although the estimated coefficient of $\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$ is small for users with credit scores around 400 (Column 1), it is noisily measured since few users have credit scores in that range, shown in Figure 2.

[17]For users in the 800 credit score group, they are more likely to check their score. The marginal effect can be interpreted as 17.6% more likely to check their score if their score crosses the hundred threshold and decreases. This can be seen since the $\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$ coefficient is positive

Table 4: Robustness Checks: RD Results for Each 100 Cutoff

	Dependent variable:				
	$\mathbb{1}\{CheckNextMonth\}$				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\{LeftDigitChanged\}$	-0.0003	0.012^{***}	-0.004^{**}	-0.011^{***}	-0.029^{***}
	(0.017)	(0.004)	(0.002)	(0.003)	(0.004)
$\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$	-0.005	-0.023^{***}	-0.009^{***}	-0.009^{**}	0.005
	(0.023)	(0.005)	(0.003)	(0.004)	(0.007)
$\mathbb{1}\{\Delta CS < 0\}$	-0.034^{**}	-0.006^{*}	-0.022^{***}	-0.029^{***}	-0.029^{***}
	(0.014)	(0.003)	(0.002)	(0.003)	(0.003)
$\lvert\Delta CS_{it}\rvert$	0.001^{***}	0.0002^{***}	0.0001^{**}	-0.0002^{***}	-0.001^{***}
	(0.0001)	(0.00005)	(0.00004)	(0.00005)	(0.0001)
$AccountAge$	0.004^{***}	0.002^{***}	0.001^{***}	0.001^{***}	0.002^{***}
	(0.001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Rounded CS	400	500	600	700	800
Observations	10,979	176,132	342,238	208,534	115,919
R^2	0.027	0.006	0.010	0.016	0.019
Avg. Prob.	0.478	0.536	0.701	0.720	0.669
% dec. cross	8.27	3.63	3.94	5.21	5.01
Marginal Effect %	15.6	384.9	42.1	31.1	-17.6
Note:					*p<0.1; **p<0.05; ***p<0.01

Notes: This table presents the OLS estimates of Equation 1 for each 100 cutoff. The bandwidth size is 10. For example, Column 1 reports only observations with credit scores between 390 and 409. Each specification includes day of week and month fixed effects. The row labeled "Avg. Prob." represents the average probability of checking one's credit score in each subsample. The row labeled "% dec. cross" is the percentage effect of the score decreasing and crossing the threshold on checking. The row labeled "Marginal Effect" is the marginal effect of crossing the threshold combined with the score decreasing on the probability of checking one's credit score.

a credit score above a certain threshold.[18] For example, individuals with a credit score over 700 may qualify for better mortgage rates, which the consumer can shop for on the website. We believe this is unlikely to be the case, since most lenders use the FICO credit score, which differs slightly from the VantageScore credit score.[19] An individual with a VantageScore credit score of 700 does not necessarily have a FICO score of 700, as the two

[18] Note that for some credit cards and loans there could be VantageScore 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. Therefore, 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 on in order to apply for credit cards and loans. However, we expect this to only affect a small percentage of users, since only 0.28% of actions on the website are applications for credit.

[19] `https://www.experian.com/blogs/ask-experian/what-does-my-credit-score-need-to-be-to-get-approved-for-a-mortgage/`

credit score are computed using different formulas.[20] Additionally, we find different effects for individuals with similar credit scores but different trends. In other words, an individual whose credit score increased to 701 has a different likelihood of checking her report next month than another individual whose score decreased to 701, which cannot be explained by having a different set of available loans or credit cards. Lastly, we note that the mechanisms behind these results are consistent with both (i) information avoidance, in which users may believe that a decrease in credit score this month may signal that their credit score will continue to drop next month, and (ii) demotivation. However, the results are not able differentiate between the two.

5 Causal Impact of Receiving Information

In this section, we explore whether information avoidance can be beneficial in terms of improving a user's credit score. Consider the following example: a user checks her credit score and sees that her score has decreased. The previous section shows that she is more likely to avoid checking her updated score than if her score had improved. This leads to a follow-up question: how does avoiding information about her credit affect her credit score, as opposed to the counterfactual in which she receives information about her updated credit report?

Specifically, we measure the heterogeneous effect of information on future credit scores for users who had received positive (a non-declining credit score) or negative (declining credit score) information. The correlation between whether a user checks her credit score and how her credit score changes is not enough to causally measure the effect due to the endogeneity of when a user checks her score, as shown in the previous section. For example, a user may be more likely to check her report when she take actions to improve her score, which would result in a selection bias. Therefore, to measure the causal impact of information on future credit scores, we need variation in whether a user checks her credit report that is not correlated with factors that may affect her score.

To identify the causal effect of information, we use the variation in the likelihood of a user checking her score generated from email experiments that the firm runs. The identification strategy comes from the fact that some users receive emails that are more effective than others at getting the user to check her credit report. Thus, email copies are instruments that shift the likelihood that the user checks her report. This type of analysis is in lines

[20] For more information on how the scores are calculated for Vantage 3.0 and FICO, `https://www.experian.com/blogs/ask-experian/the-difference-between-vantage-scores-and-fico-scores/`, (accessed August 2020)

with an "encouragement design" in which participants are not made to take the treatment (in this case checking one's credit score) rather just encouraged or nudged to do so. In the following section, we describe how the firm implemented these email experiments, the identification strategy and the data used in this analysis.

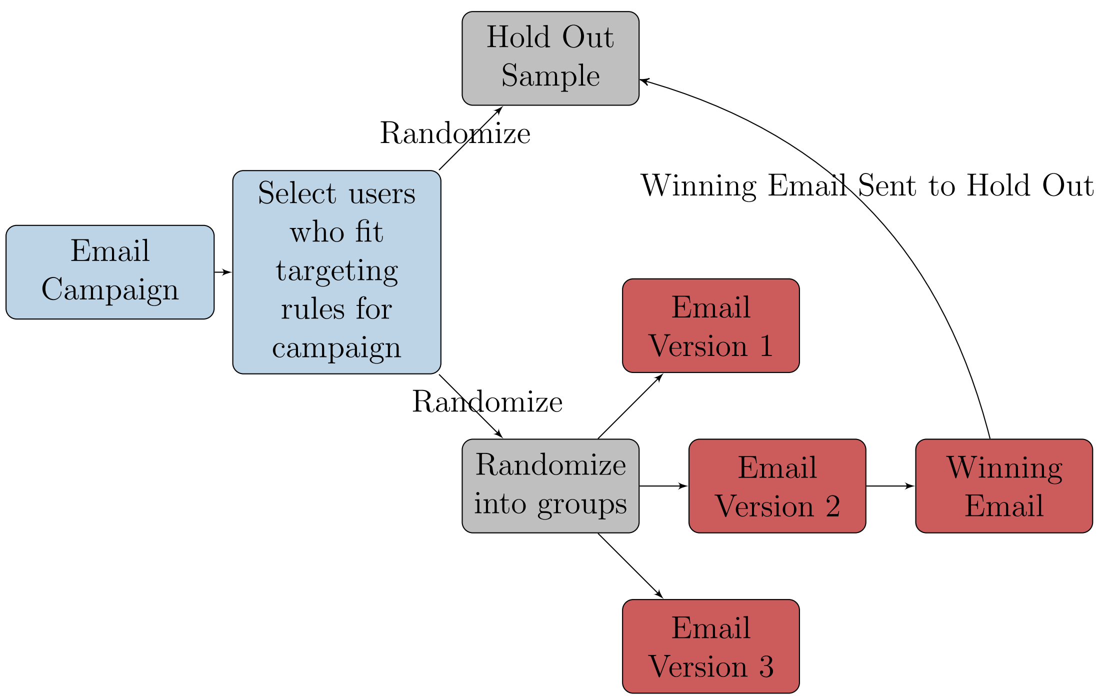
5.1 Overview of Email Experiments

One way that the firm can influence whether a user logs on to her account and checks her report is by sending an email campaign. A "campaign" refers to a type of email message that the firm sends to a large audience base. Many of these email campaigns notify users that their credit information has been updated. 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?" A campaign consists of one theme, but to determine the most effective messaging, the firm A/B tests multiple versions, which we will refer to as email copies. First, the firm selects which users will receive a campaign based on a set of targeting rules. The firm selects a proportion of users to include in the email experiment and within the targeted users they randomize users into a test group. The remainder of users are saved for a hold-out sample. Then, they further randomize all users in the test group into separate treatment groups and each treatment group receives a different email copy. For example, if the email campaign's message is to let a user know that her credit score has been updated, email variants within the campaign may have subject lines such as: "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 hold-out sample. Between 2014 and 2017, the firm tested thousands of campaigns. This process is depicted in Figure 5.

5.2 Email Copies as Instruments

Before detailing the data sample, we describe the identification strategy, and more specifically, how we use the email copies as instruments. The question we aim to answer is, if a user checks her credit report, how does that affect her future credit score? Furthermore, how does this effect differ for those users who had declining credit scores prior to checking their reports (i.e., users who were more likely to avoid information) compared to those users who had non-declining credit scores (less likely to avoid information)? The ideal experiment to answer this question would involve a 2x2 design, in which users are split into two samples, based on whether they had increasing or decreasing credit score trends. Within

Figure 5: Email Campaign Process



Notes: This depicts the process of how an email campaign is implemented and how users are split into groups.

each group, one half would receive the treatment, which is information about their updated credit report (treatment group), and the other half would not receive the treatment (control group). Then, within each group, the treatment effect would be measured by comparing how credit scores changed in the next month (when credit scores are updated) to before treatment (or lack of treatment) assignment for users in the treatment versus control group.

The main challenge in adapting our setting to the ideal experiment is the endogeneity of when users decide to check their credit report. The implications of this challenge are the following. First, we do not have random assignment in which users check their credit report. Second, if a user does not log on to the website in a given calendar month, we do not observe her credit report, leaving us with missing outcome data.

We address the first implication by using email copies as instruments for whether the user checks her credit report in a given month. The campaign email experiments test emails of varying effectiveness, thus inducing variation in whether the user logs on to her account and checks her updated credit report.[21] We will refer to the emails that we use in the following analysis as a "focal" email, as the users also receive other emails during this time that are not part of the IV analysis. The data used in this analysis comes from only users who fit the targeting rules for a campaign but are not in the hold-out sample. Thus, within an email campaign, the email copy that a user receives is randomly assigned. In the following section, we show that these emails are valid and not-weak instruments. We address the second implication by conducting robustness checks that involve holding constant the length of time between checking credit reports and imputing the missing data.

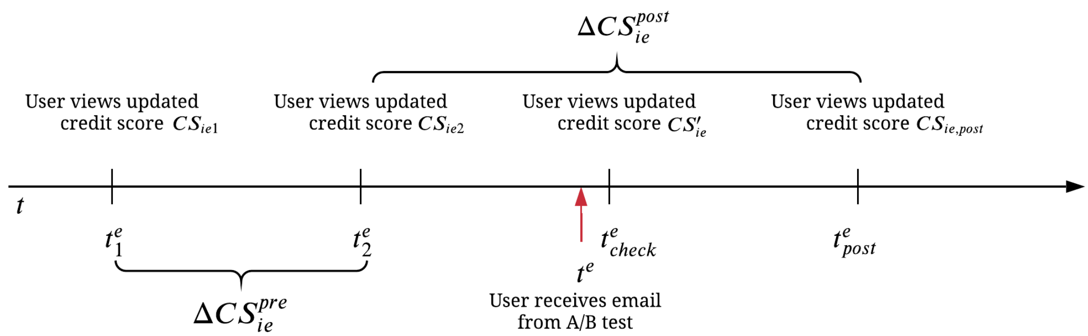
5.2.1 Variable Creation and Timeline of User Events

Figure 6 shows the timeline of events for a single user, which illustrates how the variables are created. As a reminder, we only observe credit scores when a user logs on that calendar month, meaning the length of time between credit score updates will vary by each user with a minimum of one calendar month in between credit score updates. Therefore, t denotes not a calendar month, but the month when a user checks her credit report. We define a user checking her score as when she logs on for the first time in that particular month. While the user can log on additional times, she sees the same credit score from earlier that month and thus does not learn new information. Therefore, the results should be interpreted as the effect of learning "new" information.

Let t^e be the time at which the focal email is sent to the user i. We construct the

[21]Figure 13 in Section G.2 in the Appendix, which plots the histogram of click rates for each email copy conditional on opening, shows that there is varying effectiveness in email copies. Click rates, conditional on opening, range from 10% to 53%.

Figure 6: Timeline for Variable Construction



Notes: Each observation is a user-email (ie) combination. Time periods are relative to the time that email e is received. Variable notation and definitions are described in Table 5.

Table 5: Explanation of Time Indices and Variable Construction

Name	**Variable Construction**
t^e	Time that the email e is sent
t^e_{check}	Time that the user checks her report in response to receiving $e^\dagger$
CS_{ie2}	i's last credit score before receiving the email at t^e
CS_{ie1}	i's last credit score before observing her score at $t = 2$
CS'_{ie}	i's credit score in the month after she receives the email†
$CS_{ie,post}$	i's credit score the first time it is checked at least two months after the email was sent
ΔCS_{ie}^{pre}	$CS_{ie2} - CS_{ei1}$: i's change in score before receiving e
ΔCS_{ie}^{post}	$CS_{ie,post} - CS_{ie2}$: i's change in score at least two months after receiving e
$\mathbb{1}\{Check_{ie} = 1\}$	Indicator for whether a user checks her score within 31 days after t^e

Notes: †: Only observed by i if she checks her report within 31 days after receiving email e.

"pre-trend" variable by comparing the last two credit score updates before the email was sent, at time t_2^e and t_1^e respectively.[22]

$$\Delta CS_{ie}^{pre} = CS_{ie2} - CS_{ei1} \quad (2)$$

We then consider whether the user checked her updated report within one month of the email being sent. If she checks within this time window, we classify that user as having checked her report ($\mathbb{1}\{Check_{ie} = 1\}$). Note that the credit report check at t^e_{check} only exists

[22]The pre-trend is defined in the same way as in Section 4.

for users who check their report within 31 days of receiving email e.

The outcome of interest is how a user's score changes after checking her report. To allow the user time to take action to change their score, we use the first credit score check that occurs at least two months after the email was sent to construct the ΔCS_{ie}^{post} variable. For example, if a user checks her report two days after the email was sent, one month after the email was sent and two months after the email was sent, the score at two months after t^e will be used in constructing the ΔCS_{ie}^{post} variable for that user. As another example, if a user does not check her score for four months after the email was sent, we use the credit score at the four month post email check. We use two months after, rather than just the first observed credit score after, in order to allow the user time to change their behavior and thus change their credit score.[23] Note that the post-trend variable covers different time windows for each user, which we account for in robustness checks. We define the post-trend as:

$$\Delta CS_{ie}^{post} = CS_{ie,post} - CS_{ie2}. \tag{3}$$

5.2.2 Two-Stage Least Squares

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.[24] Instead, we pool campaigns and use each email copy as an instrument, since there is significant variation in click rates in email copies across campaigns. Email copies will henceforth referred to as e. However, while the assignment to which email copy within a campaign is random, the assignment to an email campaign broadly is not randomized. Thus, there is a concern of potential correlation between email campaign click rates and users' credit scores. To control for these correlations, we obtain the email targeting rules the firm used for each campaign in this data set. The next section describes this in detail.

We assume that conditional on targeting rules, the email copy that a user receives only affects her future credit score through her propensity to check her credit report. The findings in Section B in the Appendix are consistent with this assumption. Specifically, Appendix Section B shows that 1) within each campaign, the email variants are not correlated with the users' credit scores or 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 or trend in credit score.

[23] Additionally, we choose two months to avoid cases in which only a few days elapses between t^e_{check} and t^e_{post}. This gives all users at least 31 days between t^e_{check} (if she checked her report) and t^e_{post}.

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

To estimate the heterogeneous effect of information based on the user's likelihood of avoiding information, we estimate the following 2SLS regressions separately for users with declining and non-declining credit scores prior to receiving the email. If we estimated a regression that combined the two samples, we would need to estimate the interaction variable $\mathbb{1}\{Check_{ie}\} \times \mathbb{1}\{\Delta CS_{ie}^{pre} < 0\}$ to measure the separate effects. However, this interaction variable is also endogenous and thus would require additional instruments, so we estimate the coefficients separately for each sample.

For each sample, the first stage of the IV is the following:

$$\mathbb{1}\{Check_{ie}\} = \alpha_e + \gamma X_{ie} + \eta_{ie}, \tag{4}$$

where α_e is the fixed effect for email copy e. X_{ie} is a matrix of i's covariates used in the targeting rules at time that e was sent.[25] Since we observe all the variables 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 in Table 36 in Section G.2 in the Appendix, which also includes a list of all the characteristics used for campaign targeting.

The second stage of the regression is:

$$\Delta CS_{ie}^{post} = \beta_0 + \beta_1 \mathbb{1}\{\widehat{Check_{ie}}\} + \delta X_{ie} + \epsilon_{ie}, \tag{5}$$

where $\mathbb{1}\{\widehat{Check_{ie}}\}$ are the fitted values from the first stage and X_{ie} is again the matrix of i's characteristics at the time email e was sent. We also repeat the regressions with the average monthly change in the post-trend as the outcome, $\overline{\Delta CS_{ie}^{post}}$, which is ΔCS_{ie}^{post} divided by the number of months between t_{ie2} and $t_{ie,post}$.

$$\overline{\Delta CS_{ie}^{post}} = \beta_0 + \beta_1 \mathbb{1}\{\widehat{Check_{ie}}\} + \delta X_{ie} + \epsilon_{ie}. \tag{6}$$

The following sections describe the data used to estimate these regressions, along with evidence that the instruments are valid and not weak.

5.3 Data for IV

5.3.1 Campaign Selection

To maximize the variation in credit report checks resulting from the emails, we use data from the top 25 campaigns that had the highest variation in click rates across email copies

[25]The targeting rules are defined in Table 36.

Table 6: Summary Statistics of the Email Copies

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
N Users (1000s)	3.07	4.76	5.64	21.72	11.50	179.21
Percent Opened	12.93	30.70	36.23	36.73	43.98	52.66
Percent Clicked	9.74	26.55	31.59	32.83	39.48	53.40

Notes: Each observation is at the email copy level. The first row is the number of users who receive the email copy. The second row is the percentage of users who open the email. The third row is the percentage of users who click on the email, conditional on opening. The emails included here are from the 16 campaigns we use in the following IV analysis.

and were sent to users before June 1, 2017.[26] We obtained targeting rules from the firm for 18 of the 25 campaigns. Additionally, we removed two campaigns had email copies that contained text that may violate the exclusion restriction; the subject lines contained information about the user's credit.[27] The final data set for this analysis has 16 campaigns. Table 34 in Section G.2 in the Appendix displays the subject lines for all email variants in these campaigns. There are a total of 70 different email copies, and 531,232 users on the platform received at least one of these email copies.

Table 6 provides summary statistics for the email copies in this sample. Each observation is at the email copy level. The first row is the number of users who receive the email copy, the second row is the percentage of users who open the email, and the third row is the percent of users who click on the email, conditional on opening. Each email copy is sent to between 3,000 and 180,000 users, has an average open rate of 37%, and has an average click rate of 33%, conditional on opening. The engagement rate with these emails is higher than the emails described in Table 2, which are the set of all emails that a randomly selected user receives. We believe this is because users generally pay more attention to "update" emails as compared to other types of emails, since emails that contain the words "update" or "changed" have higher open and click rates compared to emails without these words, as shown in Table 32 in Section G.1 in the Appendix.

5.3.2 Final Data Set

The final data set comprises 610,483 user email-level observations. This is a set of all users who meet all of the following criteria: 1) have received at least one of the email copies in

[26]By choosing a cutoff a year before the data observation window ends, this allows us to give time for users to return. More details in Section 5.3.2

[27]Since it is possible that the text of these subject lines may directly impact a user's credit score beyond encouraging them to check their report, we remove these types of emails. 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.

the selected 16 campaigns; 2) have checked their credit report at least twice before receiving the email, so that we can observe the pre-trend; 3) checked their credit report within three months prior to receiving the email ($t^e - t_2^e < 3$ months) so that users are likely to remember their pre-trend before checking their report again; and 4) have checked their credit report at least two months after receiving the email before the end of the data time window. To illustrate the scope of this sample, 38% of users in the random sample fit these criteria. Given that the platform has over 15 million users, this represents over 5.7 million users.

The users in this sample have an average credit score of 606 with a standard deviation of 91.7, which is slightly higher than the average of 598.5 from the random sample of users. In other matters, such as how their credit score changes, how often they check their reports, and whether they avoid information after receiving negative feedback, they resemble the random sample of users, as shown in Tables 24 and 26 in the Appendix. Table 27 in Section E.3 in the Appendix also compares users who are in the IV sample with those who would have been except that they are missing the outcome (i.e. fulfilled the first three criteria). Their pre-trend changes and number of months between the pre-trend checks have similar medians. Users in the random sample sample on average have slightly larger pre-trend changes (2.31 points versus 1.44) and wait longer between score checks (2.73 months compared to 1.68).

Furthermore, we compared users by pre-trend. The time periods between checking the report, receiving the focal email, and the time it takes to return ($t^e - t_2^e$ and $t^e_{post} - t^e$, respectively) are similar between users with declining and non-declining trends, as shown by Table 33 in the Appendix.

Among users who received one of the 16 email campaigns that we focus on, 84% received only one of these email campaigns. Additionally, only 2.7% of users received a second focal email within the same calendar month.[28]

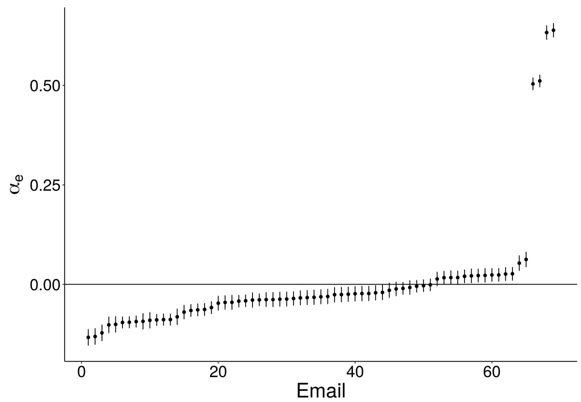
5.3.3 Instrument Validity and Strength

In order for the email e to be a valid instrument, e must be correlated with whether i checks her credit report and must satisfy the exclusion restriction. To demonstrate the former, we show that after controlling for the targeting rules, the email copies themselves explain variation in $\mathbb{1}\{Check_{ie} = 1\}$ by estimating the following regression:

$$\mathbb{1}\{Check_{ie}\} = \alpha_e + \gamma X_{ie} + \epsilon_{ie}. \tag{7}$$

[28]Users in this sample receive other emails from the firm as well. Within a week before to a week after the IV email was sent, users received an average of 10 emails from the website with a standard deviation of 7.7.

Figure 7: Estimated Email Fixed Effects



Notes: This is a plot fixed effects of each email copy on whether or not the user checks her credit report in the month after receiving the email. The y-axis is the fixed effect of a given email, ordered from the smallest to largest fixed effect. Error bars are the 95% confidence intervals.

If the email copy has no explanatory power, we would expect the email copy fixed effects α_e to be close to zero for all email copies. However, we find a wide range and statistically significant estimated fixed effects. Figure 7 displays the estimated fixed effects from Equation 7. Additionally, we verify that the email copies are not weak instruments, since the incremental F-statistic is 25 for the declining pre-trend sample, and 187 for the non-declining pre-trend sample. Details are in Section B.1 in the Appendix.

In this context, the exclusion restriction is that the email cannot affect a user's credit score except through getting a user to check her credit report. We assume the exclusion restriction to be true; below, we illustrate potential ways this restriction is violated, and provide justifications for why this assumption may hold.

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 this is unlikely in this context, given that the emails also do not contain prescriptive advice on how to change a credit score.

Another way that emails could violate the exclusion restriction is if they serve as a reminder that a user's 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 on other platforms, this firm sends a large number of emails to users who have not opted out. As shown in the summary statistics, users in this analysis receive almost 31 emails per month. Therefore, the focal emails in these campaigns are unlikely to serve as reminders about the existence of their credit score, since users are receiving many other emails from

Table 7: Intent-to-Treat Effect

	Dependent variable:	
	ΔCS^{post}	
	(1)	(2)
Click Rate	-1.591^{**}	1.167^{*}
	(0.677)	(0.667)
ΔCS^{pre}	< 0	≥ 0
Campaign Targeting Controls	Y	Y
Observations	223,913	386,570
R^2	0.017	0.022
Note:	*p<0.1; **p<0.05; ***p<0.01	

Notes: This table presents the OLS estimates of Equation 8. Each observation is at the user-email copy level. All columns include controls for the targeting rules. Standard errors are clustered at the email level.

the firm.

Lastly, because we observe credit scores of only those who log in, the exclusion restriction is violated if the instruments directly impact whether the user returns at t^e_{post} or the composition of users who return to the platform. In other words, the worry is that the instrument leads to selection on the outcome. To provide evidence that this assumption is reasonable, we show that the instrument does not impact average return rates and conduct a robustness check that imputes missing credit scores, in which we find similar effects. More details can be found in Section 6.4. However, we cannot conclusively rule out the possibility that the instrument does not impact the composition of users who return, and therefore assume that the instruments do not differentially impact retention.

Additionally, we assume monotonicity (no defiers), meaning that there are no users for whom receiving the email causes them to be less likely to check their credit report.

6 Results

6.1 Effect of checking information on credit score

Before presenting the local average treatment effect recovered by the IV, we first report the intent-to-treat effect. Instead of estimating the coefficients of each email on the change in credit score, we estimate the coefficient on the effectiveness of the email, proxied by the email's click rate. Assuming that more effective emails have higher click rates, we estimate the effect of receiving a more effective email on changes in credit score with the following regression:

$$\Delta CS_{ie}^{post} = \alpha + \beta ClickRate_e + \delta X_{ie} + \epsilon_{ie}. \tag{8}$$

Table 7 displays the OLS estimates of Equation 8. The first column reports the estimates for those with declining credit scores, and the second for those with non-declining credit scores. We see that for users who had declining scores, receiving a more effective email leads to a statistically significant 1.5 decrease in credit score. However, for users who had non-declining scores, receiving a more effective email leads to an increase in credit score. Table 35 in the Appendix Section G.2 displays the heterogeneous effects of the intent-to-treat effect by users' credit scores prior to receiving the email. We find that for users with declining credit scores, the negative effect is driven by users who have credit scores in the bottom tercile of the sample (below 578). For users with non-declining scores, users in the middle tercile (between 578 and 661) experience the largest increase in credit score in response to receiving an effective email, but the magnitudes are similar across all terciles.

Next we present the local average treatment effects estimated by the IV analysis. Table 8 displays the two-stage least squares estimates of Equation 5. 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 non-declining credit scores prior to checking. The coefficients for the campaign control variables are omitted for better readability.[29] We find that for users who are on declining credit score trends prior to receiving the email, checking their credit reports caused their future credit scores to drop substantially, by 23 points. On the other hand, for users with non-declining credit scores, the credit scores increase by nine points if they check their credit reports. Additionally, we find that these effects are robust to the length of time that the user takes to return to check their report, as shown by Columns 3 and 4, in which the dependent variable is the average monthly change in credit score.

6.2 Robustness Checks

This effect is robust to various specifications. The estimates in Table 8 are only for users who have checked their report within three months prior to receiving the focal email, but we find that the effect still holds across users who have checked more and less recently. In Section G.2 in the Appendix, we estimate Equation 5 for users who logged in more recently

[29]The full table, with the campaign control coefficients, is in Table 37 in Section G.2 in the Appendix.

Table 8: Two-Stage Least Squares Estimates of Equation 5 and Equation 6

	ΔCS^{post}		$\overline{\Delta CS^{post}}$ monthly	
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-23.247***	9.084***	-4.409**	1.436***
	(6.928)	(2.094)	(2.108)	(0.479)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Observations	223,913	386,570	223,913	386,570

Notes: This table shows the 2SLS estimates of Equation 5 and Equation 6. Targeting characteristic coefficients are omitted for readability. The full regression can be found in Table 37 in the Appendix. Each observation is at the user-email level. Standard errors are clustered at the email level.

(one and two months) and less recently (six months). The coefficient estimates of $\mathbb{1}\{Check\}$ share the same sign and are similar in magnitude across all samples.

As noted, some users received more than one focal email in the instrumental variable analysis. We repeat the regression removing any users who received more than one focal email and again for users who received more than one focal email in the frame of one month. The results are qualitatively similar in the specification that removes all users who received more than one focal email, but the effect is no longer statistically significant for users with declining pre-trends due to the large decrease in sample size. When only users who have received multiple focal emails in the same calendar month are dropped, the effect for users with declining pre-trends remains negative and statistically significant. The coefficient estimates are displayed in Tables 41 and 42 in Section G.2 the Appendix. In Tables 47 - 54 Appendix G.3 we also estimate other specifications of the IV in which we segment users by covariates.

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

The above results show that on average, credit report information does not help improve credit scores for users experiencing a downward trajectory in their credit scores. But is this effect mainly driven by users who have low credit scores? To test this, we estimate the 2SLS regression for users who had poor, medium, or good credit scores before receiving the email in Table 9. We classify users into these bins based on the tercile of their last credit score prior to receiving the email. In this data, the 33 and 66 percentile of credit scores are 578

Table 9: 2SLS IV Splitting Users by Credit Score Tercile

	ΔCS^{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Check\}$	-21.521***	-0.504	-12.152**	-4.223**	19.629***	21.602***
	(6.618)	(7.584)	(6.134)	(2.118)	(2.839)	(2.483)
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
Tercile	Bottom	Middle	Upper	Bottom	Middle	Upper
ΔCS^{pre}	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
Observations	51,185	81,078	91,650	71,803	134,781	179,986

Notes: This table presents the 2SLS results by credit score tercile. The bottom tercile includes users with credit scores below 578, the middle tercile includes users with credit scores between 578 and 661, and the top tercile includes users whose credit scores are above 661. 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. Standard errors are clustered at the email level. Controls for email targeting are omitted from output for better readability. The full regression output is in Table 45 in the Appendix.

and 661, respectively. Users in the "poor" credit score bin have a credit score below 578, while "medium" scores are between 578 and 661, and "good" scores are greater than 661.

Users in the top and bottom terciles, who were also on a declining credit score trend prior to the focal email, still have a significant and negative additional credit score drop, with the lowest tercile having the biggest drop of over 21 points. For users with non-declining credit score trends, users in the top two terciles have a significant positive increase in credit score upon checking their score. However, users with non-declining trends who had credit scores in the bottom tercile experience a statistically significant drop in credit score, but not as large as the drop for users with similar credit scores who had declining trends. Table 10 reports the same regression as Table 9 but with the monthly change in credit score as the outcome. We see similar patterns overall, with the following differences: the upper tercile for users on declining pre-trends is still negative but no longer significant, and the effect of checking is statistically significant and positive for all users in the non-declining sample.

6.4 Discussion and Limitations

Overall, we find 1) users avoid information after receiving the unwanted information that their credit score has decreased, and 2) information further decreases credit scores for those

Table 10: 2SLS IV Splitting Users by Credit Score Tercile - Monthly Credit Score Change As Outcome

	$\overline{\Delta CS^{post}}$ month					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Check\}$	-3.795**	1.450	-1.995	5.047***	2.190***	1.262***
	(1.853)	(2.365)	(1.912)	(0.658)	(0.584)	(0.481)
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
Tercile	Bottom	Middle	Upper	Bottom	Middle	Upper
ΔCS^{pre}	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
Observations	51,185	81,078	91,650	71,803	134,781	179,986

Notes: This table presents the 2SLS results of Equation 6 by credit score tercile. The outcome is the monthly credit score change. The bottom tercile includes users with credit scores below 578, the middle tercile includes users with credit scores between 578 and 661, and the top tercile includes users whose credit scores are above 661. 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. Standard errors are clustered at the email level. Controls for email targeting are omitted from output for better readability.

whose credit was on a declining trend. These findings together suggest that information avoidance may help users improve their credit score. Additionally, those with low credit scores are more negatively impacted by bad news, which is consistent with the existing literature showing that bad news has the largest impact on individuals with the lowest performance (Goulas and Megalokonomou, 2015). However, sample selection and the non-random missing outcomes affect the generalizability of these results. Below, we discuss these points and how we address them.

Given the way this data set must be constructed to conduct the IV analysis, the IV results are estimated on a selected sample of users that are not representative of the general population. The sample consists of users who do not opt out of emails and those who have checked their credit report at least three times. In Tables 23-27 in the Appendix E, we show that while users in the selected sample engage in information avoidance similar to the general population, the selected sample is generally more engaged with the platform. Although the sample is unique, users that fit this sample selection criteria represent 38% of the firm's general population, or 5.7 million users.

Additionally, due to the way that the firm updates credit reports, we do not observe credit reports for individuals who do not log on in a given month. The severity of the implications of this problem increases with differential attrition between users with declining and non-declining trends, and the impact of the emails on attrition. However, in this setting, the observed attrition rates between users with declining pre-trends and non-declining pre-trends are similar, as shown in Section C in the Appendix. Furthermore, we do not find evidence that the instrument impacts retention in this setting. This lack of evidence reduces, but does not rule out, selection bias.[30] To further address selection bias, we treat this issue as a missing data problem and impute the missing credit scores. We impute assuming that the data is missing at random and also assuming that the users who do not return tend to have lower credit scores, and find that the results still hold. Details of the imputation and the results can be found in Section C.3 in the Appendix.

Lastly, similar to most instrumental variable analysis, the results from the IV are interpreted as a local average treatment effect (LATE) (Angrist and Imbens, 1995). In other words, the estimated effects are relevant only for compliers, which are users who are induced by more effective email copies to check their report. There may be a subset of users for whom these emails are not enough to overcome their information avoidance to which our results do not apply. This is a limitation of this paper. However, we believe the LATE may be the effect of interest to the firm, as only users who are responsive to the firm's emails (compliers) will be impacted by the firm's email campaigns. In other words, the firm's emails can only impact behavior of those who are swayed by the email to view their credit report. This is similar to Wolak and Korolev (2019), in which the average treatment effect is not relevant as there is no such treatment that would get all participants to comply. Additionally, we find that approximately 30% of the users in this sample are compliers.[31] As 38% of users in the general population fulfill the criteria for the IV sample, if we assume similar rates of compliers in the overall population of users of the platform, the LATE results apply for 1.7 million users of the platform.[32] Section D in the Appendix reports details on how the compliance rate is determined.

6.5 Managerial Implications

The effects of information suggest that the common policy of sending retention/reminder emails to those who are least likely to return (those with decreasing credit scores) may not be the most effective at increasing retention. A reminder to check their score to these

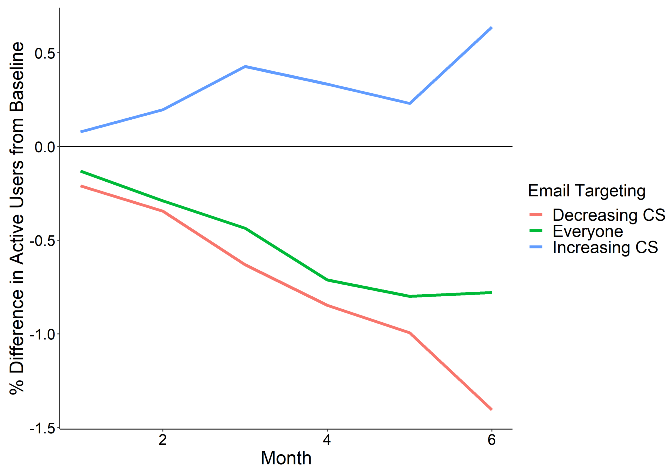
[30]Details in Appendix Section C.1.

[31]Other papers using IV analyses report compliance rates of 10% (Wolak and Korolev, 2019), and 13-88% (Bernstein et al., 2019)

[32]The number 1.7 million comes from the general population of 15 million$\times 0.38 \times 0.3$.

users causes a further decline in score and leads them to be less likely to check their score in the future. To better illustrate the managerial implications, we compare the impacts of several email targeting policies on retention, given the effects of information and information avoidance. Specifically, via simulation, the three targeting policies we consider are sending emails to all users, sending emails only to users with increasing scores, and sending emails only to users with decreasing scores.

Figure 8: Simulation Results: Retention



Notes: Results from one million simulated users. This figure displays the retention metric, which is the percentage difference in the number of users who check their report in a given month compared to the baseline policy where no emails are sent.

We simulate credit scores and credit score checking behavior for one million users. In each time period (month), the user decides whether to check her score. The probability that she checks her score is determined from the regression estimates in Table 3, which account for information avoidance resulting from a decrease in credit score. If a user receives an email at time t, her score at time $t+1$ is adjusted by the intent to treat effects in Table 7. We make some simplifying assumptions in the simulations, such as all emails have a click rate of 100%, which reduces the need to model emails of varying click rates, and that once a user chooses to not check her score in period t, she drops out of the sample for remaining time periods, which allows us to not model a user's belief about her credit score trajectory. Due to these simplifying assumptions, we note that the goal of the simulations is to be able to compare across email targeting policies, as all policies are simulated under the same assumptions, rather than to evaluate the magnitude of the effects. Details of the simulations are reported in Appendix F. Figure 8 displays the effects of different email policies on retention compared to the baseline policy of no emails. The figure confirms that sending reminder emails only to users with decreasing credit scores has the lowest retention rate compared to emailing everyone or only those with increasing scores.[33]

7 Conclusion

This paper studies individuals' demand for information in the context of consumer finance and estimates the heterogeneous causal effect of information on individuals' propensity to avoid information. Using novel data from a website that provides its users with free monthly credit reports, we observe that individuals with low credit scores are the most likely to attrit. Extending prior literature on information avoidance that focuses on avoiding potentially-unwanted information, we measure the causal effect of receiving unwanted information on the propensity to seek information in the future. We leverage left-digit bias, with the assumption that people with credit scores that end in 99 and 00 are very similar, to measure the causal effect of changes in credit score trends on checking one's information in the next month. Controlling for the absolute change in credit score, we find that individuals are less likely to check their report if their score falls below the 00 threshold than if it does not fall below the 00 threshold, indicating that individuals avoid information when they observe credit score declines.

To measure the causal effect of checking information on future credit scores, we instrument for 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 that the firm uses to target campaigns, induces exogenous variation in an individual's propensity to log on to her account and check her credit report. We find that individuals who are most likely to avoid information (those that had a declining credit score prior to checking their report) experience a significant decrease in their credit score after checking. These effects suggest that in information-tracking settings, the common practice of firms sending reminder/retention emails to those who are least likely to return may not be the most effective at increasing retention.

A limitation of this paper is that due to the way that the firm retrieves credit reports, we observe credit reports only for users who log on. Thus, the generalizability of the effect

[33]Note that this simulation assumes that the emails sent at time t do not impact whether the user returns at time $t+1$, similar to the setting in this paper. However, it may be reasonable to assume that the email increases future retention in other settings. Figure 12 in Appendix F shows simulation results for when emails sent at t impact whether the user returns at $t+1$. In that case, retention is highest when emails are sent to everyone, but the common policy of sending emails only to those least likely to return (those with decreasing credit scores) still has the lowest retention rate.

of information on credit scores is limited to returning users who have not opted out of emails. In addition, the instrumental variable analysis can only allow us to measure the local average treatment effect. In other words, the estimated effect of checking information is the treatment effect for compliers, users who are able to be persuaded by the email copy to check their information. Additionally, this paper measures the effect of information on credit scores, but does not explore the separate effects of positive versus negative information on credit scores. Future work on this question can further inform firms on not only which individuals should be targeted based on their historic trends but also on the type of information they receive. Finally, 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. We are not able to explore how the content of the emails can influence the effect of checking information. Perhaps including information about peers or exerting 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, E. T. and D. I. Simester (2003): "Effects of $9 price endings on retail sales: Evidence from field experiments," *Quantitative Marketing and Economics*, 1, 93–110.

Angrist, J. D. and G. W. Imbens (1995): "Identification and estimation of local average treatment effects," Tech. rep., National Bureau of Economic Research.

Bernstein, S., E. Colonnelli, and B. Iverson (2019): "Asset Allocation in Bankruptcy," *Journal of Finance*, 74, 5–53.

Bhattacharya, U., C. W. Holden, and S. Jacobsen (2012): "Penny wise, dollar foolish: Buy–sell imbalances on and around round numbers," *Management Science*, 58, 413–431.

Centers for Disease Control and Prevention (1997): "HIV counseling and testing in publicly funded sites: 1995 summary report," *Atlanta: US Department of Health and Human Services.*

Golman, R., D. Hagmann, and G. Loewenstein (2017): "Information Avoidance," *Journal of Economic Literature*, 55, 96–135.

Goulas, S. and R. Megalokonomou (2015): “Knowing who you are: The effect of feedback information on exam placement,” *Working Paper*.

Huang, S.-C. (2018): “Social information avoidance: when, why, and how it is costly in goal pursuit,” *Journal of Marketing Research*, 55, 382–395.

Huck, S., N. Szech, and L. Wenner (2018): “More effort with less pay: On information avoidance, optimistic beliefs and performance,” *Working Paper*.

Imbens, G. W. and J. D. Angrist (1994): “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62, 467.

Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman (2016): “Getting to the top of mind: How reminders increase saving,” *Management Science*, 62, 3393–3411.

Karlsson, N., G. Loewenstein, and D. Seppi (2009): “The ostrich effect: Selective attention to information,” *Journal of Risk and uncertainty*, 38, 95–115.

Kőszegi, B. (2003): “Health anxiety and patient behavior,” *Journal of Health Economics*, 22, 1073–1084.

Lacetera, N., D. G. Pope, and J. R. Sydnor (2012): “Heuristic thinking and limited attention in the car market,” *American Economic Review*, 102, 2206–36.

Lee, D. S. and T. Lemieux (2010): “Regression discontinuity designs in economics,” *Journal of Economic Literature*, 48, 281–355.

Liu, X., A. Montgomery, and K. Srinivasan (2018): “Analyzing Bank Overdraft Fees with Big Data,” *Marketing Science*, 37, 855–882.

Lyter, D. W., R. O. Valdiserri, L. A. Kingsley, W. P. Amoroso, and C. R. Rinaldo Jr (1987): “The HIV antibody test: why gay and bisexual men want or do not want to know their results.” *Public Health Reports*, 102, 468.

Maestas, N., K. J. Mullen, and A. Strand (2013): “Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI Receipt,” *American Economic Review*, 103, 1797–1829.

Marbach, M. and D. Hangartner (2020): “Profiling Compliers and Noncompliers for Instrumental-Variable Analysis,” *Political Analysis*, 28, 435–444.

Masatlioglu, Y., A. Y. Orhun, and C. Raymond (2017): "Intrinsic information preferences and skewness," *Working Paper*.

Medina, P. C. (2017): "Selective attention in consumer finance: Evidence from a randomized intervention in the credit card market," *Working Paper*.

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

Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus (2015): "Financial attention," *The Review of Financial Studies*, 29, 863–897.

Stango, V. and J. Zinman (2009): "What Do Consumers Really Pay on Their Checking and Credit Card Accounts ? Explicit , Implicit , and Avoidable Costs," *The American Economic Review*, 99.

Sweeny, K., D. Melnyk, W. Miller, and J. A. Shepperd (2010): "Information Avoidance: Who, What, When, and Why," *Review of General Psychology*, 14, 340–353.

Thomas, M. and V. Morwitz (2005): "Penny wise and pound foolish: the left-digit effect in price cognition," *Journal of Consumer Research*, 32, 54–64.

Thornton, R. L. (2008): "The demand for, and impact of, learning HIV status," *American Economic Review*, 98, 1829–63.

Wolak, F. A. and I. Korolev (2019): "Instrumental Variables Estimators and Bounds on Treatment Effects," *Working Paper*.

Appendix

A Additional Evidence for Information Avoidance

Another way to measure the impact of credit score trends on whether the user avoids information is to utilize the panel structure of the data, through which individual-level differences can be controlled for. We estimate the following linear regression for a user i when she checks her report at year-month t, for users in the sample who have checked their reports at least twice.[34] The regression is:

$$\begin{aligned} \mathbb{1}\{CheckNextMonth\}_{it} &= \alpha_m + \alpha_w + \alpha_i + \delta_1 \mathbb{1}\{\Delta CS_{it} < 0\} + \delta_2 |\Delta CS_{it}| \\ &+ \delta_3(\mathbb{1}\{\Delta CS_{it} < 0\} \times |\Delta CS_{it}|) + \delta_4 NChecks_{it} + \delta_5 AccountAge_{it} + \delta_6 CS_{it} + \epsilon_{it}, \end{aligned} \tag{9}$$

where $\mathbb{1}\{CheckNextMonth\}_{it}$ is an indicator for whether the user checks her score in the following calendar month. CS_{it} is i's credit score at time t, and ΔCS_{it} is the change in i's credit score between t and $t-1$ ($CS_{i,t} - CS_{i,t-1}$).[35] $\mathbb{1}\{\Delta CS_{it} < 0\}$ is an indicator for whether the user is currently on a declining credit score trend. $|\Delta CS_{it}|$ is the absolute value of ΔCS_{it}, and $NChecks_{it}$ is how many times the user has previously checked her score. Since there is selection bias in who checks their report, we include individual-level fixed effects α_i. We also control for the user's current credit-score level, account age (in months), and seasonality with α_m and α_w, which are fixed effects for month and day of week, respectively, for when the user checks her report.

Table 11 displays the OLS estimates of Equation 9. Column (3) contains the main specification and columns (1) and (2) show that the results are robust to various specifications. The coefficient for $\mathbb{1}\{\Delta CS_{it} < 0\}$ is negative and statistically significant across all columns, indicating that if a user checks her report and sees that her score has decreased since her last credit report, she is less likely to check her score in the following calendar month. The negative coefficient of $\mathbb{1}\{\Delta CS_{it} < 0\} \times |\Delta CS_{it}|$ in column (3) shows that the greater the decrease in score, the less likely the user is to check her updated report next month. Table 12, which displays estimates for Equation 9 for each tercile of credit score, shows that this effect exists for users in each credit score tercile, indicating that even if the overall credit score is high, the trend in score is relevant for whether a user checks. Tables 28 and 29 in Appendix G.1 include robustness checks with various random and fixed effects removed and for different time intervals for the pre-trend (difference between t and $t-1$). In all specifications, the coefficient for $\mathbb{1}\{\Delta CS < 0\}$ remains negative and statistically significant.

[34]The user needs to have checked at least twice in order to construct the current credit score trend

[35]Note that $t-1$ may not always be the previous calendar month.

Table 11: OLS Estimates of Equation 9

	Dependent variable:		
	$\mathbb{1}\{CheckNextMonth\}$		
	(1)	(2)	(3)
$\mathbb{1}\{\Delta CS < 0\}$	-0.019^{***}	-0.017^{***}	-0.013^{***}
	(0.0004)	(0.0004)	(0.001)
$\mathbb{1}\{\Delta CS < 0\} \times \lvert\Delta CS\rvert$			-0.0003^{***}
			(0.00002)
$AccountAge$	-0.004^{***}	0.006^{***}	0.006^{***}
	(0.00004)	(0.0001)	(0.0001)
CS	0.001^{***}	0.001^{***}	0.001^{***}
	(0.00001)	(0.00001)	(0.00001)
$NCheck$		-0.014^{***}	-0.014^{***}
		(0.0001)	(0.0001)
|CSChange|			0.0002^{***}
			(0.00001)
Person FE	Y	Y	Y
Day of Week FE	Y	Y	Y
Month FE	Y	Y	Y
Observations	4,339,410	4,339,410	4,339,410
R^2	0.343	0.346	0.347
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: Each observation is at the user-credit (it) report level. The sample consists of 561,864 users who have checked their credit report at least twice before May 1, 2018. Users who check their reports in May 2018 are dropped because the data is truncated at June 12, 2018, thus we do not observe whether a user returns. Standard errors are clustered at the individual level. Column (1) omits controls for NCheck and the $|\Delta CS_{it}|$. Column (2) includes the control for NCheck, and Column (3) reports the full specification in Equation 9.

B Email Randomization Check

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

First, we test the randomization of the email copies/variants within a single 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 9 displays a histogram of the p-values from ANOVA tests, in which the null hypothesis

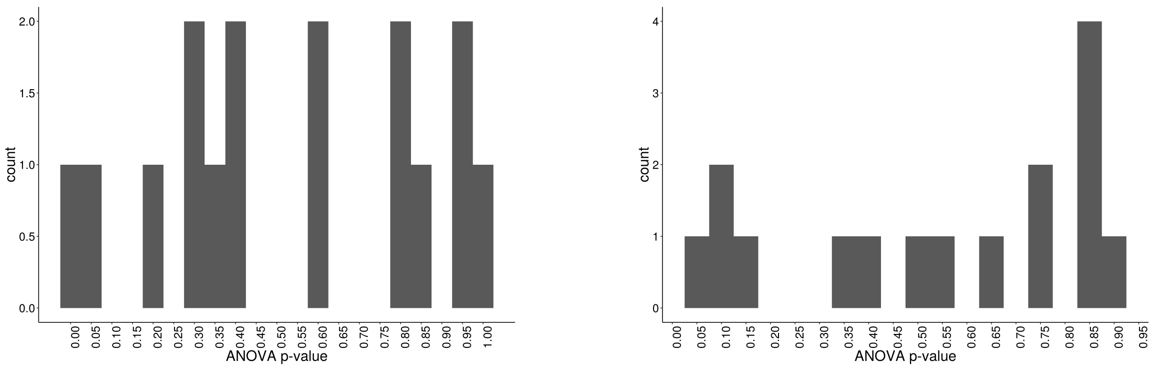
Table 12: OLS Estimates of Equation 9 by Credit Score Tercile

	Dependent variable:		
	$\mathbb{1}\{CheckNextMonth\}$		
	(1)	(2)	(3)
$\mathbb{1}\{\Delta CS < 0\}$	-0.014^{***}	-0.013^{***}	-0.009^{***}
	(0.001)	(0.001)	(0.001)
$\mathbb{1}\{\Delta CS < 0\} \times \|\Delta CS\|$	0.0002^{***}	-0.0004^{***}	-0.001^{***}
	(0.00004)	(0.00004)	(0.00004)
$AccountAge$	0.013^{***}	0.009^{***}	0.001^{***}
	(0.0002)	(0.0002)	(0.0002)
CS	0.001^{***}	0.0004^{***}	-0.00001
	(0.00002)	(0.00002)	(0.00002)
$NCheck$	-0.023^{***}	-0.017^{***}	-0.008^{***}
	(0.0003)	(0.0003)	(0.0002)
|CSChange|	0.0002^{***}	0.001^{***}	0.001^{***}
	(0.00003)	(0.00002)	(0.00002)
CS Tercile	1	2	3
Person FE	Y	Y	Y
Day of Week FE	Y	Y	Y
Month FE	Y	Y	Y
Observations	1,298,633	1,469,684	1,571,093
R^2	0.423	0.404	0.364
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This table shows the OLS estimates of Equation 9 for users with credit scores in each tercile. The tercile cutoffs are 571 (33^{rd} percentile) and 654 (66^{th} percentile). The sample in the first column is comprised of users with credit scores in the bottom tercile (less than 571), the sample in the second column is users with credit scores in the middle tercile (between 571 and 654), and the third column is users with credit scores in the top tercile (above 654). Standard errors are clustered at the individual level.

is that there is no significant difference in means across email variants within campaigns. Each observation is at the campaign level. Two campaigns had significantly different rates of $\mathbb{1}\{\Delta CS < 0\}$ (p-values of 0.024 0.074) and two had significantly different credit scores prior to receiving the email (p-values of 0.066 and 0.083). However with 16 campaigns we would expect a few to have differences by chance. This suggests that within campaigns, the variants are randomized across relevant observables.

Figure 9: Histogram of the P-Values Checking Covariate Balance for Randomization



Notes: This is a 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.

To measure whether these campaigns affect the main IV estimates, we removed the campaigns that had p-values of less than 0.1. The results were very similar as seen in Tables 13 and 14. At the monthly level, for negative-pre-trend users, the change is no longer statistically significant due to a drop in sample size, however the direction and magnitude remain similar.

Table 13: IV Removing Campaigns that Failed the Randomization Check

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-18.244**	8.386***
	(8.783)	(2.563)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	107,588	214,568

Notes: The table presents the IV regression of Equation 5 with only the check regressor listed. Campaigns 8, 9, 15, and 17 are removed as those had significant p-values during the randomization check. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email was sent to after with at least a two-month gap. Targeting controls are omitted for readability.*

Table 14: IV Removing Campaigns that Failed the Randomization Check

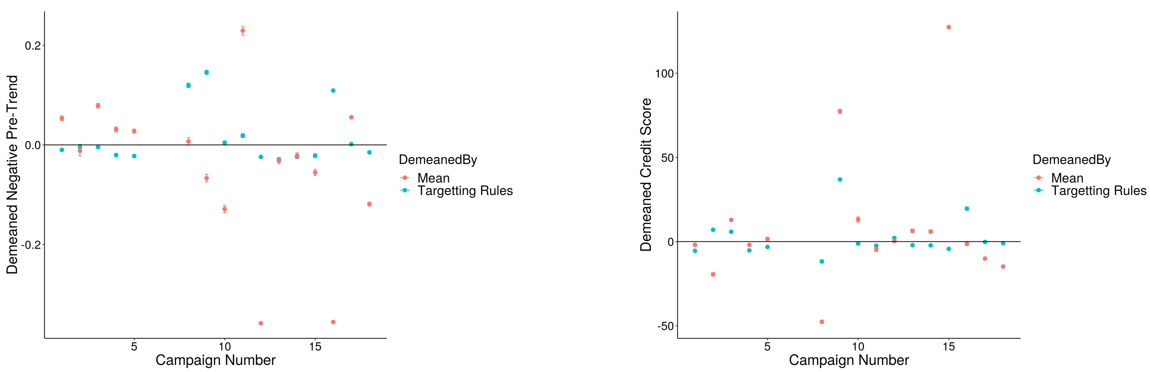
	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-3.936	1.315**
	(2.659)	(0.635)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	107,588	214,568

Notes: The table presents the IV regression of Equation 5 with only the check regressor listed. Campaigns 8, 9, 15, and 17 are removed as those had significant p-values during the randomization check. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email was sent to after with at least a two-month gap.*

Second, we show that after controlling for the campaign targeting rules, there is no correlation between the campaign the user received and her credit score or trend in credit score. In other words, we show that the targeting rules explain the vast majority of 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.[36]

Figure 10 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 and credit score is eliminated after controlling for the campaign targeting rules.

Figure 10: Variation by Group Across Campaigns



Notes: This figure shows the variation in whether users have $\Delta CS^{pre} < 0$ (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, in which 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.

To compare this more quantitatively, we measure how much the incremental variation campaign fixed effects explain. We compare the adjusted R^2 between the following regressions.

$$y_{ie} = \alpha + \beta X_{ie} + \epsilon_{ie}$$

$$y_{ie} = \alpha_c + \beta X_{ie} + \eta_{ie}$$

[36]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 36, excluding ΔCS^{pre}. The demeaned variables are then the residuals of this regression, $y_i - \hat{y}_i$.

where α_c are campaign fixed effects and the outcomes y_i are $\Delta CS^{pre} < 0$ and credit score. We find that adding campaign fixed effects explains only 0.52% more variation in credit scores and 0.55% more variation in whether the user has a declining credit score.[37] Since these numbers are so low, we conclude that the campaign assignments are not correlated with pre-trends and absolute credit scores after controlling for the targeting rules.

B.1 Instrument Strength

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 declining and non-declining credit scores prior to checking). The unrestricted and restricted regressions are as follows:

$$\mathbb{1}\{Check_{ie}\} = \alpha_e + \gamma X_{ie} + \eta_{ie} \tag{10}$$

$$\mathbb{1}\{Check_{ie}\} = \gamma X_{ie} + \eta_{ie}. \tag{11}$$

The regression estimates are reported in Table 15. The incremental F-statistic for the comparison between the restricted and unrestricted regressions above is 25 for users with declining credit scores and 187 for users with non-declining credit scores, indicating that the emails are not weak instruments for whether the user checks her credit information.[38] An alternative measure is the change in the R^2. The inclusion of email fixed effects increases the R^2 by 3% (0.222 to 0.228) for users with declining scores, and by 27% (0.182 compared to 0.231) for users with non-declining scores.

[37]For the regression on $\Delta CS^{pre} < 0$, adding campaign fixed effects increases the adjusted R^2 from 0.658 to 0.662, and 0.676 to 0.679 for the regression on credit score.

[38]The F-statistics were computed using a Wald test which compared the unrestricted versus restricted regressions (Equations 10 versus Equation 11) for both groups.

Table 15: OLS Results for the Restricted (First Column) and Unrestricted (Second Column) First-Stage Regression

	Dependent variable:			
	$1\{Check\}$			
	(1)	(2)	(3)	(4)
Homeowner	0.016***	0.015***	0.011***	0.008***
	(0.002)	(0.002)	(0.002)	(0.001)
ccclicks	0.042***	0.029***	0.028***	0.023***
	(0.002)	(0.002)	(0.001)	(0.001)
active	0.012***	0.009***	0.046***	0.029***
	(0.003)	(0.003)	(0.003)	(0.003)
refreshSameMonth	0.173***	0.104***	0.288***	0.164***
	(0.002)	(0.005)	(0.001)	(0.003)
CreditUtilizationGradeAB	−0.011	−0.009	−0.023***	−0.002
	(0.011)	(0.011)	(0.006)	(0.005)
CreditUtilizationGradeDF	0.014	0.017**	−0.064***	−0.006
	(0.008)	(0.008)	(0.005)	(0.005)
CSgreater700	0.033***	0.027***	0.002	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
ChangeinDebt	0.067***	0.074***	0.050***	0.074***
	(0.003)	(0.004)	(0.002)	(0.002)
Collections	−0.022***	−0.023***	−0.040***	−0.028***
	(0.002)	(0.002)	(0.002)	(0.001)
PaymentHistoryGradeCDF	0.032***	0.029***	0.044***	0.038***
	(0.006)	(0.006)	(0.004)	(0.004)
CreditAgeGradeAB	−0.559***	−0.559***	−0.353***	−0.566***
	(0.013)	(0.013)	(0.007)	(0.007)
CreditAgeGradeCDF	−0.552***	−0.553***	−0.368***	−0.552***
	(0.014)	(0.014)	(0.007)	(0.007)
CreditUtilizationNotA	−0.006	−0.007	0.050***	0.005
	(0.010)	(0.010)	(0.005)	(0.005)
ChangeinCreditUtilizationGrade	0.807***	0.264**	0.273***	0.062***
	(0.126)	(0.134)	(0.005)	(0.005)
DebtIncrease	−0.006***	−0.004**	−0.025***	−0.018***
	(0.002)	(0.002)	(0.002)	(0.002)
CSIncrease			0.062***	−0.020***
			(0.002)	(0.002)
CSIncrease2016	0.103***	0.104***	−0.050***	0.042***
	(0.002)	(0.002)	(0.002)	(0.002)
Email FE	N	Y	N	Y
ΔCS^{pre}	< 0	< 0	≥ 0	≥ 0
Observations	254,223	254,223	451,594	451,594
R^2	0.222	0.228	0.182	0.231
Adjusted R^2	0.222	0.227	0.182	0.231
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: This table tests the strength of the instruments by estimating the restricted and unrestricted (adding in email fixed effects) regressions. The outcome variable is whether a user checks her score within 31 days of receiving the email.

C Addressing Selection

C.1 Differential Attrition

While we cannot extend the results to users who do not return after receiving an email or users who opt out of all emails, we can check if users who receive positive or negative information have different retention rates, regardless of whether they check their information. Although users with declining and non-declining 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 23.2% drop out rate and the negative information group having a 21.5% drop out rate.[39]

C.2 Retention

Another source of selection bias comes from unobserved credit scores of those who do not return, in combination with the potential for email to impact retention. To illustrate, consider the case that compliers have persistent downward trends in credit scores, regardless of whether they check their score. If more effective emails are sent, then more compliers will remain on the platform. More compliers will then lead to a higher retention of users with declining scores, and therefore lend to the interpretation that more effective emails lead to a decrease in credit score. Through simulations, we find that this form of bias exists when the instruments impact retention. To measure whether this is the case in this setting, we estimate the following regression:

$$\mathbb{1}\{\text{Return}\}_{ie} = \beta ClickRate_e + \gamma X_{ie} + \epsilon_{ie}. \tag{12}$$

The dependent variable $\mathbb{1}\{\text{Return}\}_{ie}$ is an indicator variable for whether i has an observed post-trend credit score change (returns to the platform). $ClickRate_e$ is email e's click rate, or effectiveness. The click rates range from 0 to 1, with 1 meaning that all users who receive e click on the email. X_{ie} are the campaign controls. Table 16 displays the OLS estimates of Equation 12. We find that for users on a declining trend, a 100% increase in the email's click rate increases the probability that the user has an observed post-trend by 0.6 percentage points. This effect is significant at the 10% level. For users with a non-declining trend, the

[39]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 (0.016, 0.19) 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.

effect is close to zero and not statistically significant. The measured effect sizes are similar when the email's open rate, instead of click rate, is used as a measure of effectiveness. Therefore, in this setting, the instruments do not impact whether the outcome is observed.

A potential reason that we do not see any effect on retention is that because the metric is whether the user returns in two or more months, and that variation in the emails' effectiveness are not enough to induce retention over this period. Receiving an email compared to not receiving an email may be a stronger nudge to increase retention, but the variation in this regression comes from receiving a more effective email compared to receiving a less effective email, which may not enough to create long-term retention. Additionally, due to the randomization of the email copy in this month, users may receive another email in two months with an uncorrelated effectiveness to the focal email, which would have a stronger impact on whether the user checks in that month.

This, with the effects from imputing the missing outcomes in Appendix Section C.3, create more confidence that the effects of information are not entirely driven by selection bias resulting from unobserved outcomes.

Table 16: Impact of Email Effectiveness on Retention

	Dependent variable:	
	1{Return}	
	(1)	(2)
Click Rate	0.006*	−0.002
	(0.003)	(0.003)
DV Mean	0.88	0.86
ΔCS^{pre}	< 0	≥ 0
Campaign Targeting Controls	Y	Y
Observations	254,223	451,594
R^2	0.509	0.418
Note:	*p<0.1; **p<0.05; ***p<0.01	

Notes: OLS estimates of Equation 12. Sample consists of all users who viewed their report within three months of receiving a campaign email. Each observation is at the user-email level. Standard errors clustered at the email level.

C.3 Imputation

Due to the way the data provider pulls credit reports only when a user logs on, we have missing outcomes. For users in the IV sample (who have a non-missing pre-trend and logged onto the platform within the last three months before they received the focal email), 13.5% having missing outcomes. In general, it is not advisable to impute missing outcomes since they are not likely to be missing at random. Users who do not return to the platform after receiving one of the focal emails might be inherently different than users who remain, or at least a sample of them are. However, to measure how sensitive the results are and to provide some bounds, we impute these missing credit scores using two different imputation strategies.

First, we impute data as if the credit scores were missing at random. Second, since the data is not likely to be not missing at random, we impute under the assumption that users with lower credit scores are less likely to return. We match only on users whose post-trend is below the median, in which the median is a one-point increase in credit score. Given that we find users are less likely to return if they are doing badly, this seems like a reasonable counterfactual. The results are not sensitive to the missing data.

The imputation algorithm matches two users who are similar on the following list of covariates. One user has a missing outcome while the other does not. The missing outcome will then be imputed by the user who is similar on the covariates but has an observed outcome. We use the following covariates (as defined before the email was sent) to match users:

- Most recent credit score
- Number of times the user has checked her score
- First credit score pulled by the website
- Year and month
- Number of months the user has been on the website
- Pre_trend
- Indicator for whether the user has a negative_pre_trend
- Number of open accounts
- Total balances
- Total monthly payments
- Number of public records
- Number of delinquent accounts
- Number of derogatory accounts
- Number of inquiries that affect credit score
- Length of time since user last logged on
- The campaign that the user is assigned to
- Number of emails the user has received from the website since she signed up
- Click rate of the emails for that user
- Average rate at which the user clicks through on emails that she has opened
- Homeowner indicator*

- Whether the credit score increased in 2016*
- Whether the score increased in the past year*
- Whether the credit score changed since last report*
- Collections*
- Indicator for a credit score of greater than 700*
- Indicator for active on the website*
- Indicator for whether the debt has changed*
- Indicator for whether the debt has increased*

The * denotes the targeting rules used for email campaigns.[40]

Both imputation methods are done using predictive mean matching with the MICE (Multivariate Imputation via Chained Equations) package in R. For both strategies, five sets of imputed data are created. We estimate the IV regression five times, once on each set of imputed data. The coefficient shown is the average of the five coefficients. The standard errors are pooled across the five regressions accounting for imputation variance.

Table 17 displays the summary statistics for the observed post-trends, and the imputed values using each imputation method. The imputed values are similar to the overall distribution of original values, and the values imputed with below median values are lower than the observed, as expected. Table 18 displays the 2SLS estimates of Equation 5 and Equation 6 for this sample in which the missing values are imputed at random. Table 19 displays the estimates for those assuming the missing values are below the median. The estimated coefficients are robust to imputation strategies and are similar to the estimates on the non-imputed data in Table 8.

[40]We do not use all the targeting rules in these covariates because the imputation method is not able to handle a large number of covariates. Therefore, we also select the campaign the user was assigned to as a covariate to control for all the targeting rules.

Table 17: Summary Statistics for Observed and Imputed ΔCS_{ie}^{post}

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Original ΔCS_{ie}^{post}	-405.00	-20.00	1.00	-2.63	17.00	349.00
Imputed ΔCS_{ie}^{post}	-405.00	-19.00	1.00	-2.302	18.00	349.00
Imputed with Below Median Values	-405.00	-47.00	-21.00	-32.86	-6.00	0.00
Original and Imputed Combined	-405.00	-19.00	1.00	-2.309	18.00	409.00
Imputing All	-405.00	-19.00	1.00	-2.38	18.00	349.00

Notes: The sample used for this table is the same as that used in the main IV specification. The first row are the observed, non-missing ΔCS_{ie}^{post} summary statistics for users who have checked their credit score within three months of receiving the email, as is the sample in the main IV specification. The second row are the summary statistics just for the imputed values. The third row is for imputed values based on users who have below median outcomes. The fourth row are the summary statistics combining rows one and two for one of the five imputed data sets. The "imputing all" row imputes the outcome variable for all users based on the non-missing data. In all, the imputations are repeated five times and the summary statistics reflect an average of these five imputations.

Table 18: Two-Stage Least Squares IV Imputed Outcomes

	ΔCS^{post}		$\overline{\Delta CS^{post}}$ monthly	
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-25.40***	10.25***	-5.28**	2.64***
	(7.557)	(2.045)	(2.284)	(0.517)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Observations	254,196	451,54	254,196	451,545

Notes: This table presents the 2SLS estimates of Equation 5 and Equation 6. Targeting rules omitted for readability. Standard errors are clustered at the email level. Missing outcome values are imputed by matching on the targeting covariates. In this analysis we assume that the data is missing at random.

Table 19: Two-Stage Least Squares IV Imputed Outcomes using Users with Below Median Outcomes

	ΔCS^{post}		$\overline{\Delta CS^{post}} monthly$	
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-22.84***	28.78***	-4.90**	2.18***
	(7.479)	(3.587)	(2.301)	(0.479)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Observations	247,003	438,587	247,003	438,587

Notes: This table presents the 2SLS estimates of Equation 5 and Equation 6. Targeting rules omitted for readability. Standard errors are clustered at the email level. Missing outcome values are imputed with matching on email targeting covariates and only on users with below median outcomes.

In order to check the accuracy of the imputation, we also impute credit scores for all users, regardless of whether their actual credit score is observed. To do this, we replicate all users who do not have missing outcomes and remove their outcomes for both the total change in credit score, and the average monthly change in credit score. Imputation is done assuming outcomes are missing at random. The results, shown in Table 20, show that the estimates are similar when using this imputation strategy as well.

Table 20: Two-Stage Least Squares IV Imputed Outcomes For All Observations

	ΔCS^{post}		$\overline{\Delta CS^{post}} monthly$	
	(1)	(2)	(3)	(5)
$\mathbb{1}\{Check\}$	-27.92***	9.79***	-5.49***	2.48***
	(8.577)	(2.473)	(2.302)	(0.110)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Observations	254,196	451,545	254,196	451,545

Notes: This table presents the 2SLS estimates of Equation 5 and Equation 6. Targeting rules omitted for readability. Standard errors are clustered at the email level. Outcome values are imputed with matching on email targeting covariates even on the observations that are not missing.

D Compliers

In order to understand what population the LATE results hold for, we estimate the percentage of compliers in the sample. To calculate this, we follow steps first modeled by Imbens and Angrist (1994) and adapted by Maestas et al. (2013). This method has been used before with non-binary instruments. According to Bernstein et al. (2019),

> As pointed out by Maestas et al. (2013), when the treatment variable is binary and the instrument varies between zero to one, the size of the population that is on the margin and sensitive to the instrument is equal to the first-stage coefficient.

Furthermore, from Maestas et al. (2013), "More precisely, in the case of a binary treatment... the size of the marginal population is the first-stage coefficient times the range of [instruments]."

The treatment variable is whether or not a user checked her credit score, which is binary. To make the instrument vary between zero and one, rather than using the email copy, we use the average rate of checking one's credit score upon receiving a given email e. The average rate of checking for each email will be the new instrument, Z. Again, we split the sample into those with declining and those with non-declining pre-trends. Z is also calculated separately for each sub-sample. In other words, we calculate the mean of $\mathbb{1}\{Check\}$ for each email copy within users with declining pre-trends and again for those with non-declining pre-trends.

We then estimate the following OLS regression:

$$\mathbb{1}\{Check_{ie}\} = \beta_0 + \beta_1 Z_i + \nu_i. \tag{13}$$

The estimates of Equation 13 are in Table 21. Next, we take the coefficient from this regression and multiply it by the range of values of Z within each subsample. Within this range, there are some emails where the user would have checked their score and some where they would not have. In other words, there is some email that would have changed their treatment outcome.

Table 21: Estimates of Equation 13: First Stage Using Average Rate of Checking

	Dependent variable:	
	$1\{Check\}$	
	(1)	(2)
Average Check Rate By Email	0.545***	0.710***
	(0.016)	(0.011)
ΔCS^{pre}	< 0	≥ 0
Campaign Targeting Controls	Y	Y
Observations	223,913	386,570
R^2	0.126	0.110
Note:	*p<0.1; **p<0.05; ***p<0.01	

Notes: This table is based on the sample of users that we use in the IV estimation. Targeting rules are omitted for better readability.

Therefore, the percentage of compliers for each declining and non-declining sample is the following:

$$\mathbb{1}\{\Delta CS_{ie}^{pre} < 0\} : (max(Z_{\mathbb{1}\{\Delta CS_{ie}^{pre} < 0\}}) - min(Z_{\mathbb{1}\{\Delta CS_{ie}^{pre} < 0\}})) * 0.545 = 28.7\%$$

$$\mathbb{1}\{\Delta CS_{ie}^{pre} >= 0\} : (max(Z_{\mathbb{1}\{\Delta CS_{ie}^{pre} \geq 0\}}) - min(Z_{\mathbb{1}\{\Delta CS_{ie}^{pre} \geq 0\}})) * 0.710 = 30.0\%.$$

Other work has been done to characterize compliers in more detail, but, to our knowledge (as well as according to Marbach and Hangartner (2020)), these methods have not been extended to include a non-binary instrument. Still, similar to Bernstein et al. (2019), we can conduct some characterization of the compliers by splitting the sample into different groups based on other covariates of interest, reconstructing Z, and re-estimating the first stage. We further split the samples into users who had pre-trend credit score changes of ten points or fewer or more than ten points in magnitude (in each pre-trend subsample). The percentage of compliers in each group is Table 22.

Table 22: Percentage of Compliers by Pre-Trend Direction and Magnitude

	$\Delta CS^{pre} < 0$	$\Delta CS^{pre} \geq 0$
Change > |10|	24.7%	26.4%
Change <= |10	30.2%	32.1%

In both the compliance rates by pre-trend, and by the magnitude of the change in credit score within pre-trend samples, we find that there are slightly more compliers in the positive/flat pre-trend group, which makes sense given the initial evidence of users who are on increasing trends are more likely to return. It is interesting to note that the percentage of compliers does not vary drastically across groups.

E Comparing Across Different Samples of Users

E.1 Users Who Opt Out of Emails

We selected a random sample of 100,000 users who opt out of emails entirely. After removing users with data errors, we have 83,857 users.[41] Users who opt out of emails have slightly higher credit scores than the random sample of users, with an average score of 626 when they first log on and a standard deviation of 96.6 (compared to 599 and 91 for the random sample). The average score across all observed credit reports is 655 and the standard deviation is 93.1. 75% of these users (62,934) return to the website at least once past their initial sign up. Table 23 displays summary statistics for these users. Changes in credit score and the rate at which users check their score are similar to the random sample.

Table 23: For Users Who Opt Out of Emails - Summary Statistics for Changes in Credit Score and How Often Users Check Updated Credit Reports

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Change in CS	-376.00	-6.00	1.00	1.07	10.00	309.00
Months Between Checking CS	0.00	0.87	1.07	1.68	1.57	47.20

Notes: The table is based on a random sample of 83,857 users who opt out of email communication. An observation is a user and credit report check. "Change in CS" is the difference in credit score between the consecutive reports, and "Months Between" is the number of months between when the user checks her updated reports.

[41] We removed users with credit scores < 300 or >850.

E.2 IV vs. Random Sample

Next we compare users in the IV sample with those in our random sample. Table 24 shows that users in the IV sample check their credit score more often on average and have higher credit scores upon signing up for the platform compared to the general population.

We also looked at email engagement in Table 25. The first row is the number of emails that a user receives a month over the course of our data window. To be in the IV analysis, users have to at least receive the focal instrument emails. However, in the random sample, users can opt-out of email communication. Therefore, we see that users in the IV sample receive more 3 emails per month. Additionally, they have a 6% higher open rate for emails. Both groups have similar rates of clicking on the emails, conditional on opening.

We also measure whether users in the IV sample engage in information avoidance after receiving negative feedback. We estimate Equation 9 on the IV sample, using all observed credit scores from these users.[42]

Table 24: Summary Statistics for Users in IV and Random Sample

Variable	Group	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Change in CS	IV Sample	-148	-1.95	0.68	0.54	3.38	131.5
	Random Sample	-386	-6	1	1.01	10	409
Months Between Checking CS	IV Sample	0.39	1.17	1.53	2.07	2.33	18.78
	Random Sample	0	0.87	1.07	1.72	1.53	37.93
Number of CS Checks	IV Sample	3	10	17	18.22	25	81
	Random Sample	1	1	2	5.91	7	81
First CS	IV Sample	300	551	613	618.05	680	839
	Random Sample	300	524	581	594.21	653	839
Average CS	IV Sample	300	499	558	568.91	635	839
	Random Sample	300	512	560	575.91	635	839

Notes: This table consists of the sample of users who are in the IV analysis as well as the random sample of 969,254 users from Section 3. "Change in CS" is the difference in credit score between the consecutive reports, and "Months Between Checking CS" is the number of months between when the user checks her updated reports. "Number of CS Checks" is the number of times the user has ever checked her score. "First CS" is the user's credit score the first time she logs onto the platform. "Average CS" is the user's average credit score across the data set.

[42] We do not estimate the left-digit-bias regression, as this is a drastically smaller sample than the random sample, and are not able to get enough observations close to the 00 cutoff.

Table 25: Email Engagement Summary Statistics for Users in IV and Random Sample

Variable	Group	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Emails Per Month	IV Sample	0.15	9.22	21.15	30.64	39.28	431.65
	Random Sample	0	11.31	25.07	26.45	41.87	84
Percent Opened	IV Sample	0.18	12.07	24.66	31.42	45.48	100
	Random Sample	0.18	6.52	16.76	25.29	36.98	100
Percent Clicked	IV Sample	0	6.71	15.35	21.81	30.51	100
	Random Sample	0	3.11	12.5	21.06	30.56	100

Notes: This table consists of the sample of users who are in the IV analysis as well as the random sample of 969,254 users from Section 3. Emails Per Month is the average number of emails a user receives in our data window. Percent opened is the percentage of those emails received that the user opened. Percent clicked is the percent of those emails that the user clicked on, conditional on opening.

Table 26: OLS Estimates of Equation 9 for Users who Receive IV Emails

	Dependent variable:		
	$\mathbb{1}\{$Check Next Month$\}$		
	(1)	(2)	(3)
$\mathbb{1}\{\Delta CS < 0\}$	-0.136^{***}	-0.148^{***}	-0.065^{***}
	(0.002)	(0.002)	(0.002)
$\mathbb{1}\{\Delta CS < 0\} \times \lvert\Delta CS\rvert$			-0.005^{***}
			(0.0001)
Account Age	-0.003^{***}	-0.047^{***}	-0.048^{***}
	(0.0001)	(0.0003)	(0.0003)
CS	0.002^{***}	0.001^{***}	0.001^{***}
	(0.00002)	(0.00002)	(0.00002)
NCheck		0.064^{***}	0.066^{***}
		(0.0003)	(0.0003)
$\lvert\Delta CS\rvert$			0.002^{***}
			(0.0001)
Person RE	Y	Y	Y
Day of Week FE	Y	Y	Y
Month FE	Y	Y	Y
Observations	7,539,647	7,539,647	7,539,647
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This table replicates Table 11 for the final sample of users who are in our main IV specification.

E.3 Users with Missing Outcomes

Lastly, in Table 27, we compare users in the IV sample with users who would have been in the IV sample, except they did not return to the website two months after receiving the email. Users who have a missing outcome tend to check their reports less frequently, and a more positive change in credit scores prior to receiving the email.

Table 27: Summary Statistics of Users in IV and Users With Missing Outcome Variable

Variable	Sample	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Pre Trend Change	Missing Outcome	-318	-4	1	2.31	14	271
	IV Sample	-352	-6	1	1.44	10	315
# of Months Between	Missing Outcome	0	0.88	1.34	2.73	3.3	20.95
Pre Trend Checks	IV Sample	0	0.9	1.08	1.68	1.66	20.85

Notes: The "IV sample" are the 464,752 users who are in the main IV analysis (received an email from one of the selected campaigns, checked their report at least twice before receiving the email, and checked their credit report within three months of receiving the email, and have an observed credit score at least two months after receiving the email). The "missing outcome" sample are the 90,565 users who fulfill the first three criteria to be in the IV sample but do not have an observed credit score at least two months after receiving the email.

F Simulating Credit Scores and Retention

In this section, we simulate the evolution of user retention and the credit scores of these users over time, given the estimated effects of checking and avoiding information. The intuition of these effects is the following. The estimated effect of checking credit reports on credit score is negative for users with declining scores, but positive for those with increasing scores. Additionally, users with declining scores are less likely to check their report in the next month. These two effects lead to the following cycle: for users who start off with declining scores, if they receive an email, their score further decreases, causing them to be less likely to return to the platform. Upon receiving an email, of those that do return, their score further declines. On the other hand, those with increasing scores have a further increase their credit score after receiving the email, and are more likely to return to the platform. The goal of these simulations is to combine these effects and isolate their impact on user retention and credit scores.

To be comparable to the random sample of users' data in this paper, we generate a sample of one million users, and simulate their credit scores for two consecutive time periods. In the first period, their credit scores are drawn from empirical distribution of credit scores of the random sample, as shown in Figure 2. We then generate the credit score transition matrix ρ using the observed transitions in consecutive months of the random sample. The two credit scores allow us to observe the users' initial credit score trend. Thus, the entire sample consists of users who have checked their credit score at least twice. We then simulate their activity for the next six time periods (months).

We simulate four email targeting policies: no emails (baseline), emails sent to everyone, emails sent to only people with increasing credit scores, and emails sent only to people with decreasing credit scores. We follow a similar timeline as in Figure 6. In the simulations, t is analogous to t_2^e, $t-1$ to t_1^e, and $t+1$ to t_{post}^e. In all these simulations, except the baseline, emails are sent out between t and $t+1$. Note that t_{check}^e in Figure 6 is not a time period in the simulations because not everyone checks their score after receiving the email. Additionally, another difference is that in the simulations, t denotes a calendar month, where in t in Figure 6 is an instance when a user checks their report.

We generate the users' decisions to view their report and their subsequent credit score evolutions with the following steps. For each user,

1. Given the user's trend in credit score between $t-1$ and t, estimate the probability π the user will check her report at $t+1$ using regression estimates from Table 3.
2. Randomly draw whether the user checks her report at $t+1$, given π.
3. If she checks her report at $t+1$, generate the user's credit score at $t+1$, given her

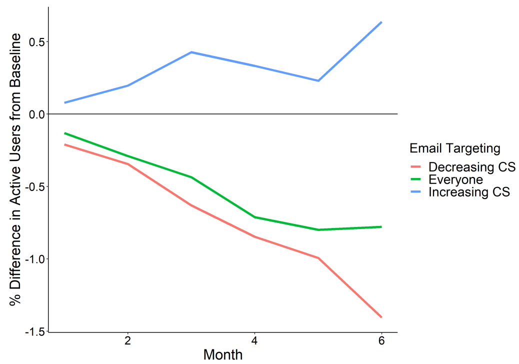
score at t, using the transition matrix ρ and whether she received an email between t and $t+1$. If she receives an email, then her score at $t+1$ is adjusted by the intent to treat effects estimated in Table 7.

4. Repeat steps 1-3 until the user chooses not to check her score in the next period.

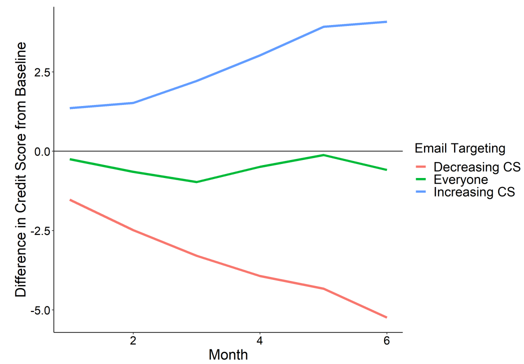
Note that this simulation makes the following simplifications. First, it includes only users who check *consecutively.* Once a user chooses to not check her score in the next month, she drops out of the sample for the remaining time periods. This simplification allows us to not model a user's belief about her credit score if she does not check in a given month. Second, we assume that all emails are "effective" emails, meaning they have a click rate of 100%. This allows us to simplify the data generating process by ignoring modeling emails of varying click rates, and assuming that each email received decreases future credit scores by 1.6, or increases future credit scores by 1.2, depending on the prior credit score trend. Note that receiving an effective email between t and $t+1$ does not guarantee the user checks her report at $t+1$. Rather, a 100% click rate means that the user checks her report at the time of the email.

Below, we report the results from two sets of simulations. In the first set of simulation results (which we refer to as "Information effects only"), emails do not have any impact on the probability that the user checks her report in the next time period. This allows us to isolate the effects from checking information and information avoidance only. The second set of results (referred to as "Information and email effects") assumes that emails increase the probability that users check their reports, which is closer to reality, but by itself is more difficult to parse the effects of information from receiving an email. Since we do not observe the causal effect of receiving an email on whether the user checks her report, as there is no exogenous variation in whether a user receives an email, we assume that receiving an email this month increases the probability that a user checks her report next month by 8%. This estimate is derived from Table 32, which shows that users are 8% more likely to click on an "Update" email rather than a "Non-Update" email.

Figure 11: Simulation Results: Information Effects Only



(a) Retention



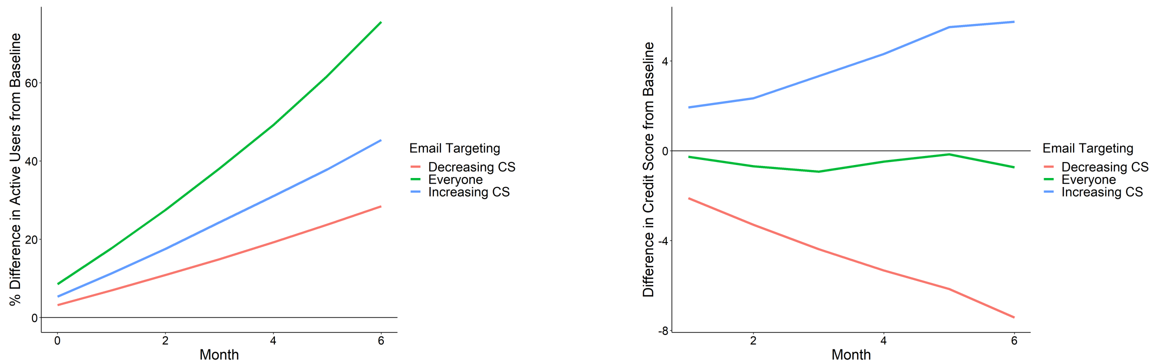
(b) Credit Scores

Notes: Results from 1 million simulated users. Receiving an email does not boost the probability the user checks her report. Panel (a) displays the retention metric, which is the percentage difference in the number of users who check their report in a given month compared to the baseline policy where no emails are sent. Panel (b) displays the difference in the average credit scores of users who check their scores in a given month, relative to the baseline policy.

Figure 11 displays the simulation results, relative to the baseline of the firm not sending any emails. Figure 11a displays the percentage difference in the number of users who are checking their credit report in a given month, relative to the the baseline. At month six, if the email was sent to only those with increasing credit scores each month, there would be 0.5% more active users on the platform compared to the baseline. This illustrates that the targeting rule of sending emails to only those with increasing scores leads to a 2% higher retention rate, or 20,000 more active users on the platform, compared to if the platform sent emails only to those with decreasing scores. Figure 11b displays the difference in the average credit scores across the different policies. Similarly, sending emails only to users with decreasing trends leads to an average of eight points lower than if the emails were sent only to users with increasing trends.[43]

[43]Comparing the distance between the red and blue lines.

Figure 12: Simulation Results: Information and Email Effects



(a) Retention (b) Credit Scores

Notes: Simulation results from 1 million simulated users. Receiving an email increases the probability the user checks her report by 8%. Panel (a) displays the retention metric, which is the percentage difference in the number of users who check their report in a given month compared to the baseline policy where no emails are sent. Panel (b) displays the difference in the average credit scores of users who check their scores in a given month, relative to the baseline policy.

Figure 12 displays the simulation results when emails boost the probability that users check their report in the next time period. Since receiving an email boosts retention, the retention rate increases for all targeting policies compared to the baseline, with the smallest increase in retention for the policy of sending emails only to users with decreasing trends. However, in Panel (b), the credit score disparity between the increasing-only policy and decreasing-only policy is larger: the average credit score for active users under the decreasing trend email policy is 12 points lower than that of the increasing trend email policy.[44] This disparity is caused by the email's effect on retention. If a user's credit score was decreasing (increasing), she is more (less) likely to stop checking her report. However, if she receives an email, she is more (less) likely to check her report, causing her score to drop (increase) further. Due to the differences in retention and credit scores across policies, these simulations suggest that the credit score trend is an important piece of information to take into consideration for email targeting.

[44]Comparing the distance between the red and blue lines.

G Additional Tables and Figures

G.1 Evidence of Information Avoidance

Table 28: Robustness Checks: OLS Estimates of Equation 9

	Dependent variable:			
	$\mathbb{1}\{CheckNextMonth_t\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\Delta CS < 0\}$	-0.029^{***}	-0.030^{***}	-0.030^{***}	-0.003^{***}
	(0.0005)	(0.0004)	(0.0004)	(0.001)
$\mathbb{1}\{\Delta CS < 0\} \times \lvert\Delta CS\rvert$				-0.001^{***}
				(0.00002)
$AccountAge$				-0.013^{***}
				(0.0001)
CS				0.0001^{***}
				(0.00000)
$NCheck$				0.021^{***}
				(0.0001)
|CSChange|				0.001^{***}
				(0.00001)
$\mathbb{1}\{CheckNextMonth_{t-1}\}$				0.214^{***}
				(0.001)
Person FE	N	Y	Y	N
Day of Week FE	N	N	Y	Y
Month FE	N	N	Y	Y
Observations	4,339,410	4,339,410	4,339,410	4,339,410
R^2	0.001	0.338	0.338	0.115
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: This table presents alternative specifications of Equation 9 by iteratively removing person fixed effects and controls for seasonality. The last column displays the estimates when the lagged dependent variable is used, instead of individual fixed effects. Standard errors clustered at the individual level.

Table 29: Robustness Checks: OLS Estimates of Equation 9 by Time Interval

	Dependent variable:			
	$\mathbb{1}\{CheckNextMonth\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{\Delta CS < 0\}$	-0.011^{***}	-0.013^{***}	-0.018^{***}	-0.013^{***}
	(0.001)	(0.001)	(0.004)	(0.003)
$\mathbb{1}\{\Delta CS < 0\} \times \lvert\Delta CS\rvert$	-0.00001	-0.0002^{***}	-0.001^{***}	-0.0005^{***}
	(0.00003)	(0.00004)	(0.0001)	(0.0001)
$AccountAge$	0.014^{***}	0.015^{***}	0.024^{***}	0.016^{***}
	(0.0002)	(0.0002)	(0.001)	(0.0003)
CS	0.0004^{***}	0.001^{***}	0.001^{***}	0.0004^{***}
	(0.00001)	(0.00001)	(0.0001)	(0.00004)
$NCheck$	-0.020^{***}	-0.023^{***}	-0.047^{***}	-0.053^{***}
	(0.0003)	(0.0003)	(0.001)	(0.001)
|CSChange|	0.0002^{***}	0.0003^{***}	0.001^{***}	0.0003^{***}
	(0.00002)	(0.00002)	(0.0001)	(0.0001)
t - $(t-1)$ (Months)	1	2	3	4+
Person FE	Y	Y	Y	Y
Day of Week FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
Observations	2,072,852	1,534,876	300,412	430,824
R^2	0.411	0.430	0.704	0.691
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: This table displays results of Equation 9 for different sample based on the time interval of the pre-trend. Observations in the first column are users in which t - $(t-1)$ is 1 month, second column is 2 months, etc. Standard errors clustered at the individual level.

Table 30: Robustness Checks: OLS Estimates of Equation 1

	Dependent variable:		
	$\mathbb{1}\{CheckNextMonth\}$		
	(1)	(2)	(3)
$\mathbb{1}\{LeftDigitChanged\}$	-0.002	-0.008^{***}	0.019^{***}
	(0.001)	(0.001)	(0.001)
$\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$	-0.020^{***}	-0.009^{***}	-0.014^{***}
	(0.002)	(0.002)	(0.002)
$\mathbb{1}\{\Delta CS < 0\}$	-0.025^{***}	-0.020^{***}	-0.001
	(0.001)	(0.001)	(0.001)
$\lvert\Delta CS_{it}\rvert$		0.0001^{***}	0.0003^{***}
		(0.00002)	(0.00003)
$AccountAge$		0.001^{***}	-0.004^{***}
		(0.0001)	(0.0001)
Person FE	N	N	Y
Day of Week FE	N	Y	Y
Month FE	N	Y	Y
Rounded CS FE	N	Y	Y
Observations	854,329	854,329	854,329
R^2	0.001	0.024	0.528
Adjusted R^2	0.001	0.024	0.289
Note:			*p<0.1; **p<0.05; ***p<0.01

Notes: This sample consists of users with credit scores within 10 points to the nearest hundred (scores between X90 and Y09). Both specifications include controls for the rounded credit score to the nearest hundred, month, and day of week. Second column has individual fixed effects.

Table 31: Robustness Checks: RD Results with Different Bandwidths

	Dependent variable:			
	$\mathbb{1}\{CheckNextMonth\}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{LeftDigitChanged\}$	0.001	-0.005^{***}	-0.008^{***}	-0.009^{***}
	(0.002)	(0.002)	(0.001)	(0.001)
$\mathbb{1}\{LeftDigitChanged\} \times \mathbb{1}\{\Delta CS < 0\}$	-0.011^{***}	-0.010^{***}	-0.009^{***}	-0.010^{***}
	(0.004)	(0.003)	(0.002)	(0.002)
$\mathbb{1}\{\Delta CS < 0\}$	-0.020^{***}	-0.021^{***}	-0.020^{***}	-0.018^{***}
	(0.003)	(0.002)	(0.001)	(0.001)
$\lvert\Delta CS_{it}\rvert$	0.0001^{***}	0.0001^{***}	0.0001^{***}	0.0001^{***}
	(0.00004)	(0.00003)	(0.00002)	(0.00002)
$AccountAge$	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.00004)
Bandwidth	3	5	10	15
Person FE	N	N	N	N
Day of Week FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
Rounded CS FE	Y	Y	Y	Y
Observations	256,903	425,520	854,329	1,284,648
R^2	0.024	0.023	0.024	0.023
Adjusted R^2	0.024	0.023	0.024	0.023
Note:				*p<0.1; **p<0.05; ***p<0.01

Notes: This table presents the OLS estimates of Equation 1 with different bandwidths. For example, a bandwidth of 3 in the first column means that only credit scores that are within 3 of the hundred threshold are included (497 to 502, 597 to 602, 697 to 702, 797 to 802).

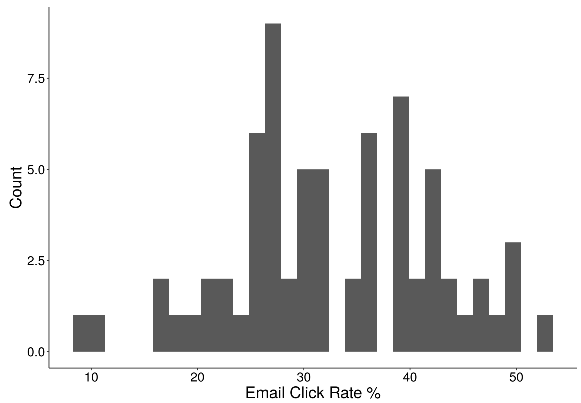
Table 32: OLS Results of Whether the Email is an "Update" Email on Open and Click Rates.

	Dependent variable:	
	Open Rate (%)	Click Rate (%)
	(1)	(2)
$\mathbb{1}\{\text{Update Email}\}$	4.905***	1.116***
	(0.245)	(0.151)
Constant	34.298***	13.120***
	(0.102)	(0.063)
Observations	26,240	26,240
R^2	0.015	0.002
Adjusted R^2	0.015	0.002
Note:		*p<0.1; **p<0.05; ***p<0.01

Notes: This table presents the OLS results of $y_e = \alpha + \beta \mathbb{1}\{\textit{Update Email}\}_e + \epsilon_e$, where y_e is the open rate (Column 1) and click rate conditional on opening (Column 2). Each observation is at the email (e) level. Sample consists of all emails sent by the firm to at least 1000 users before June 1, 2018. An "update" email is an email in which the subject line contains "update" or "changed".

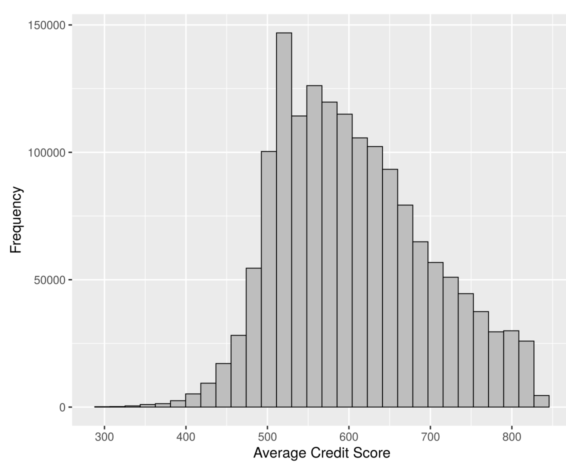
G.2 IV Validity

Figure 13: Histogram of Click Rates for the IV Email Copies



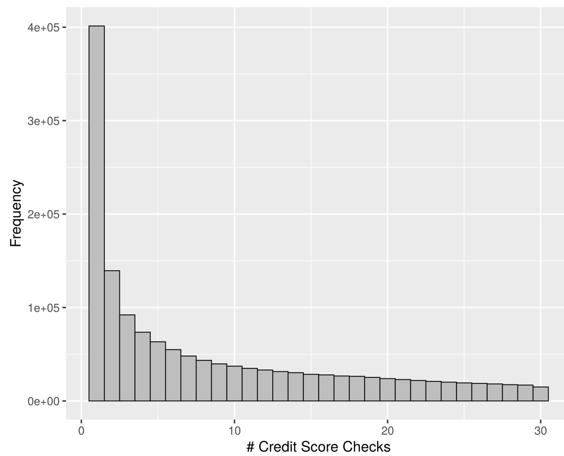
Notes: In this table, each observation is an email copy in the IV analysis. For each email, we calculated the percentage of users who receive the email who click on the email, conditional on opening, which is on the x-axis. The y-axis is how many of the emails are in that click rate bucket. The click rates range from 10-53%.

Figure 14: Average Credit Score for Users in the IV Analysis



Notes: This table is based on the sample of users who are in the IV analysis. Each observation is at the individual level. A user's average credit score is averaged over all her observed credit reports.

Figure 15: Histogram of the Number of Times a User in the IV Analysis Sample Checks her Score



Notes: This histogram consists of the sample of users who are in the IV analysis. Number of times a user checks her credit score, truncated at 30 checks.

Table 33: Summary Statistics of Time Before and After Receiving the Email

Pre-Trend	Variable	Min	Median	Mean	Max	SD
Declining	$t^e - t^e_2$ (Months)	0.0	0.7	0.9	3.0	0.7
	$t^2_{post} - t^e$ (Months)	2.1	2.9	4.2	27.3	3.3
Non-declining	$t^e - t^e_2$ (Months)	0.0	0.5	0.8	3.0	0.7
	$t^2_{post} - t^e$ (Months)	2.1	2.8	3.9	27.4	3.1

Notes: Each observation is at the user-email level. The first row is the number of months the user last checks her credit report before receiving her email ($t^e - t^e_2$) and the second row is the difference in time between her credit report for the post period and receiving the email ($t^e_{post} - t^e$).

Table 34: Subject Lines for Each of the Emails Sent Out over All 16 Campaigns

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]##, we've 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.

Notes: Each campaign has at least 2 email variants. "##[F][S][P][first]##" is the user's first name.

G.3 Additional IV Robustness Checks

Table 35: Heterogeneity in Intent-to-Treat Effects by Credit Score

	Dependent variable:					
	ΔCS^{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
Click Rate	-3.665^{***}	1.125	-0.681	1.165	1.895^{**}	1.097
	(1.234)	(1.320)	(0.692)	(1.120)	(0.753)	(0.838)
ΔCS^{pre}	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
CS Tercile	Bottom	Middle	Upper	Bottom	Middle	Upper
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
Observations	75,187	75,126	73,600	109,938	128,951	147,681
R^2	0.013	0.029	0.014	0.014	0.033	0.019
Note:						*p<0.1; **p<0.05; ***p<0.01

Notes: This table replicates Table 7, but instead further separates samples by the user's credit score tercile before receiving the email. The tercile cutoffs are 578 and 661.

Table 36: OLS Results for the First Stage in the IV Estimation

	Dependent variable:	
	$1\{Check\}$	
	(1)	(2)
Homeowner	0.016***	0.011***
	(0.002)	(0.002)
ccclicks	0.042***	0.028***
	(0.002)	(0.001)
active	0.012***	0.046***
	(0.003)	(0.003)
refreshSameMonth	0.173***	0.288***
	(0.002)	(0.001)
CreditUtilizationGradeAB	−0.011	−0.023***
	(0.011)	(0.006)
CreditUtilizationGradeDF	0.014	−0.064***
	(0.008)	(0.005)
CSgreater700	0.033***	0.002
	(0.002)	(0.002)
ChangeinDebt	0.067***	0.050***
	(0.003)	(0.002)
Collections	−0.022***	−0.040***
	(0.002)	(0.002)
PaymentHistoryGradeCDF	0.032***	0.044***
	(0.006)	(0.004)
CreditAgeGradeAB	−0.559***	−0.353***
	(0.013)	(0.007)
CreditAgeGradeCDF	−0.552***	−0.368***
	(0.014)	(0.007)
CreditUtilizationNotA	−0.006	0.050***
	(0.010)	(0.005)
ChangeinCreditUtilizationGrade	0.807***	0.273***
	(0.126)	(0.005)
DebtIncrease	−0.006***	−0.025***
	(0.002)	(0.002)
CSIncrease		0.062***
		(0.002)
CSIncrease2016	0.103***	−0.050***
	(0.002)	(0.002)
Constant	0.430***	0.421***
	(0.004)	(0.004)
ΔCS^{pre}	< 0	≥ 0
Observations	254,223	451,594
R^2	0.222	0.182
Note:	*p<0.1; **p<0.05; ***p<0.01	

Notes: This table is the regression of Equation 4. The outcome variable is whether or not a user checks her score within 31 days of receiving the email. The independent variables in this regression are the targeting rules, which are all indicator variables. We describe the targeting rules with unclear names: ccclicks (whether the user has clicked on a credit card application in the same month), active (user has created an account, or opened or clicked on an email in the last three months), refreshSameMonth (viewed their credit report more than once in that month), CSIncrease (an indicator for whether the user's last observed credit score has increased since their previous score), CSIncrease2016 (whether the user's credit score has increased in 2016).

Table 37: IV with All Regressors Displayed

	ΔCS^{post} (1)	ΔCS^{post} (2)
$\mathbb{1}\{Check\}$	-23.247***	9.084***
	(6.928)	(2.094)
active	-1.393***	-0.439
	(0.354)	(0.352)
Homeowner	2.540***	1.949***
	(0.199)	(0.184)
ccclicks	-3.043***	-3.848***
	(0.539)	(0.255)
refreshSameMonth	6.714***	1.092
	(1.655)	(0.688)
CreditUtilizationGradeAB	0.762	-2.492**
	(2.963)	(1.068)
CreditUtilizationGradeDF	12.702***	4.621***
	(3.852)	(1.476)
CSgreater700	-3.892***	-5.429***
	(0.556)	(0.480)
ChangeinDebt	-0.746	-4.026***
	(0.960)	(0.484)
Collections	3.406***	2.908***
	(0.544)	(0.465)
PaymentHistoryGradeCDF	14.106***	0.333
	(2.668)	(1.003)
CreditAgeGradeAB	-19.203***	4.694***
	(5.564)	(1.670)
CreditAgeGradeCDF	-26.049***	-0.636
	(5.466)	(2.149)
CreditUtilizationNotA	-7.125***	0.779
	(2.350)	(1.590)
ChangeinCreditUtilizationGrade	7.031	-3.060***
	(9.629)	(1.025)
DebtIncrease	-5.759***	-5.852***
	(0.232)	(0.201)
CSIncrease	0.000	-2.834***
	(.)	(0.271)
CSIncrease2016	3.993***	-2.230***
	(0.987)	(0.327)
Constant	14.152***	-2.292**
	(2.638)	(1.115)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	223,913	386,570

Notes: This table presents the IV regression of Equation 5 with all regressors. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent.

Table 38: Number of Emails Users Receive and Open Each Month

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Emails Per Month	0.15	9.22	21.15	30.63	39.26	431.65
Percent Opened	0.18	6.84	16.73	24.53	35.56	100.00
Percent Clicked	0.00	3.79	11.85	19.53	27.38	100.00

Notes: The sample is users who are in the IV analysis, i.e., any user who received at least one of the 16 campaigns. The statistics presented are the same as in Table 2 to compare these to a random sample of users. Each observation is the user-month level.

Table 39: IV with Different Samples of when a User Last Logged On Before the Email

	ΔCS^{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Check\}$	-14.923**	4.911**	-25.536***	9.814***	-24.784***	12.607***
	(6.403)	(2.345)	(6.879)	(2.365)	(6.777)	(2.620)
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0	< 0	≥ 0
Prev Login Window	1 Month	1 Month	6 Months	6 Months	No Cutoff	No Cutoff
Observations	143,836	267,934	259,867	441,246	283,461	474,492

Notes: This table presents the IV regression of Equation 5 with only the check regressor listed. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for the email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent. "Prev Login Window" refers to the last time a user checked her score before the email was sent. For one month, this means that a user last checked her score within a month before the email was sent. The main specification allows a user to have checked her score within the last three months. The last columns do not restrict the sample.*

Table 40: IV with Different Samples of when a User Last Logged On Before the Email

	ΔCS^{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Check\}$	-3.051	0.785	-4.282**	1.727***	-3.616*	2.348***
	(2.009)	(0.608)	(2.029)	(0.505)	(2.005)	(0.505)
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0	< 0	≥ 0
Prev Login Window	1 Month	1 Month	6 Months	6 Months	No Cutoff	No Cutoff
Observations	143,836	267,934	259,867	441,246	283,461	474,492

Notes: This table presents the IV regression of Equation 6 with only the check regressor listed. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for the email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. "Prev Log on Window" refers to the last time a user checked her score before the email was sent. For one month, this means that a user last checked her score within a month before the email was sent. The main specification allows a user to have checked her score within the last three months. The last columns do not restrict the sample.

Table 41: Two-Stage Least Squares IV With Sample of Users Who Received One Focal Email

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-5.373	6.007***
	(5.909)	(1.847)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	77,736	138,904

Notes: This table presents the IV regression of Equation 5 with only the check regressor listed. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent. This analysis restricts the sample to users who received one of the focal emails which is 84% of the users in the sample.

Table 42: Two-Stage Least Squares IV With Sample of Users Who Received One Focal Email Within a Month

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-19.478***	9.490***
	(6.090)	(2.065)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	183,938	318,288

Notes: This table presents the IV regression of Equation 5 with only the check regressor listed. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent. This analysis restricts the sample to users who received at most one of the focal emails within a month which is 97.3% of the users in the sample.*

Table 43: Two-Stage Least Squares IV With Sample of Users Who Received One Focal Email - Monthly Change Outcome

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	0.213	0.421
	(1.528)	(0.469)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	77,736	138,904

Notes: This table presents the IV regression of Equation 6 with only the check regressor listed. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. This analysis restricts the sample to users who received one of the focal emails which is 84% of the users in the sample.*

Table 44: Two-Stage Least Squares IV With Sample of Users Who Received One Focal Email Within a Month - Monthly Change Outcome

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-3.649**	1.615***
	(1.808)	(0.446)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	183,938	318,288

Notes: This table presents the IV regression of Equation 6 with only the check regressor listed. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. This analysis restricts the sample to users who received at most one of the focal emails within a month which is 97.3% of the users in the sample.

Table 45: Two-Stage Least Squares IV By Tercile - Full Regression

	ΔCS^{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Check\}$	-21.521***	-0.504	-11.610*	-4.223**	19.629***	21.511***
	(6.618)	(7.584)	(6.232)	(2.118)	(2.839)	(2.537)
active	-4.387***	-3.712***	-0.448	-1.896**	-2.014***	-1.575***
	(0.978)	(0.734)	(0.487)	(0.846)	(0.731)	(0.346)
Homeowner	3.231***	5.033***	3.545***	2.174***	3.812***	3.292***
	(0.629)	(0.458)	(0.297)	(0.529)	(0.452)	(0.222)
ccclicks	-0.967	-3.758***	-4.725***	-1.957***	-3.840***	-5.043***
	(0.724)	(0.768)	(0.608)	(0.314)	(0.362)	(0.504)
refreshSameMonth	3.980***	7.753***	5.452***	3.082***	2.450**	-0.170
	(1.360)	(1.823)	(1.405)	(0.661)	(0.990)	(0.710)
CreditUtilizationGradeAB	1.563	2.420	3.298	0.257	-1.389	-1.797
	(5.500)	(6.181)	(4.652)	(3.720)	(1.312)	(1.298)
CreditUtilizationGradeDF	-9.204	-2.396	-7.477*	0.011	-1.533**	-2.753***
	(5.642)	(4.479)	(4.366)	(2.972)	(0.755)	(1.036)
ChangeinDebt	1.715**	-0.741	0.483	-1.852***	-1.986***	-0.582
	(0.803)	(1.182)	(1.366)	(0.419)	(0.633)	(0.610)
Collections	-0.569	-3.173***	-2.266***	-3.085***	-2.900***	-1.335***
	(0.480)	(0.365)	(0.507)	(0.667)	(0.269)	(0.368)
PaymentHistoryGradeCDF	13.701***	1.294	-2.736	-6.390**	-2.767***	-3.003***
	(3.948)	(4.534)	(2.915)	(2.777)	(0.759)	(0.858)
CreditAgeGradeAB	0.757	-1.219	-8.945	16.438***	11.677***	9.459***
	(11.050)	(10.654)	(6.437)	(5.669)	(2.634)	(2.098)
CreditAgeGradeCDF	-4.294	-15.228*	-19.194***	8.704	6.314***	3.203
	(10.212)	(8.018)	(7.400)	(5.456)	(2.070)	(1.991)
CreditUtilizationNotA	5.864	-0.948	-5.138*	2.703	1.003	0.971
	(7.552)	(6.948)	(3.013)	(3.355)	(2.062)	(1.368)
ChangeinCreditUtilizationGrade	-18.596**	0.090	36.626***	0.163	-5.040***	-4.146***
	(7.673)	(10.331)	(8.070)	(2.582)	(0.530)	(1.066)
DebtIncrease	-6.874***	-7.102***	-2.738***	-3.610***	-7.126***	-5.354***
	(0.647)	(0.338)	(0.284)	(0.447)	(0.285)	(0.274)
CSIncrease	0.000	0.000	0.000	1.041	-3.534***	-3.620***
	(.)	(.)	(.)	(0.730)	(0.432)	(0.246)
CSIncrease2016	4.262***	3.052***	2.853***	-0.909	-2.718***	-1.097**
	(1.394)	(1.102)	(0.828)	(0.769)	(0.388)	(0.428)
Constant	30.307***	0.611	-0.532	17.429***	-9.875***	-17.801***
	(2.174)	(3.008)	(2.642)	(1.505)	(1.548)	(1.513)
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
Tercile	Bottom	Middle	Upper	Bottom	Middle	Upper
ΔCS^{pre}	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
Observations	51,185	81,078	91,650	71,803	134,781	179,986

Notes: This table presents the IV regression of Equation 5. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent.

Table 46: Two-Stage Least Squares IV By Tercile - Full Regression - Monthly Change

	ΔCS^{post}					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{Check\}$	-3.795**	1.450	-2.106	5.047***	2.190***	1.186**
	(1.853)	(2.365)	(1.925)	(0.658)	(0.584)	(0.493)
active	-0.589***	-0.584***	-0.152	-0.312*	-0.373***	-0.258***
	(0.112)	(0.174)	(0.115)	(0.165)	(0.129)	(0.076)
Homeowner	0.719***	1.029***	0.753***	0.443***	0.789***	0.724***
	(0.187)	(0.093)	(0.061)	(0.120)	(0.089)	(0.042)
ccclicks	-0.400*	-1.218***	-1.433***	-0.673***	-1.070***	-1.404***
	(0.211)	(0.222)	(0.178)	(0.092)	(0.111)	(0.153)
refreshSameMonth	2.650***	1.315**	0.700*	0.838***	0.482**	-0.229
	(0.390)	(0.575)	(0.422)	(0.210)	(0.229)	(0.163)
CreditUtilizationGradeAB	0.142	0.440	0.140	0.380	-0.173	-0.700*
	(0.415)	(0.501)	(0.301)	(0.966)	(0.304)	(0.386)
CreditUtilizationGradeDF	-0.249	-0.169	-0.568*	0.103	-0.494***	-0.717***
	(0.330)	(0.299)	(0.313)	(0.749)	(0.188)	(0.214)
ChangeinDebt	0.388	-0.432	-0.208	-0.390***	-0.419**	0.033
	(0.248)	(0.365)	(0.384)	(0.118)	(0.170)	(0.159)
Collections	-0.277*	-0.877***	-0.553***	-0.939***	-0.805***	-0.264***
	(0.145)	(0.092)	(0.108)	(0.189)	(0.078)	(0.078)
PaymentHistoryGradeCDF	0.872***	0.466	0.013	-1.146	-0.880***	-0.614***
	(0.287)	(0.303)	(0.207)	(0.823)	(0.189)	(0.199)
CreditAgeGradeAB	-3.511***	0.178	-0.703	2.038***	2.722***	2.441***
	(1.008)	(1.514)	(1.198)	(0.747)	(0.455)	(0.591)
CreditAgeGradeCDF	-3.660***	-0.275	-1.137	0.734	1.590***	1.337**
	(0.943)	(1.425)	(1.245)	(0.995)	(0.536)	(0.565)
CreditUtilizationNotA	0.108	0.394	-0.272	1.686*	1.138***	0.807***
	(0.495)	(0.548)	(0.204)	(0.942)	(0.308)	(0.297)
ChangeinCreditUtilizationGrade	2.049	-7.389**	6.311	0.137	-1.789***	-1.809***
	(1.576)	(3.629)	(4.070)	(0.705)	(0.113)	(0.156)
DebtIncrease	-1.501***	-1.574***	-0.622***	-0.808***	-1.840***	-1.351***
	(0.168)	(0.101)	(0.090)	(0.117)	(0.097)	(0.073)
CSIncrease	0.000	0.000	0.000	0.286*	-0.812***	-1.004***
	(.)	(.)	(.)	(0.164)	(0.118)	(0.064)
CSIncrease2016	1.470***	0.707**	0.679***	-0.493***	-0.566***	-0.055
	(0.370)	(0.342)	(0.236)	(0.183)	(0.129)	(0.110)
Constant	5.284***	0.044	0.379	1.130***	-0.269	-1.531***
	(0.646)	(0.917)	(0.804)	(0.364)	(0.322)	(0.283)
Campaign Targeting Controls	Y	Y	Y	Y	Y	Y
Tercile	Bottom	Middle	Upper	Bottom	Middle	Upper
ΔCS^{pre}	< 0	< 0	< 0	≥ 0	≥ 0	≥ 0
Observations	51,185	81,078	91,650	71,803	134,781	179,986

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores.

Table 47: Two-Stage Least Squares IV Including Pre Trend As Covariate

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-16.463**	7.697***
	(6.783)	(1.996)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	223,913	386,570

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. An additional covariate of the pre-trend credit score change is included.

Table 48: Two-Stage Least Squares IV Only Users Whose Pre-Trend Change was 10 Points Or Fewer In Magnitude

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-11.783**	5.016**
	(5.959)	(2.247)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	104,693	235,417

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. Only users whose pre-trend is ten points or fewer in magnitude are included.

Table 49: Two-Stage Least Squares IV with Last Credit Score as Covariate

	ΔCS^{post}	
	(1)	(2)
$\mathbb{1}\{Check\}$	-13.654**	13.393***
	(6.130)	(1.675)
Campaign Targeting Controls	Y	Y
ΔCS^{pre}	< 0	≥ 0
Observations	223,913	386,570

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. The users last credits score observed before the email is added an additional covariate.

Table 50: Two-Stage Least Squares IV Grouped by Refresh Indicator

	ΔCS^{post}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-11.747*	5.482**	-5.482	10.788
	(6.204)	(2.565)	(8.022)	(7.213)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Refresh Same Month	Y	Y	N	N
Observations	118,759	235,881	105,154	150,689

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. Refresh Same Month is an indicator for whether the user had more than one credit score refresh that month (which was an extra credit score report). Yes indicates that the user had this and No that they did not.

Table 51: Two-Stage Least Squares IV Grouped by Active Indicator

	ΔCS^{post}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-22.835***	8.881***	-12.421	7.646
	(6.862)	(2.139)	(12.023)	(7.091)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Active	Y	Y	N	N
Observations	208,889	365,515	15,024	21,055

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. Active is a campaign control targeting covariate. It is an indicator for whether the user has created an account, opened or clicked on an email in the last three months.

Table 52: Two-Stage Least Squares IV Grouped by Credit Card Offer Click Indicator

	ΔCS^{post}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-5.687	2.353	-22.227***	9.972***
	(8.778)	(4.526)	(6.763)	(2.029)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Credit Card Clicks	> 0	> 0	0	0
Observations	11,518	28,422	212,395	358,148

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. Click is an indicator for whether the user has clicked on a credit card application in the same month.

Table 53: Two-Stage Least Squares IV Grouped by All Activity Measures

	ΔCS^{post}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-22.835***	8.881***	-12.421	7.646
	(6.862)	(2.139)	(12.023)	(7.091)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	≥ 0	< 0	≥ 0
Activity	Y	Y	N	N
Observations	208,889	365,515	15,024	21,055

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. Activity is an indicator for whether the user had refresh same month, active, and click all zero or all one.*

Table 54: Two-Stage Least Squares IV Grouped by Average Rate of Checking Credit Report

	ΔCS^{post}			
	(1)	(2)	(3)	(4)
$\mathbb{1}\{Check\}$	-14.478**	-15.436**	8.284***	10.416***
	(6.285)	(6.589)	(2.788)	(2.371)
Campaign Targeting Controls	Y	Y	Y	Y
ΔCS^{pre}	< 0	< 0	≥ 0	≥ 0
Rate of Checking Score	Above	Below	Above	Below
Observations	118,083	105,812	215,789	170,698

Notes: This table presents the IV regression of Equation 6. $\mathbb{1}\{Check\}$ *is an indicator for whether a user checked her score within 31 days of receiving the email. This regressor is instrumented for which email version a user received. The outcome is the change in credit score from before the email to the first credit score checked at least two months after the email was sent divided by the number of months between the credit scores. The average rate number of checks per month that all users have is calculated which is, 0.89. There are many points above one as some users get extra refreshes within a month, or if a user checks on June 1 and July 1 it counts as twice in a month due to the way it is calculated. The users are then split according to whether they check more or less frequently per month than the average. Users who check their credit score often before the email is sent, are less likely to have their retention impacted by the email. Therefore, we split users into whether or not they check their credit score more or less often than average (compared to other users who receive a focal email).*