Using AI and Behavioral Finance to Cope With Limited Attention and Reduce Overdraft Fees

Daniel Ben-David[a,b], Ido Mintz[b], and Orly Sade[a]

Abstract

In a randomized field experiment using a large personal financial management platform operating in the United States and Canada, we investigate mechanisms to reduce overdraft fees. A sample of users identified via an artificial intelligence (AI) algorithm as having a propensity to overdraw their accounts were sent reminder notices to test the efficacy of different framings in reducing the number of overdraft fees. Employing parametric identifications, as well as time-to-event semi-parametric analysis to learn that sending a reminder proved effective in and of itself, and the impact was significantly enhanced by simplifying the message. A negative framing of the simplified version elicited greater engagement and had a stronger impact than a positive framing. Significant effects are seen predominantly among users with medium to high annual incomes. We relate our findings to the literature on limited attention and the ostrich phenomenon. Our work also contributes to the literatures on financial technology, AI, and human-computer interaction.

Keywords: artificial intelligence, behavioral finance, overdraft, limited attention, ostrich

[a] Finance Department, Jerusalem School of Business, The Hebrew University of Jerusalem, Israel. [b] Intuit. E-mail addresses: daniel.ben-david@mail.huji.ac.il (D. Ben-David), ido_mintz@intuit.com (I. Mintz), and orlysade@huji.ac.il (O. Sade).
We have benefited from comments by Ori Heffetz, Avner Kalay, Olivia Mitchell, Lev Muchnik, Michaela Pagel, Eyal Winter, Yishay Yafeh, Stephen Zeldes, and Asaf Zussman as well as the participants at the 2019 AFE University of Chicago, 2019 Exeter Behavioral Prize Workshop, Exeter, Britain, 2019; the Intuit AI Summit Conference, 2019; and the Hebrew University of Jerusalem finance and behavioral economics seminars. Sade acknowledges financial support provided for this research by the Krueger Center of Finance at the Hebrew University of Jerusalem.

1. Introduction

Overdrafts are a widespread global phenomenon. An overdraft occurs when a transaction leads to a balance below zero in an individual's bank account. Banking systems around the world offer different services for dealing with their customers' overdrafts. In the United States, a common service is overdraft protection, whereby the bank will cover certain transactions that trigger an overdraft, mostly in return for what is termed an "overdraft fee." A non-sufficient funds (NSF) fee is imposed when the bank declines the transaction. These two fees apply irrespective of whether the customer is registered for such a protection service.[1] In addition to the overdraft or NSF fee, banks may charge a daily fee or interest for the negative balance amount remaining. Overdraft and NSF fees are widespread in the retail banking system: In the United States alone, retail banks charge $17 billion annually as NSF fees,[2] which accounts for approximately 6% of their total revenues (Stango & Zinman, 2014) and for two thirds of the income they earn from fees.[3]

The economic literature offers several competing explanations for why people overdraw their accounts and as a consequence pay overdraft fees. The first is based on complete rational information modeling. Accordingly, fully rational individuals may consciously decide to overdraw their accounts and pay the fees after evaluating all the benefits of consumption and the associated overdraft costs. A second line of explanations assume limited attention or

[1] NSF and overdraft fees are commonly aligned and usually range between $24 and $35; one or the other fee is paid for each transaction leading to an overdraft. In the context of our inquiry, no distinction is made between the payment of overdraft fees and NSF fees, and these two terms are used interchangeably.

[2] According to the Consumer Financial Protection Bureau (2017), in the United States, overdraft fees amounted to $17 billion. Several private institutions estimate that amount to be higher: In 2017, consumers paid $34.3 billion, and in 2016 $33.3 billion (Moebs Services and Products, 2017). According to Melzer and Morgan (2015), in 2007, consumers in the United States paid $23 billion.

[3] According to the Consumer Financial Protection Bureau (2017), overdraft fees constituted around two-thirds of all retail bank fees.

rational inattention.[4] Because the process of acquiring information can be costly or complex, one might decide to overdraw and pay the associated fees, after considering only incomplete information. A third, behavioral, explanation follows the line of research that explores the so-called ostrich effect and information-dependent utility models.[5] Roughly, individuals may overdraw and as a result pay overdraft fees because they tend to avoid acquiring information at unpleasant times.

The focus of this project is to investigate competing treatments aimed at reducing the tendency to pay fees that occur when overdrawing one's account for people who are prone to paying them, while addressing different explanations suggested by the literature. For this we use Mint, a web-based personal financial management (PFM) application with over 20 million users in the United States and Canada. We explore whether we can reduce the number of overdraft fees paid by Mint customers who have been screened for a high likelihood of paying an overdraft fee in the immediate future. Mint's customers connect their bank accounts and credit cards to the application, which produces a personalized integrative interface. It also offers advanced services such as personal financial insights, recommendations, and budget planning. We focus on Mint's insight service called Overdraft Prediction (ODP), which is a machine learning algorithm that computes the likelihood of an individual running a below-zero balance in the coming week. If the score exceeds a certain threshold, the algorithm triggers an e-mail message to be sent to the individual concerned.

In a randomized field experiment, we test whether sending recurrent reminder messages changes customers' behavior, and whether the phrasing and structure of such reminders make

[4] Inattention has been theoreticaly investigatigated with regard to economic decisions in various contexes, e.g., Chetty et al., 2009; DellaVigna, 2009; Gabaix et al., 2006; Hirshleifer & Teoh, 2003; Karlan et al., 2016; Sims, 2010. A more comprehensive review of the potential explanations is found in Section 3.

[5] The ostrich effect bias was initially coined by Galai and Sade (2006) and later extended by others. For examples of information-dependent utility models, see Koszegi and Rabin (2009) and Andries and Haddad (2017), among others. We provide a more elaborated review of the literature in Section 3.

a difference (the experimental procedure is laid out in detail in Figure 1). We hypothesize that receiving a reminder may have an impact on individuals with a propensity for "limited attention," but also that certain designs of the reminder messages, especially simplicity, may intensify this effect. Furthermore, we conjecture that framing the e-mail notifications differently, with a negative or positive valence, will have different effects on individuals who are prone to limited attention or to ostrich behavior.

The experiment was conducted over four months in 2018. The Mint ODP users sampled were assigned to one of four conditions: (1) The Control group received no message; (2) the Base group received a message previously used by Mint (Figure 2, Base); (3) the Simplified Avoid group received a negatively framed simplified message (Figure 2, Simplified Avoid); and (4) the Simplified Save group a positively framed simplified message (Figure 2, Simplified Save). The motivation for each of these rubrics is provided in the literature review.

[Figure 2]

The sample comprised 39,607 unique Mint users, randomly assigned to the above four groups. Of these, 22,775 were return ODP users and 16,832 (~42.5%) were first-time ODP users. Prior to the commencement of the experiment, the return users had received, at least once, the Base message issued by the Mint ODP service, while first-timers received their first message as part of the experiment. We control for the first-time effect in the main empirical analysis (including a propensity score matching analysis). The sample had a mean of 0.885 NSF fees one week after receiving their first reminder, and 81% experienced at least one NSF/overdraft fee during the experiment.

To analyze the experimental results, we use three key indicators: the open rate,[6] the click-to-explore rates,[7] and the reduction in number of NSF/overdraft fees. The first two indicators are important for better understanding the mechanism by which the reminders impact the final fee outcome.[8] For our estimation of the third, goal indicator, reduction of NSF/overdraft fees, we employ parametric identification, as well as a time-to-event (TTE) semi-parametric analysis.[9]

We find that sending a reminder (the Base message), in and of itself, proved a significant help in addressing inattention and reducing overdraft fees in the short term. Compared to the Control group, the mean number of NSF fees per person in the Base group was reduced by 5.3% one week after receiving their first reminder.[10]

Next, we focus on the structure of the reminder. Simplifying the message substantially enhanced its influence by further reducing the cost of acquiring financial information. Receiving a simplified message (keeping the Avoid, negative framing) improved the open rate of e-mails sent to the participants by 12.1%[11] and the click-to-explore rate (via a link in the message) by 78.4%.[12] The simplified message reduced the mean number of NSF fees per person, compared to participants in the Control group, by 9.25% one week after receiving the first notice and increased its impact by 74.5% compared to the effect of the Base message.[13]

6 Users are assigned the value of 1 if they opened at least one reminder in the course of the study and 0 otherwise. The rate is calculated by dividing the sum of all participants in a group by the group's total number.

7 Users are assigned the value of 1 if they entered the app by clicking a link in any reminder e-mail and 0 otherwise. The rate is calculated by dividing the sum of all participants in a group by the group's total number.

8 The users who open the reminder alert receive further information about potential actions they can pursue; clicks on links to the Mint app can attest to the willingness to explore the advice.

9 We follow Prentice et al.'s (1981) conditional risk set model for recurrent events.

10 Different time frames are presented in the Results section (and 3.3% for the duration of the entire experiment of four months). This result is consistent with Caflisch et al. (2018), Karlan et al.(2016), and Medina (2017), among others.

11 The open rate was 30.85% in the Simplified Avoid group compared to 27.5% in the Base group.

12 The click-to-explore rate was 2.5% in the Simplified Avoid group compared to 1.4% in the Base group.

13 Different time frames are presented in the Results section.

Interpreting the results in U.S. dollar, across the entire experiment (four months), sending the Base and Simplified Avoid reminders saved each individual, on average, $9.33 and $25.16, respectively, compared to the Control group, which did not receive any treatments. Unlike other field experiments, which test the impact of a one-time reminder, we investigate the efficacy of an algorithm that creates several reminders, as required, over the course of the experiment. Accordingly, each of the reminder treatments used in our experiment may not only trigger an action related to that specific event, but also affect the frequency of additional, subsequent reminders. Hence, we elected to use TTE analysis for recurrent events to investigate the effect of each treatment on the risk of paying NSF/overdraft fees during the entire experiment. The results of the TTE semi-parametric analysis indicate that the simplified reminder significantly changed participants' behavior and thereby lowered the risk of their incurring overdraft fees for the duration of the experiment.

A comparison of the differently framed versions, the Simplified Avoid (negative) and the Simplified Save (positive) message, yielded the following results: (1) There was no significant difference in the open rates for the e-mail messages[14]; (2) the Avoid framing produced a higher click-to-explore rate, increasing exploration of the Mint ODP service by a significant 65.6% compared to the Save version[15]; and (3) importantly, the Avoid version emerged as significantly more effective, by 81.6%, in reducing the number of NSF fees in the long term than the Save version.[16] Similar results were obtained for the TTE analysis, in which the Simplified Avoid message outperformed the Simplified Save message in changing the risk for additional NSFs over time. For the full 4-month duration of the experiment, in monetary

[14] The open rate was 30.85% and 30.51%, for the Simplified Avoid and Simplified Save message, respectively.
[15] The click-to-explore rate was 2.5% and 1.51% for the Simplified Avoid and Simplified Save message, respectively.
[16] There was a 9.25% mean reduction after one week and 8.9% after four months in the Simplified Avoid message and 9.44% but only 4.9% for the Simplified Save message, over the same time frames.

terms, receiving the Simplified Avoid message saved users, on average, $11.33 more than the Simplified Save message.[17]

Our research also tests the effect of the reminders on individuals with different levels of income. In particular, we seek (a) to establish whether the reminders created by the algorithm are more effective in preventing fees for individuals with lower versus higher annual incomes, and (b) to relate our findings to the limited-attention and ostrich-effect literature.[18] First, we compare the impact of receiving any reminder to receiving no reminder (the Control condition) across six income levels.[19] Next, we attempt to identify the potential "ostriches" by analyzing the difference in responses to a negatively versus positively framed message, on the assumption that ostriches will selectively read a larger number of positive messages.

The results show that the algorithm-driven reminders had the greatest effect among individuals with incomes above the 25th percentile.[20] Moreover, in the long term, we established a behavior change among the higher income groups, in the 75th percentile and above. Specifically, participants with higher incomes opened the reminders at a greater rate, which increased the overall effectiveness of the reminder messages sent. This result can be driven by several salient factors: (1) Low-income individuals may have a greater propensity for "ostrich" behavior; (2) they may have substantial additional burdens, such that they experience higher levels of cognitive load that diminish their ability to pay due attention and handle multiple stimuli (e.g., Mani et al., 2013); and (3) they may simply overdraw their accounts in the absence of better credit alternatives.

[17] There was a reduction of $25.16 vs. $13.85 for the Simplified Avoid and Simplified Save message, respectively, compared to no message (Control). Such an outcome aligns with the goal-framing literature and adds to Karlan et al.'s findings (2016) in respect to savings.
[18] The parametric identification method is similar to that used by Olafsson and Pagel (2017, 2018).
[19] We utilize the four quartiles and split the lower and upper quartiles into two, using the bottom and top deciles, to explore potential notable changes at the tails of the distribution.
[20] Here the percentiles are calculated based on individual income. Following Baker (2018), we repeated the exercise using household income and report it in Section 8.

We initially hypothesized that positive messages would be more effective in reducing ostrich behavior among lower income groups, given the general hardships they are likely to experience. However, comparing the Simplified Avoid and Simplified Save conditions, we find no consistent differences in the open rates across the different income percentiles.

Our project contributes to the literature on overdraft fees, AI, and behavioral economics. First, to the best of our knowledge, we are the first to document the effect of using an machine learning (ML) algorithm to produce recurrent reminders designed to reduce overdraft fees on a population with a tendency to pay such fees. Second, we propose a time-to-event semi-parametric estimation method, which has been used predominantly in clinical studies, that takes into account both the time and the NSF/overdraft fee events to evaluate the effectiveness of the treatments. Third, we document the additional effect of applying behavioral-finance-based treatments beyond the advantages of reducing the information-gathering costs obtained by using the AI algorithm. Specifically, in addition to investigating the effect of reminder alerts as such, we also compare the impact of different frames based on the behavioral economics literature; we find (a) that simplification plays a major role in addressing limited attention and (b) that the influence of negatively framed messages translates into better engagement and endures longer than that of positively framed ones. Fourth, we contribute to the literature that investigates the effect of reminders on different segments of the population, as we also find an indication that low-income individuals—who generally display a propensity to pay overdraft fees, possibly as a result of their complex financial situation—tend to behave as ostriches, and that therefore, for them, reducing the costs of obtaining information may not be a sufficient solution.

The rest of the paper is organized as follows. Next we elaborate on the overdraft phenomenon (Section 2), possible reasons for such transactions (Section 3), and the study settings (Section 4). These rubrics are followed by the research hypotheses (Section 5). The data and the experimental design are presented in Section 6. We then proceed by detailing the results (Section 7) and formulating a summary and offering concluding remarks (Section 8).

2. Overdrafts as a general phenomenon and in the United States

An overdraft occurs when a withdrawal causes the balance of a person's bank account to fall below zero. Banking systems around the world offer different services for dealing with customers' overdrafts. One such service is the stipulation of a limit amount, that is, a line of credit that establishes the maximum loan amount the customer can borrow before overdraft fees and/or interest are charged. When a customer causes the account's balance to fall below zero or exceeds the credit limit, if applied, the bank decides whether to allow or reject further transactions and what fees and interest rates to charge—these are for the most part fixed and relatively high. In addition to the initial fee imposed on customers for exceeding the limit and overdrawing their accounts, some banks charge periodic fees for maintaining an overdraft.

In the United States and Canada, a common service banks offer their customers is *overdraft protection*, whereby the bank will cover certain transactions that trigger an overdraft, and usually charge an overdraft fee. When the bank does not accept a transaction, an NSF fee is charged, irrespective of whether the customer is registered for the overdraft protection service. In addition to the NSF fee, banks usually charge a fee or interest for each day the account remains overdrawn. The overdraft protection service automatically approves costs such as periodic transactions, automated teller machine (ATM) withdrawals, checks, debit transactions, and more. As of 2010, ATM cash withdrawals or debit transactions require the customer's explicit authorization.[21] Small U.S. banks offer a similar service called *bounce protection*, which is usually managed by a third party according to a specific bank policy, such that not all overdraft transactions are automatically covered.

[21] The Federal Reserve's opt-in Regulation E.

In addition to the services described above, some banks offer their customers *overdraft transfer protection*. This service connects a person's accounts or credit cards, such that one account or card can be used to cover an overdraft on another. Customers pay a fee of $5 to $10 per activation for this service (Stango & Zinman, 2014).

As specified above, overdraft and NSF fees are pervasive in the retail banking system: In the United States alone, retail banks charge $17 billion annually as NSFs, which account for approximately 6% of their total revenues (Stango & Zinman, 2014) and for 2/3 of their income accrued from fees. Elsewhere in the world, in February 2018, the U.K. Competition and Markets Authority imposed a regulation whereby personal current accounts are automatically enrolled in a program of reminders to notify them when they have exceeded or attempted to exceed a pre-agreed credit limit and will incur a charge (Caflisch et al., 2018).

3. Rationality, limited attention, selective attention, and the ostrich effect

As already stated, the literature offers several explanations for overdraft behavior, based on the complete relational information modeling, on limited-attention models for acquiring information, and selective-attention models such as the ostrich effect. According to the rational models, people make consistent decisions after evaluating all available information and fully maximizing their total lifetime utility. With respect to overdrafts, individuals are assumed to make a conscious choice to pay the fees after thoroughly evaluating all the benefits of consumption, the associated overdraft cost, and alternative financing sources. Our data indicate, however, that more than 50% of NSF fee payments occur as a result of transactions involving less than $50. In the rational model, an individual in our sample would evaluate all the information and decide to pay a $35 overdraft fee in order to buy a $4 cup of coffee. Though such a scenario is possible, it is not very plausible. According to the rational model, sending reminders regarding overdraft fees would be unnecessary as it would not affect people's decisions.

Limited-attention or rational-inattention models posit that the process of acquiring information is complex (Bossaerts et al., 2019; Franco et al., 2018) or costly (direct or indirect costs), so people may make financial decisions after considering only a portion of the information available (e.g., Chetty et al., 2009; DellaVigna, 2009; Gabaix et al., 2006; Hirshleifer & Teoh, 2003; Karlan et al., 2016; Sims, 2010, among others).[22] These models are supported by extensive empirical literature (e.g., Ashraf, Karlan, & Yin, 2006; Medina, 2017; Stango & Zinman, 2014, among others).[23] Regarding overdrafts, rational limited attention can serve as a possible explanation as some may find it costly to constantly monitor their bank balances or find it complicated to correctly predict the probability of an overdraft. In such cases, financial decisions may be based on partial information, resulting in overdraft fees. In this approach, a reminder sent by an algorithm that predicts overdrafts may reduce the costs associated with ongoing monitoring of one's accounts. Furthermore, a clearly phrased, simple message (as opposed to a longer, more technical text) may reduce the complexity of the message and facilitate comprehension and hence is likely to have a stronger effect in reducing fees. It is important to note that limited attention may be exacerbated in the digital era: Obtaining information is relatively cheap yet at the same time may be overwhelming, as

[22] Hirshleifer and Teoh (2003) study the effects of different pro forma disclosures on market prices; Gabaix et al. (2006) discussed costly information acquisition, modeling, and comparison between the cognition model and the rational model in the laboratory; DellaVigna (2009) provides models and a survey of empirical evidence in behavioral economics; Chetty et al. (2009) explore limited attention regarding salience and taxation; Karlan et al. (2016) study theory and evidence regarding the impact of reminders on limited attention and on savings. For a broad review of the literature on inattention, see Caplin (2016) and Gabaix (2017).

[23] Stango and Zinman (2014) present empirical work that is based on administering to panelists questionnaires on overdraft; the objective is to investigate the effect of the survey on overdraft fee payments. Medina (2017) presents empirical work on limited and selective attention in the case of credit card fees. Ashraf et al. (2006) present field experiments on the commitment savings product.

individuals are constantly exposed to multiple and frequent stimuli (e.g., Benartzi & Lehrer, 2015; Cowan, Johnson, & Scott Saults, 2005; Lindgaard, Fernandes, Dudek, & Brown, 2006).

A third possible explanation to consider rests on information-dependent utility models, such as those by Kőszegi and Rabin (2009), Nieuwerburgh and Veldkamp (2009), and Andries and Haddad (2017). These researchers posit that individuals choose to ignore relevant information even though this can result in a cost.

A behavioral explanation relevant to our inquiry is the ostrich bias effect, defined as the avoidance of acquiring information regarding risky or unpleasant situations (Galai & Sade, 2006; Karlsson et al., 2009). Phrased metaphorically, individuals could be willing to pay a price for the bliss of ignorance. Since an overdraft is a result of a withdrawal transaction leading to a negative account balance, the ostrich effect may propel individuals to eschew gathering information that could help them avoid overdrafts and paying penalties. We expect individuals who act as ostriches not to respond to reminders, as they are likely to entertain the vain hope that if they ignore the information, the menace will disappear. Among this population, responsiveness to reminders might be correlated with income, as the latter may be negatively correlated with financial concerns (e.g., Olafsson & Pagel, 2017). Furthermore, to the extent that ostriches accept economic pain to avoid negative news, they may respond better to positive messages, phrased in terms of saving, than to the negative framing, that is, warnings about penalty fees.

4. Setting

We use Mint, a large web-based Personal Financial Management application (PFM) that enables the monitoring of bank transactions, credit cards, investments, and loans. As already explained, Mint's customers connect their credit cards and bank accounts to the application, which produces an integrated personal interface. In addition, Mint offers advanced services such as personal financial insights, recommendations, and budget planning.

As elaborated above, one of Mint's advanced services is ODP, an AI algorithm that computes the likelihood of an individual initiating a transaction that exceeds an account's available funds, thereby putting the person in the "negative zone" in the coming week. If the score produced by the algorithm exceeds a certain threshold, an e-mail reminder message is sent to the user (see Figure 2, Base). The AI algorithm predicts these events a week ahead, allowing customers time to take action to prevent an overdraft. The algorithm is a gradient boosting decision tree model with periodic re-training.[24] The calculation of the score involves a set of features that can be conceptually divided into three groups, related to (a) user profile (e.g., the number of accounts owned, periodic income), (b) balance features (e.g., cashflow behavior and balance trends), and (c) transactions (e.g., previous NSF fees, number of small debit and credit transactions, frequency of purchases).

[24] During the experiment, the model was not re-trained to avoid selection on participants entry or with the notification messages sent.

5. Research questions, motivation behind the experimental hypotheses, and the manipulations tested

Our project builds on the behavioral economics studies that have shown that messages of various types can shape people's decisions and mitigate limited attention in different contexts. The focus of our randomized field experiment investigations is to learn whether the structure and framing of recurrent e-mail reminder messages make a difference in terms of overdraft fees.

We focus on (1) a simple versus more elaborate message and (2) a positive versus negative framing. In what follows, we further elaborate on the motivation and the specific hypotheses that propelled us to investigate each of these manipulations.

5.1. Simplicity

While receiving a reminder, in and of itself, may mitigate the limited-attention problem through reducing information cost and complexity, the phrasing of such a message can further influence these outcomes. An important aspect of the "information architecture" of a notification is its simplicity (e.g., Benartzi & Lehrer, 2015; Reinecke et al., 2013; Tuch et al., 2010; Zheng et al., 2009), especially given the abundance of screen stimuli today. Thus, Cowan and colleagues (2005) and Mormann and colleagues (2012) document that individuals tend to form impressions and make decisions based on information presented and processed in a very short time and with very little attention investment. In our experiment, we investigate the effect of a message designed and used by Mint ODP (Base), compared to that of a simpler message created for the purposes of this research. The simplification of the message involves (1) shortening the text, (2) giving the fee amount more salience by displaying it as part of the headline, (3) providing a related art that is associated with savings, (4) increasing the size of the "explore account" link, and (5) creating a related actionable text. The details are presented in Figure 2.

For robustness, we ran an online experiment survey using Amazon's Mturk platform[25] with a representative sample of 107 participants from the United States to establish that the new e-mail reminder is indeed perceived as simpler. The results confirm that the simplification of the reminder was perceived as intended by a significant margin; see the Appendix for details.

We hypothesize that, compared to receiving no treatment (Control condition), receiving reminders will affect individuals with a tendency toward limited attention. In addition, we hypothesize that, compared to long text messages, short and simple messages will be even more effective, as they further reduce the cost of gathering and processing information. At the same time, we expect that the level of simplicity of the reminders will have no effect on individuals prone to the ostrich bias, who may not even bother opening them. We also hypothesize that an increase in attention triggered by a reminder will result in actions that will reduce the mean number of NSF fees.

[25] Although there is an ongoing debate in the literature over Mturk sample representativeness and fit for specific tasks, it is a commonly used platform for economics, business, and psychology experiments. We use the Mturk sample for our treatment robustness check. Unlike some other survey methods or interviews, the participants of Mturk are web and internet users. This characteristic is of interest, given that in our primary empirical investigation, we investigate app users. Also, Mint users, as well as the Mturk sample, have heterogeneous characteristics, enabling us to obtain answers from a broad range of participants. Last, given that we provided Mturk participants with a straightforward task, we did not anticipate major barriers to perform it. See Goodman et al. (2013) for a detailed discussion about the advantages of using Mturk for tasks such as ours.

5.2. Positive versus negative framing

The idea of framing is introduced and developed by Tversky and Kahneman (1981), who show that a negative framing of a message, emphasizing a potential loss, triggers a different response from a positive framing, with a focus on potential benefits.[26] A rich discourse has subsequently evolved on the positive versus negative valence of messages: Some studies show that a positive utterance has a greater impact (e.g., Davis & Bobko, 1986; Levin & Gaeth, 1988, among others), while others maintain that responses to a negative utterance are more acute (e.g., Block & Keller, 1995; Meyerowitz & Chaiken, 1987, among others). A framing typology for the different situations explored was developed by Levin and colleagues (1998).[27]

Our project focuses on what the persuasion literature terms "goal framing" (Levin et al., 1998), which differentiates between emphasizing the benefit of obtaining a positive outcome versus avoiding a negative consequence. To test the goal-framing effect, we conduct the following manipulation: While keeping the body of the simplified message unchanged, we use two versions of the heading, one framed negatively (Avoid) and the other positively (Save). Although a substantial part of the literature on goal framing suggests that a negative framing has the stronger impact, within the setting of our experiment one might hypothesize that ostriches will be more affected by the positive message, since they tend to avoid negative news.

6. Data and experimental design

The experiment was carried out over four months in 2018 and involved two main stages (see Figure 1). First, the Mint ODP algorithm calculated a score for the Mint users sampled. If the score exceeded a certain threshold, the user was randomly assigned to one of the four

[26] See Levin et al. (1998) for a literature review.

[27] The typology distinguishes three types of framing: (1) the original risky choice framework developed by Kahneman and Tversky, (2) attribute framing for relatively good or bad potential outcomes, and (3) goal framing.

experimental groups[28]: (1) the Control group: users do not receive any reminder message; (2) the Base group: users receive the original message (Figure 2, Base); (3) the Simplified Avoid group: users receive a message with a simplified negative text framing (Figure 2, Simplified Avoid), and (4) the Simplified Save group: users receive a reminder with a simplified positive text framing (Figure 2, Simplified Save). Users exposed to one of the above treatments did not receive any additional messages during the following week. Each user could receive up to four messages over one month. All messages sent to the same user were of the same type. Overall, 39,607 unique users participated in the experiment. During the first month of the experiment, new participants could be assigned to one of the treatments, but after the first month, no additional participants were included. We observed all the participants in the sample for at least three additional months and for up to four months total.[29] An illustration of the time frame is presented in Figure 3.

[Figure 3]

The experiment involved two types of participants: return and first-time ODP users. The return users had received, at least once, the Base message prior to the commencement of the experiment, while first-timers received messages for the first time as part of the experiment. The larger proportion of participants (22,775) were return ODP users and about 42.5% (16,832) were ODP first-timers. We control for this important characteristic in the empirical analysis. Tables 1 and 2 present the descriptive statistics of the sample.

[Tables 1 and 2]

[28] The randomization was carried out using a hash function over a random ID number and not over stratified characteristics. ODP is a service provided as part of the product, no opt-in was required.
[29] No additional preliminary conditions were set for the experiment participants, such as the number of linked accounts, minimum app use duration, demographics, etc. Among the users in the experiment, no attrition was documented.

Of the 39,607 participants, 10,972 were randomly assigned to the Control group, 9,422 to the Base group (the original message), 9,725 to the Simplified Avoid group and 9,488 to the Simplified Save group. Over the four months of the experiment, 81% of the participants experienced at least one NSF or overdraft fee; the average number of fees for the total sample stood at 7.8 (SD = 11.01) with a median of 4. The open rate[30] was approximately 29.64% (SD = 0.45) and the average click-to-explore rate[31] was approximately 1.81% (SD = 0.16). Over the four months of the experiment, the average number of reminders sent to each participant was 3.2 (SD = 2.77).

The average annual income for the sample was approximately $108,000 ($SD$ = $1.02M) with a median of $53,200. The division of income into six groups was carried out according to the U.S. 2018 individual income quantiles.

6.1. The experiment and the data: Advantages and challenges

Our unique proprietary data and the use of a field experiment provide several advantages. (1) Our access to Mint's data enabled us to observe and investigate detailed data on many users: Specifically, we could observe the sample's bank activity, credit cards, investments, loans, and transactions, including overdraft and NSF fees.[32] (2) Within the framework of a controlled field experiment, we knew exactly when each reminder was sent and could directly monitor the participants' subsequent behavior. (3) Participants were randomly assigned to experimental and control conditions, which enabled us to identify the causal effect of receiving a message and the marginal effect of the different versions. (4) The Mint application also enabled us to monitor user engagement subsequent to each reminder by

[30] As specified above, users are assigned the value of 1 if they opened at least one reminder in the course of the study and 0 otherwise. The rate is calculated by dividing the sum of all participants in a group by the group's total number.

[31] As specified above, users are assigned the value of 1 if they entered the app by clicking a link in any reminder e-mail and 0 otherwise. The rate is calculated by dividing the sum of all participants in a group by the group's total number.

[32] See Gathergood et al. (2019) for a broader review of the importance of linked-data apps.

observing and measuring the open rates and the click-to-explore rates. (5) We were able to investigate the effect of recurrent reminder events, as some of the participants received more than one reminder. (6) Last but not least, we could test the combined effect of the financial innovation introduced by the application (the algorithm that predicts overdrafts) with the various phrasings of reminders, in an endeavor to find a more effective means to reduce overdraft fees.

The above merits notwithstanding, our data and experiment present several challenges. (1) Individuals who opt to use a financial application may not be representative of the general population,[33] as they likely possess higher levels of awareness and attention. One could therefore argue that our sample can be characterized by selection in terms of user interest and actions. Yet, while this concern could be relevant in comparing the app users to the general population, within the app user population, we employ random assignment. (2) We take into consideration only those individuals who are projected to pay NSF fees. The probability score determined by the AI algorithm and the threshold chosen for sending the message is influenced by previous overdraft fees that occurred in similar cases. Indeed, the population investigated in the experiment frequently pay overdraft fees (with a weekly mean of 0.885 NSF fees). Hence, our sample is not representative of the North American population who use financial applications or even of all Mint users. In our view, however, the focus on this particular population renders the impact of the reminders' information architecture more clearly distinguishable, since we examine this effect in complex financial contexts for individuals who are more likely to pay overdraft fees.[34] (3) The experiment was conducted over four months, and the effects targeted were monitored during that time period alone. Any subsequent activity

[33] See Baker (2018) on large account-linked-data representativeness.
[34] As stated, 81% of participants encountered at least one fee. Additional evidence was found by Caflisch et al. (2018), whose U.K. study compared a similar population with two other groups that have lower rates of overdraft occurrences.

is thus beyond the scope of this study. It has been suggested (e.g., Stango & Zinman, 2014) that the effect of notifications decreases over time. One advantage that our setting affords is that the experiment also explores the effect of reminders on first-time compared to return users. This difference, in our opinion, partially offsets the limitation associated with the short duration of the experiment. (4) Prior to the experiment, in the event of an overdraft risk, Mint ODP sent a long version of the notification via e-mail. Consequently, some of our participants had previously received the Base message, while the rest received an experimental version for the first time. The effect of reminders on first-time users may differ on account of novelty, and in addition, that population may inherently be at a lower risk of an overdraft—that is, this could be the reason they had not been sent a reminder previously. To address the potential difference between these two populations, our analysis controls for the first-time effect, and we also add a propensity score matching analysis in the robustness section. (5) Our income analysis[35] is based on accounts users registered on the application. It is of course possible that a user may not have reported all of his or her accounts, and therefore we cannot guarantee that the information gleaned from the application reflects each user's total income. Although including only partial information contravenes most people's motivation to use the application, we must be cautious in that regard. Furthermore, no clear distinction can be made between individual and partnered household incomes.

[35] We calculate the periodic income based on wages, investment interest, dividends, and rentals.

7. Results

Our main research hypotheses as described earlier are as follows:

Hypothesis 1: *Sending reminders will mitigate limited attention and reduce overdraft fees*.

Hypothesis 2: *Simplicity will help intensify the impact of the message.*

Hypothesis 3: *The Simplified Avoid version will have a different effect from the Simplified Save version.*

The impact of the reminders is estimated based on three key indicators: (1) the open rate; (2) the click-to-explore rate; and (3) the reminders' impact on the number of NSF/overdraft fees paid by individuals in the course of the experiment. Items (1) and (2) are used as proxies for interest and engagement: The users who open the reminder receive additional information about several potential actions they can pursue. Clicks to the Mint app can serve as verification of the willingness to explore the advice and hence relevant.

The impact of reminders on overdraft fees is measured in several ways. We report in the text the results of the following three measures: (a) time periods after the first reminder (TAFs): We report the effect one week after participants received the first experimental version—with the aim of measuring the immediate effect of the algorithm prediction[36]; (b) total time of the experiment (TT): for the four months of the experiment's duration, starting from the beginning; and (c) a semi-parametric time-to-event (TTE) analysis for recurrent events.[37]

In addition to the main analysis, we investigate the results obtained in light of human behavior theory, namely, the limited-attention and ostrich phenomena, and differentiate between annual income levels with respect to each.

7.1. The impact of reminders and the importance of simplicity

[36] For robustness, the regression tables also report the results for two weeks and one month after the first alert reminder.

[37] We follow Prentice et al. (1981) in applying a conditional risk set model for recurrent events. See Section 7.1 for details about the estimation procedure.

We start by examining the effect of sending the Base ODP reminder on the number of overdraft fees paid by the participants (Table 3). Compared to the Control group, the Base group showed a significant reduction of 5.3% (from 0.91 to 0.86) in the mean number of NSF fees per individual one week after the first reminder and 3.3% (from 8.08 to 7.82) at four months. The mean open rate and click-to-explore rate for any one of the Base messages was 27.5% and 1.4%, respectively, at four months.

[Table 3]

Next, we document the effect of simplifying the message. The Simplified Avoid message, which kept the negative valence of the Base message, increased the open rate (30.85%) by a significant 12.1% and the click-to-explore rate (2.5%) by 78.4% compared to the Base message. Compared to no message (Control condition), the Simplified Avoid message reduced the mean number of NSF fees per person by a significant 9.25% one week after the first reminder and by 8.9% for the full four months. This means that compared to the Base reminder, the Simplified Avoid version changed the impact of the notices by 74.5% after one week and by 113.6% for the experiment period (four months). Translating the impact into dollar terms, for the four months of the experiment, sending the Base and Simplified Avoid reminders saved each individual, on average, $9.33 and $25.16, respectively.

We proceed by conducting a multivariate analysis while controlling for the type of message and for the first-time effect. For the regression analysis of the open rates and click-to-explore rates, we use the following probit models[38]:

(1) $Open_i = \alpha + \beta_1 Simplified_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$

(2) $Click\ to\ explore_i = \alpha + \beta_1 Simplified_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$

[38] The results are not sensitive to different regression models such as ordinary least squares (OLS) or logit.

where the outcome variable $Open_i$ in model (1) and $Click\ to\ explore_i$ in model (2) equals the value of 0 or 1 for at least one reminder for each individual i. $Simplified$ is a dummy variable for a simplified treatment where the base alternative is the original message (Base). $First$ is a dummy for first-timers and $Simplified\_First$ is the interaction variable between the two. In both models, $\log(Yearly\ income)$ represents the individual's log yearly income. Table 4 shows the results of the analysis.

[Table 4]

Regression results indicate that the Simplified Avoid version significantly increased the open rates and the click-to-explore rates compared to the Base message. These results support our hypothesis that simplicity increases engagement.[39]

For the NSF fees regression estimation, we use the same specification for the TAF and TT analysis (cross-section for different time frames) of the tobit model on the log (1+ NSF period) as the dependent variable (the values are listed below).[40] We start by testing, in model (3), the impact of the Base message compared to no message (Control condition) and proceed by comparing the simplified version with the Base message in model (4).

$$(3)\quad \log\left(1 + NSF_{Period,i}\right) = \alpha + \beta_1 Base_i + \beta_2 First_i + \beta_3 Base\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$$

$$(4)\quad \log\left(1 + NSF_{Period,i}\right) = \alpha + \beta_1 Simplified_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$$

where the dependent variables are the log of the periodic NSF fees + 1. The periods are one week, two weeks, and one month after receiving the first reminder, and the full four-month experimental period.

[39] The analysis is not sensitive to whether we use a measure of open and click-to-explore rates that is based on at least one of the several reminders that were sent to each participant as specified above or on taking into account every single reminder.

[40] The regression results are not sensitive to the estimation method. Results are sustained with OLS regressions, with and without the yearly income variable. These effects are analyzed below.

In model (3) the variable $Base$ is a dummy variable for the original Base message (where the underlying alternative is the Control group). In model (4), the variable $Simplified$ is a dummy variable for the simplified version with the negative valence (where the underlying alternative is the original Base message). $First$ is a dummy for first-timers and $Base\_First$/$Simplified\_First$ are the interaction variables between the two. In both models, $\log(Yearly\ income)$ represents the individual's log yearly income. The results are displayed in Table 5.

Unlike other field experiments, which focus on the outcomes of a one-time reminder, we investigate the effect on NSF fees of an algorithm that creates several reminders, as required during the experiment period. Accordingly, each of the treatments implemented may not only trigger an action related to a specific event but also affect the frequency of further treatments. Hence, in addition to the cross-sectional analysis of the different time frames, we follow Prentice and colleagues (1981), who developed a variation on the traditional survival analysis, and apply conditional risk set models of TTE analysis.[41] This procedure enables us to analyze the potential differences in NSF fees between the groups while accounting for the time at which the event (NSF fee) occurred. These models incorporate time-dependent covariates (for each participant) with recurrent and ordered events (NSF fees over time in which the risk for the next NSF fee event is only for those who experienced an event on the previous stratum). We use the semi-parametric identification of the Cox model to relax the predefined distribution of the hazard to pay an NSF fee. In such an analysis, the dependent variables are both the number of NSF fees and the time at which the events occurred.

Hence, our hazard estimation function is as follows:

$$(5)\ \lambda_{ik}(t) = \lambda_{0\mathrm{k}}(t) * e^{X_{ik}\beta}$$

[41] For a more detailed account of the methodology and comparisons, see Amorim and Cai (2015) and Prentice et al. (1981).

Where $\lambda_{ik}(t)$ is the hazard function for individual *i* for event *k* (NSF fee occurrence) at time *t*. As mentioned above, we do not assume a base distribution of the hazard function λ. β is a vector that includes, as in models (3) and (4), the following: the treatment dummy variables, $Base$ for model (3) where the underlying alternative is no message (Control condition) and $Simplified$ for model (4) where the underlying alternative is the original Base message; $First$ is a dummy for first-timers; and $Base\_First$/$Simplified\_First$ are the interaction variables between the two. In both cases, *Yearly income* represents the participant's yearly income. The standard errors are robust.

In our specification, the time periods are set as days. We measure time to event *k* from time *t*, where *t* stands for the time of the previous event or of the participant's entry to the experiment. Insofar as the unit of time is set as 1 day, the TTE analysis treats all NSF or overdraft fees paid during a single day as one event. The results are displayed in Table 6.

[Table 5, 6]

The Base version significantly reduced the number of NSF fees paid during all the time periods investigated (one week, two weeks, one month, and four months), for both TAF and TT estimations. Simplifying the message significantly increased its impact compared to the Base version for all time periods, in both TAF and TT estimations. These results likewise emerged as significant when controlling for first-time effects and income.[42] Within the TTE analysis, we find that the Base alternative reduced the number of NSF fees, albeit not significantly; the reduction achieved by the simplified version, on the other hand, emerged as significant.

[42] The effect of yearly income was significantly positive on both the open rate and NSF fees in models (1), (3), and (4), which means that people with a higher income opened reminders at a greater rate and paid significantly more NSF/overdraft fees. These results were found to be significant in estimations (6) and (8) as well. *First*—the first-timer dummy variable—was found to be negatively significant for model (1) on the open rate and for models (3) and (4) on NSF fees. Participants who received the reminder for the first time opened it less on average and paid significantly fewer NSF fees. These results were obtained in all the other estimations, as evident in Tables 7–9, 12–14.

According to our hypothesis, as well as the previous literature, sending a reminder makes a difference, in and of itself, at least in the short term; specifically, it significantly helps address inattention, which, in the financial context, may result in overdraft fees. In the experiment, receiving reminders increased individuals’ attention to their financial situations and reduced the number of NSF fees. Simplifying the message substantially enhanced this impact, changing participants' behavior by further reducing the cost of acquiring salient financial information.

Yet, while the number of NSF fees was reduced, the treatment did not eliminate them by a long margin. This can be accounted for with the rational model, that is, the estimation of the importance of the purchase relative to the fee anticipated, additional cost barriers, or ostrich behavior.

7.2. Comparison of the different framings: Positive (Save) and negative (Avoid)

Next, we explore possible differences in the impact of differently framed reminders by changing the heading of the simplified message from Avoid (negative) to Save (positive). As in the previous analyses, we begin by gauging the differences in the open rates and click-to-explore rates as proxies for interest and engagement triggered by the reminders and then examine the differences in the number of NSF and overdraft fees paid per person. Descriptive statistics for both these examinations are presented in Tables 1 and 2.

On comparing the simplified alternatives bearing the Save and Avoid headings, no significant difference in the open rates emerged (30.51% and 30.85%, respectively). In terms of click-to-explore rates, we observe that the Avoid version increased engagement significantly compared to the Save version (2.5% and 1.51% for Avoid and Save, respectively). Significant differences between the two groups were observed in reducing NSF and overdraft fees, at least in the long term, that is, at least one month. Compared to the Control group, the mean reduction for the Simplified Avoid group stood at 9.25% as measured after one week and 8.9% as measured for the full experiment (four months), respectively; the parallel statistics obtained for the Simplified Save group were 9.44% and 4.9%, respectively. In dollar terms and for the experiment duration, receiving the Simplified Avoid reminder saved, on average, $11.33 more per user than receiving no message (reduction of $25.16 vs. $13.85 for the Simplified Avoid and Simplified Save group, respectively, compared to the Control group). The Simplified Save version appears to be subject to a stronger devaluation in the long term, that is, at least one month.

For the regression analysis of the open rate and click-to-explore rate, we use the same probit models as in models (1) and (2)[43]:

$$(6)\quad Open_i = \alpha + \beta_1 Simplified\ Avoid_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$$

$$(7)\quad Click\ to\ explore_i = \alpha + \beta_1 Simplified\ Avoid_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$$

where the outcome variables $Open_i$ in model (6) and $Click\ to\ explore_i$ in model (7) take the value of either 0 or 1 for at least one reminder during the period of investigation for each individual, and $Simplified\ Avoid$ is a dummy variable for the simplified treatment where the underlying alternative is Simplified Save . $First$ is a dummy for first-timers and $Simplified\_First$ is the interaction variable between the two. In both models, $\log(Yearly\ income)$ represents the individual's log yearly income. The results of the analysis are displayed in Table 3.

In terms of the positive versus negative framing of the simplified message, the regression results indicate no significant differences in the open rates and a significant increase in the click-to-explore rates for the Simplified Avoid compared to Simplified Save message. For the NSF fees regression, we follow the same three approaches as described above. Model (8) represents the TAF and TT analyses and the cross-section for different time frames. We use a tobit analysis on the log(1+ NSF period) as the dependent variable (the values are described below).[44] Model (9) represents the TTE analysis for both NSF fees and time as dependent variables. Betas are as in model (8).

[43] The results are not sensitive to different regression models such as OLS or logit. The analysis is not sensitive to whether we use open and click-to-explore rates as acting on at least one of the several reminders that were sent for each participant divided by the total group participants, as in the model, or as a rate for every single reminder.

[44] The results are not sensitive to different regression models such as OLS.

(8) $\log(1 + NSF_{Period,i}) = \alpha + \beta_1 Simplified\ Avoid_{\ i} + \beta_2 First_i + \beta_3 Simplified\_First_i + \beta_4 \log(Yearly\ income)_i + \varepsilon_i$

(9) $\lambda_{ik}(t) = \lambda_{0k}(t) * e^{X_{ik}\beta}$

In model (8), the time periods are one week, two weeks, and one month after receiving the first reminders and the four months of the entire experiment. In model (9) the dependent variable is the hazard for NSF fees that encompasses both NSF fee occurrences and the time to such events. For both models, $Simplified\ Avoid$ is a dummy variable for the Simplified Avoid message (where the underlying alternative is the Simplified Save reminder). $First$ is a dummy for first-timers and $Simplified\_First$ are the interaction variables between the two. In both models, $\log(Yearly\ income)$ represents the individual's log yearly income. The results for the TAF and TT analyses are displayed in Table 7 and for the TTE in Table 8.

[Table 7, Table 8]

On comparing the positive versus negative framing of the simplified message, we find significant difference for the Avoid version after one month, for the total experiment (four months), and when changing the risk in the TTE analysis. In both analyses, these significant impacts of the Avoid message appear to last longer in the course of the experiment.

Overall, the simplified message with the Avoid heading engendered a significantly higher interest in exploring the account and emerges as more effective in reducing the number of NSF fees compared to the positive Save framing. These results correspond to the goal-framing literature and add to Karlan and colleagues' (2016) findings in the context of savings.

7.3. Limited attention and ostrich effects: Identification across different income levels

In addition to the main analysis, we explore the effect of the reminders for different levels of annual income. Our objective is to explore whether individuals with a lower yearly income tend to display more ostrich behavioral and whether the framing of the message mitigates such behavior. To this end, we first compare the impact of receiving a reminder (any type) for each of the income groups studied relative to the Control group. Next, we attempt to identify ostrich behavior by analyzing the different influence of the negative versus positive framing across income groups; specifically, we test whether individuals in different income groups opened the Save reminder at a greater rate than the Avoid reminder.

We divide the sample into six income percentile groups:[45] (1) total annual income of less than $10,000; (2) up to the 25th percentile, $10,000–$20,000; (3) up to the 50th percentile, $20,000–$40,000; (4) up to the 75th percentile, $40,000–$70,000; (5) up to the 90th percentile, $70,000–$115,000; and (6) an annual income of $115,000 and above. Figure 4 shows the distribution of the average NSF fees one week after receiving the first reminder for the Treatment group - receiving any type of reminder, compared to the Control group with respect to the six income groups.

[Figure 4]

Within the experimental sample—which, as already noted, comprises individuals with a higher probability of overdrawing an account than the general population—we observe that participants with a higher total annual income pay more NSF fees. In the lowest income level

[45] Based on the U.S. Census Bureau's Current Population Survey (Census CPS) Annual Social and Economic Supplement; data are similar to that of the IPUMS-CPS, University of Minnesota analysis (2019). We use the four personal income quartiles and split the lower and the upper quartiles with approximate bottom and top deciles to explore potential notable changes at the tails of the distribution. We excluded 13 participants from this analysis because we do not have their income information. In addition, we censored the sample from outliers of yearly incomes above $1 million to have the same income distribution across treatments. Notwithstanding that the full results remain the same when using, for robustness, the entire sample as well. As one might argue that it is debatable if the relevant grouping from the distribution should be based on individual income or household income, we also provide in Section 8 household-level income cuts with a similar analysis. The quality of our results remains.

(less than $10,000), the mean total number of NSF fees paid stands at 4.5 and 4.77 for the control and the treatment populations, respectively. As income rises, this number steadily and significantly increases, culminating in the respective means of 10.46 and 9.43 for the highest income group.

Concerning the mean open rate, we observe a positive trend, such that the open rate for the lowest income group is 23.17%, rising to 31.17% for the fifth income group. In the highest income group, the open rate stands at 30.65%.

Differences in the number of NSF fees paid between the Control group and the treatment population (i.e., those who received any version of the reminder) with respect to income level are not significant in the first three income groups: a difference of 5.96%, 5.49%, and -2.9%, for the first, second, and third group, respectively. For the fourth income group between the 50th and 75th percentiles, we find a significant mean reduction of 6.98%. For the fifth income group, with yearly income between $70,000 and $115,000 (75th–90th percentiles), the mean reduction in the number of overdraft fees paid, compared to the Control group, is at -8.23%. This pattern endures in the sixth income group, with yearly income of $115,000 and above, where the mean reduction, compared to the Control group, is significant at -9.84%. Overall, the influence of the reminders increases with yearly income. Table 9 displays the descriptive statistics for this analysis.

[Table 9]

On comparing the differences in the effect of the Avoid versus Save versions on the different income groups relative to the control population, we observe the following: First, only the second income group registered significant differences in open rates, with higher open rates found for the Simplified Avoid version. The differences in open rates for all other income groups emerged as non-significant. Thus, the trend of increasing open rates holds for both simplified versions as well. Table 10 presents the descriptive statistics for this analysis.

[Table 10]

Second, no consistent differences are observed between the Avoid and Save versions in terms of the reduction in the number of overdraft fees paid. Specifically, no significant differences emerged in the first, lowest income group. Within the second income group, a significant 11.1% difference in reducing NSFs is observed, with the positive framing being more effective (mean number of overdraft fees of 5.5 and 4.95 for the Simplified Avoid and Simplified Save version, respectively). From the third income group onward, the opposite effect is perceptible, with the Simplified Avoid version performing better in reducing the number of overdraft fees compared to the Simplified Save version, although the differences appear significant only for the third and the sixth groups (differences in reducing the number of NSF fees, compared to the Simplified Save group, of 7.1%, 3.2%, 3.4%, and 7.4% for the third, fourth, fifth, and sixth group, respectively).

For the treatment versus no treatment regressions, we use tobit log models on the total number of NSF fees paid for the duration of the experiment and for the weekly number of NSF fees paid as of the first reminder.

$$(10) \log(1 + Period\ NSF_{Income\ group,i}) = \alpha + \beta_1 Reminder_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \varepsilon_i$$

where the outcome variables are one week and the total number of NSF fees for an individual i in each of the six income groups. *Reminder* is a dummy variable if the participant received any kind of reminder. $First$ is a dummy for first-timers and $Simplified\_First$ is the interaction variable between the two. Table 11 shows the results of the analysis.

[Table 11]

To compare the effect of the Avoid versus Save versions on the different income levels, we first consider the open rates and then analyze the number of overdraft fees paid. For the regression analysis on the open rates, we use the following probit model:[46]

$$(11) Open_{income,i} = \alpha + \beta_1 Simplified\ Avoid_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \varepsilon_i$$

where the outcome variable $Open_{income,i}$ equals either 0 or 1 for each reminder received by individual i, in each of the six income groups; $Simplified\ Avoid_i$ is a dummy variable for the simplified treatment, with Save as the underlying alternative; $First$ is a dummy for first-timers; and $Simplified\_First_i$ is the interaction variable between the two. Table 12 shows the results of the analysis.

[Table 12]

For the NSF fees regression, we use a tobit analysis on the log(1+ Period NSFs) as the dependent variable.[47]

$$(12) \log(1 + Total\ NSF_{income,i}) = \alpha + \beta_1 Simplified\ Avoid_i + \beta_2 First_i + \beta_3 Simplified\_First_i + \varepsilon_i$$

where the outcome variables are log + 1 for one week and the total number of NSF fees for individual i for the different income groups. The variable $Simplified\ Avoid$ is a dummy variable for the Simplified Avoid version (with Simplified Save as the underlying alternative).

[46] The results are not sensitive to different regression models such as OLS or logit.
[47] The results are not sensitive to the NSF fee as a dependent variable with tobit and OLS regression.

$First$ is a dummy for first-timers and $Simplified\_First$ is the interaction variable between the two. Table 13 shows the results.

[Table 13]

The income analysis reveals that sending reminders to individuals in the 25th income percentile and above significantly reduced their number of overdraft fees at least in the short term,[48] and the same effect obtains for the total experiment duration among individuals in the 75th percentile and above. Open rates observed as increasing with income for all treatments and for the Avoid version were significantly higher in the 10th–25th income percentiles compared to the Simplified Save alternative. Differences in the reduction in number of NSF fees observed between the Avoid and Save versions were not consistent, and no significant differences were found over the short term of one week. During the course of the entire experiment, the following patterns were observed: In the 10th–25th percentiles, the Save alternative outperformed the Avoid alternative with marginal significance; and in the 25th–50th percentiles a reverse outcome was obtained, such that the Avoid alternative significantly outperformed the Save alternative. This trend endures into the highest income group but becomes significant again only in the highest decile.

We consider these results as evidence that technology is able to influence the limited-attention pattern. For the Mint users who received the messages, we reduced the cost of acquiring information and consequently the number of overdraft fees. Participants who emerged as the most strongly influenced by the reminders were those with incomes above the 25th percentile in the short term and above the 75th percentile for the full course of the experiment (four months). These results differ from Stango and Zinman's (2014) findings showing a greater influence of salience shocks on overdraft fees for lower income individuals.

[48] Similar results were observed two weeks and one month after receiving the first reminder.

This connection raises the question of why all income groups were not influenced by the reminders in the same way.

Among the explanations that can be suggested for the lack of uniformity observed, four appear to be the most salient. (1) Our results are consistent with the ostrich behavior pattern. We find that the lower income groups tend to act as ostriches, in that they received the messages but were unwilling to face information that shed light on their precarious financial situation and therefore did not change their behavior. They opened the reminders at a lower rate, and as a consequence, the number of NSF fees they paid was reduced to the least extent. (2) Low-income individuals, whose marginal rate substitution should be lower, and who could be subject to the ostrich effect, may in addition choose to make a transaction resulting in an overdraft because they have no better credit alternative. (3) Low-income individuals are subject to a cognitive load that may impact their ability to allocate resources to dealing with the information they receive or it can influence their e-mail use habits. (4) For the individuals who pay fewer NSF fees, we may not have reduced information costs sufficiently to elicit a behavior change.

The results of the test regarding the effectiveness of a positive (Save) versus negative (Avoid) framing were rather inconsistent. Our results do not provide evidence that the Save version helped mitigate the ostrich behavior in dealing with overdrafts among the low-income group.

8. Robustness

8.1. Propensity score matching analysis

One might argue that although the sample population was randomly distributed across the treatments, groups may still have differed in various characteristics. For example, first-timers may have been categorically different from return ODP users, who had received reminders prior to the experiment. To address such putative differences and confirm they do not impact the main results, we conduct a propensity score matching analysis, which matches between the treatment groups on the number of NSF fees paid in the various time periods, as well as between the first-time and return users with parallel yearly incomes.[49]

We find that the results are consistent with our main results: Sending a simplified reminder significantly reduces the number of overdraft fees compared to the Base treatment, and the significant differences between the Avoid and the Save framings hold as well. These results support the randomization and suggest that the first-timers versus return users allocation did not change the results across the treatment groups.

8.2. Endogeneity of treatments

One might argue that the effect of the different versions of the reminder is contingent on the number of messages sent to each participant, as decided by the algorithm, rather than on their structure or framing. It is noteworthy, in this connection, that the group means (and standard deviations) for the number of reminders sent during the experiment were as follows: 3.25 (*SD* = 2.75), 3.15 (*SD* = 2.68), and 3.16 (*SD* = 2.71) for the Base, Simplified Avoid, and Simplified Save group, respectively. In our view, these statistics reinforce our results, since the smallest number of reminders were in the most effective, that is, the Simplified Avoid group.

8.3. Household income groups split

As mentioned above, we could not determine whether the users represent an individual or household. Given that in the Results section we analyze the results on the basis of individual income cuts, for robustness we also analyze the results using household income cuts[50] to see if the effects on the groups remain. It is not surprising that we find the same trends. Yet, now the thresholds are higher at every income level. For the immediate response of one week after the first reminder, we were able to reduce the number of NSF fees significantly above the 10th

[49] We use logit for the confounders and nearest neighbors for the propensity matching.
[50] Our analyses we based on Census CPS 2018 data. The cuts were as follows: $0–14,600, $14,600–30,700, $30,700–63,000, $63,000–114,000, $114,000–$185,000, and $185,000 and above.

income percentile. For the total experiment duration, we were able to significantly reduce the number of NSF fees paid for the 50th percentile and above.

8.4. Participants with a large number of NSFs were those impacted by the messages

As described above, participants with a higher total annual income pay more NSF fees. One might claim that the differences observed in the influence of reminders on different income groups depend on the proportion of participants with a prior tendency to pay NSF fees. In other words, the differences observed across the income groups in the impact of the reminders are not due to a particular kind of treatment but occur because income groups differ as to their members' tendency to pay NSF fees in the first place. To address this challenge, we censor our data to create closer NSF fee distributions among the income groups and we limit the number of weekly fees to seven and five, consecutively. Under both limits, the distribution of NSF fees across the different quantiles is practically the same. The results appear to hold with both censored data analyses: The influence of the reminders is significantly stronger on participants with incomes above the 25th percentile.

9. Conclusions

This study has investigated whether overdraft fees paid by individuals can be lowered using an AI algorithm that creates reminders. To this end, we conducted a field experiment involving 39,607 unique users of Mint, a large web-based personal financial management application in the United States and Canada.

Our goal was to test the possibility of reducing limited-attention behavioral patterns that affect individuals' financial situations. We conducted two manipulations: simplification and negative versus positive framing. Our findings show that sending reminders is effective in mitigating limited attention, and that simplifying the message can substantially increase this impact. The results also reveal that the negative framing performed better than the positive framing in reducing overdraft fees.

The study contributes to the literature concerned with models of selective attention and limited attention. Participants in our experiment who emerged as the most susceptible to the reminders were in the medium to high income levels. For this population, reminders—especially when phrased in simple terms—lowered the cost of gathering salient information sufficiently to trigger behavior change. Our findings show, however, that low-income people tended to open the reminders at a lower rate, and hence the messages sent were not sufficient to trigger action—a finding that we regard as an indication of a greater incidence of ostrich behavior patterns among such individuals.

A promising avenue for future research would be to test whether the impact of reminders could be enhanced through personalized messages based on individuals' personal information and whether such reminders would prove more beneficial for their financial behavior than a standardized version. We likewise leave for future research to explore the results obtained in our experiment over a longer time period. Furthermore, studies could investigate the effects of reminders on individuals' consumption behaviors and whether sending notices would change periodic purchase patterns or the use of financial resources.

References

Amorim, L. D., & Cai, J. (2015). Modelling recurrent events: A tutorial for analysis in epidemiology. *International Journal of Epidemiology*, *44*(1), 324–333. https://doi.org/10.1093/ije/dyu222

Andries, M., & Haddad, V. (2017). *Information aversion*. National Bureau of Economic Research.

Ashraf, N., Karlan, D. S., & Yin, W. (2006). Household Decision Making and Savings Impacts: Further Evidence from a Commitment Savings Product in the Philippines. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.912771

Baker, S. R. (2018). Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data. *Journal of Political Economy*, *126*(4), 1504–1557. https://doi.org/10.1086/698106

Benartzi, S., & Lehrer, J. (2015). *The Smarter Screen: What Your Business Can Learn from the Way Consumers Think Online*. Hachette UK.

Block, L. G., & Keller, P. A. (1995). When to accentuate the negative: The effects of perceived efficacy and message framing on intentions to perform a health-related behavior. *Journal of Marketing Research*, *32*(2), 192–203.

Bossaerts Peter, Yadav Nitin, & Murawski Carsten. (2019). Uncertainty and computational complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1766), 20180138. https://doi.org/10.1098/rstb.2018.0138

Caflisch, A., Grubb, M. D., Kelly, D., Nieboer, J., & Osborne, M. (2018). Sending Out an SMS: The Impact of Automatically Enrolling Consumers Into Overdraft Alerts. *Available at SSRN 3538527*.

CFPB Report Raises Concerns About Impact of Overdraft Practices on Consumers. (2017). Consumer Financial Protection Bureau. Retrieved May 4, 2019, from https://www.consumerfinance.gov/about-us/newsroom/cfpb-report-raises-concerns-about-impact-of-overdraft-practices-on-consumers/

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review*, *99*(4), 1145–1177. https://doi.org/10.1257/aer.99.4.1145

Cowan, N., Johnson, T. D., & Scott Saults, J. (2005). Capacity limits in list item recognition: Evidence from proactive interference. *Memory*, *13*(3–4), 293–299.

Davis, M. A., & Bobko, P. (1986). Contextual effects on escalation processes in public sector decision making. *Organizational Behavior and Human Decision Processes*, *37*(1), 121–138. https://doi.org/10.1016/0749-5978(86)90048-8

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, *47*(2), 315–372. https://doi.org/10.1257/jel.47.2.315

Franco, J. P., Yadav, N., Bossaerts, P., & Murawski, C. (2018). *Where the really hard choices are: A general framework to quantify decision difficulty* [Preprint]. Animal Behavior and Cognition. https://doi.org/10.1101/405449

Gabaix, X. (2017). *Behavioral inattention*. National Bureau of Economic Research.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review*, *96*(4), 1043–1068. https://doi.org/10.1257/aer.96.4.1043

Galai, D., & Sade, O. (2006). The “ostrich effect” and the relationship between the liquidity and the yields of financial assets. *The Journal of Business*, *79*(5), 2741–2759.

Gathergood, J., Mahoney, N., Stewart, N., & Weber, J. (2019). How Do Individuals Repay Their Debt? The Balance-Matching Heuristic. *American Economic Review*, *109*(3), 844–875. https://doi.org/10.1257/aer.20180288

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making*, *26*(3), 213–224. https://doi.org/10.1002/bdm.1753

Hirshleifer, D., & Teoh, S. H. (2003). *Limited Attention, Information Disclosure, and Financial Reporting*. 63.

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science, 62*(12), 3393–3411.

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty*, *38*(2), 95–115.

Kőszegi, B., & Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review*, *99*(3), 909–936.

Levin, I. P., & Gaeth, G. J. (1988). How Consumers are Affected by the Framing of Attribute Information Before and After Consuming the Product. *Journal of Consumer Research*, *15*(3), 374. https://doi.org/10.1086/209174

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational Behavior and Human Decision Processes*, *76*(2), 149–188. https://doi.org/10.1006/obhd.1998.2804

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology*, *25*(2), 115–126. https://doi.org/10.1080/01449290500330448

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science*, *341*(6149), 976–980.

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

Melzer, B. T., & Morgan, D. P. (2015). Competition in a consumer loan market: Payday loans and overdraft credit. *Journal of Financial Intermediation*, *24*(1), 25–44.

Meyerowitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, *52*(3), 500.

Nieuwerburgh, S. V., & Veldkamp, L. (2009). *Information Immobility and the Home Bias Puzzle*. 45.

Olafsson, A., & Pagel, M. (2017). *The Ostrich in Us: Selective Attention to Financial Accounts, Income, Spending, and Liquidity* (Working Paper No. 23945). National Bureau of Economic Research. https://doi.org/10.3386/w23945

Olafsson, A., & Pagel, M. (2018). The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software. *The Review of Financial Studies*, *31*(11), 4398–4446. https://doi.org/10.1093/rfs/hhy055

Prentice, R. L., Williams, B. J., & Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika*, *68*(2), 373–379.

Reinecke, K., Yeh, T., Miratrix, L., Mardiko, R., Zhao, Y., Liu, J., & Gajos, K. Z. (2013). Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2049–2058.

Services and Products. (2017). Retrieved April 28, 2019, from http://www.moebs.com/services/nobounce

Sims, C. A. (2010). Rational Inattention and Monetary Economics. In *Handbook of Monetary Economics* (Vol. 3, pp. 155–181). Elsevier. https://doi.org/10.1016/B978-0-444-53238-1.00004-1

Stango, V., & Zinman, J. (2014). Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. *The Review of Financial Studies*, *27*(4), 990–1030.

Tuch, A. N., Bargas-Avila, J. A., & Opwis, K. (2010). Symmetry and aesthetics in website design: It'sa man's business. *Computers in Human Behavior*, *26*(6), 1831–1837.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

United States Household Income Brackets and Percentiles in 2018—DQYDJ. (2018). Retrieved April 28, 2019, from https://dqydj.com/united-states-household-income-brackets-percentiles/

Zheng, X. S., Chakraborty, I., Lin, J. J.-W., & Rauschenberger, R. (2009). Correlating low-level image statistics with users-rapid aesthetic and affective judgments of web pages. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–10.

Table 1. Descriptive Statistics: Number of Non-Sufficient Funds (NSF) Fees

Condition	Observations	Number of NSF fees assessed											
		One week after 1st reminder			Two weeks after 1st reminder			One month after 1st reminder			End of experiment (four months)		
		M	*SD*	*Mdn*	*M*	*SD*	*Mdn*	*M*	*SD*	*Mn*	*M*	*SD*	*Mdn*
Control	10,972	0.91	1.65	0.00	1.56	2.53	1.00	2.91	4.22	2.00	8.08	11.41	4.00
Base	9,422	0.86	1.58	0.00	1.48	2.45	1.00	2.79	4.27	1.00	7.82	11.12	4.00
Simplified Avoid	9,725	0.83	1.54	0.00	1.41	2.30	0.00	2.63	3.79	1.00	7.36	10.45	4.00
Simplified Save	9,488	0.83	1.51	0.00	1.43	2.36	0.00	2.70	4.04	1.00	7.68	11.67	4.00

Notes: The table shows the observations, means, standard deviations, and medians for the number of NSF fees assessed one week, two weeks, one month, and four months after the first reminder was sent.

Table 2. Descriptive Statistics: Open and Click-to-Explore Rates and Yearly Income

Condition	Observations	Open rate		Click-to-explore rate		Yearly income (U.S. dollars)		
		M	*SD*	*M*	*SD*	*M*	*SD*	*Mdn*
Control	10,972	—	—	—	—	111,094.58	1,506,853.37	53,059.49
Base	9,422	28%	45%	1.4%	15.1%	110,067.38	634,350.13	52,545.41
Simplified Avoid	9,725	31%	46%	2.5%	19.5%	105,212.89	532,774.45	53,854.11
Simplified Save	9,488	31%	46%	1.5%	14.0%	99,740.31	573,513.53	53,387.42

Notes: The table shows the observations, means, and standard deviations for the reminders' open rates (any reminder within the experiment), click-to-explore rates (explore the app from any reminder within the experiment) and participants' yearly income.

Table 3. Summary Statistics: Treatment Effects

Condition	NSF fees (one week after reminder)	NSF fees (four months)	Open rate	Click-to-explore rate
	(1)	(2)	(3)	(4)
Base	-5.3% ***	-3.3%*		
Simplified Avoid	-9.25% ***	-8.9%***	12.1%***	78.4%***
Simplified Save	-9.44% ***	-4.9%***	11%***	7.8%
Control	**0.91**	**8.08**		
Base			**27.5%**	**1.4%**

Notes: Columns 1 and 2 show average treatment change compared to the Control group in number of non-sufficient funds (NSF) fees one week after receiving the first reminder and for four months experiment. Columns 3 and 4 show the average open rates and click-to-explore rates for the simplified alternatives (Avoid, Save) compared to the Base (original) message. Open rates and click-to-explore rates calculated on any reminder in the experiment. The Control and Base values below the line are means. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. Probit Analysis for the Open and Click-to-Explore Rates

Variable	Open rate		Click-to-explore rate	
	Base vs. Simplified	**Simplified Avoid vs. Simplified Save**	**Base vs. Simplified**	**Simplified Avoid vs. Simplified Save**
Simplified Avoid	0.0291***	-0.00896	0.00799***	0.00574**
	(0.00859)	(0.00874)	(0.00245)	(0.00246)
Yearly income	0.0116***	0.0123***	-0.00118	-0.00119
	(0.00349)	(0.00359)	(0.000942)	(0.000998)
First-timers	-0.0361***	-0.0545***	-0.00182	-0.00341
	(0.00960)	(0.00958)	(0.00295)	(0.00294)
Simplified Avoid × First-timers	0.0105	0.0284**	0.00211	0.00371
	(0.0133)	(0.0135)	(0.00377)	(0.00381)
N	19,147	19,213	19,147	19,213

Notes: The table shows probit regressions and marginal effects; delta-method standard errors are in parentheses. Dependent variables are open rates (any reminder within the experiment) and click-to-explore rates (explore the app from any reminder within the experiment). First-timers is a dummy variable for new Overdraft Prediction users. Yearly income is the log(yearly income). Simplified Avoid is the condition where the underlying variable is the Base or the Simplified Save condition. Simplified Avoid × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5. Tobit Regression for Log(1+ NSF Period) for the Mean Number of NSF Fees One Week, Two Weeks, and One Month After First Reminder, and for Four Months of the Experiment: Control Versus Base and Base Versus Simplified Avoid Group

Variable	**Base vs. Control group**				**Simplified Avoid vs. Base group**			
	Full Experiment (four months)	**One month**	**Two weeks**	**One week**	**Full Experiment (four months)**	**One month**	**Two weeks**	**One week**
Base	-0.0316*	-0.0386**	-0.0327**	-0.0283**				
	(0.0192)	(0.0159)	(0.0137)	(0.0116)				
First-timers	-0.807***	-0.533***	-0.393***	-0.271***	-0.761***	-0.500***	-0.367***	-0.246***
	(0.0196)	(0.0154)	(0.0127)	(0.0106)	(0.0209)	(0.0165)	(0.0136)	(0.0113)
Base × First-timers	0.0470	0.0330	0.0265	0.0248				
	(0.0286)	(0.0226)	(0.0187)	(0.0155)				
Yearly income	0.0952***	0.0594***	0.0536***	0.0417***	0.0794***	0.0534***	0.0499***	0.0384***
	(0.00787)	(0.00632)	(0.00536)	(0.00450)	(0.00827)	(0.00666)	(0.00570)	(0.00476)
Simplified Avoid					-0.0530***	-0.0399**	-0.0345**	-0.0199*
					(0.0200)	(0.0165)	(0.0141)	(0.0119)
Simplified Avoid × First - timers					-0.000569	0.0129	0.0282	0.0126
					(0.0294)	(0.0232)	(0.0191)	(0.0159)
N	20,394	20,394	20,394	20,394	19,147	19,147	19,147	19,147

Notes: The table shows tobit regressions and marginal effects; delta-method standard errors are in parentheses. Dependent variables are non-sufficient funds (NSF) fees assessed one week, two weeks, and one month after first reminder sent, and four months of the experiment. Base is the condition where the underlying variable is the Control condition. First-timers is a dummy variable for new Overdraft Prediction users. Base × First-timers is an interaction variable. Yearly income is the log(yearly income). Simplified Avoid is the condition where the underlying variable is the Base condition. Simplified Avoid × First-timers is an interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6. Time-to-Event (TTE) Analysis and Semi-Parametric Cox Model: Base Versus Control group and Simplified Avoid Versus Control group

Variable	Base vs. Control group TTE-PWP-Cox	Simplified Avoid vs. Control group TTE-PWP-Cox
Base	-0.00875	
	(0.00825)	
Yearly income	1.66e-09	-2.58e-10
	(3.98e-09)	(3.48e-09)
First-timers	-0.208***	-0.206***
	(0.0131)	(0.0131)
Base × First-timers	0.00559	
	(0.0185)	
Simplified Avoid		-0.0183**
		(0.00824)
Simplified Avoid × First-timers		-0.00858
		(0.0187)
N	114,604	113,546

Notes: The table shows results of the TTE analysis and PWP(Prentice-Williams-Peterson)-Cox semi-parametric regression; robust standard errors are in parentheses. Dependent variables are both non-sufficient funds (NSF) fees and time to NSF event (the hazard for NSF). Yearly income is the log(yearly income), First-timers is a dummy variable for new Mint Overdraft Prediction users. Base versus Control: the underlying variable is the Control condition. Simplified Avoid is the condition where the underlying variable is the Control condition. Simplified Avoid × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7. Tobit Regression for Log(1+ NSF Period) for the Mean Number of NSF Fees One Week, Two Weeks, and One Month After First Reminder, and Four Months of the Experiment: Simplified Avoid Versus Simplified Save Group

Variable	Time frame			
	Full Experiment (four months)	**One month**	**Two weeks**	**One week**
Simplified Avoid	-0.0502**	-0.0298*	-0.0160	-0.00298
	(0.0201)	(0.0166)	(0.0142)	(0.0119)
First-timers	-0.789***	-0.516***	-0.361***	-0.233***
	(0.0209)	(0.0164)	(0.0136)	(0.0112)
Yearly income	0.0273	0.0281	0.0219	-0.000920
	(0.0293)	(0.0231)	(0.0191)	(0.0158)
Simplified Avoid × First-timers	0.0754***	0.0419***	0.0359***	0.0271***
	(0.00836)	(0.00678)	(0.00568)	(0.00470)
N	19,213	19,213	19,213	19,213

Notes: The table shows tobit regressions and marginal effects; delta-method standard errors are in parentheses. Dependent variables are number of non-sufficient funds (NSF) fees one week, two weeks, and one month after first reminder sent, and four months of the experiment. First-timers is a dummy variable for new Mint Overdraft Prediction users. Yearly income is the log(yearly income). Simplified Avoid is the condition where the underlying variable is the Simplified Save condition. Simplified Avoid × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8. Time-to-Event (TTE) Analysis and Semi-Parametric Cox Model: Simplified Avoid Versus Simplified Save Group

Variable	TTE-PWP-Cox
Simplified Avoid	-0.0192**
	(0.00864)
Yearly income	6.50e-09
	(7.35e-09)
First-timers	-0.210***
	(0.0144)
Simplified Avoid × First-timers	0.00169
	(0.0195)
N	102,724

Notes: The table shows results of the TTE analysis and PWP(Prentice-Williams-Peterson)-Cox semi-parametric regression; robust standard errors are in parentheses. Dependent variables are both non-sufficient funds (NSF) fees and time to NSF event (the hazard for NSF). First-timers is a dummy variable for new Mint Overdraft Prediction users. Simplified Avoid is the condition where the underlying variable is the Simplified Save condition. Yearly income is the log (yearly income). Simplified Avoid × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9. Mean Number of NSF Fees: Treatment Versus No Treatment (Control) by Income Group

Condition	Income group					
	Up to $10,000	$10,000–20,000 (up to 25th percentile)	$20,000–40,000 (up to 50th percentile)	$40,000–70,000 (up to 75th percentile)	$70,000–115,000 (up to 90th percentile)	$115,000 and above (above 90th percentile)
Control	4.51	4.96	6.73	8.17	9.11	10.46
Treatment	4.78	5.23	6.53	7.60	8.36	9.43
Difference	5.96%	5.49%	-2.91%	-6.98%	-8.23%	-9.84%

Notes: The table displays the observations for the mean number of non-sufficient funds (NSF) fees in the six income groups. The treatment population includes all the versions of the reminders (see Figure 2 for the description of the different treatments).

Table 10. Open rates by Income Group

Condition	Income group					
	Up to $10,000	$10,000–20,000 (up to 25th percentile)	$20,000–40,000 (up to 50th percentile)	$40,000–70,000 (up to 75th percentile)	$70,000–115,000 (up to 90th percentile)	$115,000 and above (above 90th percentile)
Simplified Avoid	0.24	0.30	0.30	0.31	0.33	0.31
Simplified Save	0.26	0.25	0.29	0.31	0.33	0.32
Difference Avoid–Save	-6%	20%	1%	0%	0%	-4%

Notes: The table displays the observations for the mean open rates (opened any reminder sent) in the six income groups. See Figure 2 for the description of the different treatments.

Table 11. Tobit Regressions for Number of NSF Fees One Week After the First Reminder and Four Months of the Experiment: Reminder Versus No Reminder, by Income Group

Variable	four months of experiment						one week after reminder					
	Up to $10,000	$10,000–20,000 (up to 25th percentile)	$20,000–40,000 (up to 50th percentile)	$40,000–70,000 (up to 75th percentile)	$70,000–115,000 (up to 90th percentile)	$115,000 and above (above 90th percentile)	Up to $10,000	$10,000–20,000 (up to 25th percentile)	$20,000–40,000 (up to 50th percentile)	$40,000–70,000 (up to 75th percentile)	$70,000–115,000 (up to 90th percentile)	$115,000 and above (above 90th percentile)
Reminder	-0.0925	0.0189	-0.0389	-0.0198	-0.0678*	-0.0843**	-0.0611	-0.00823	-0.0388**	-0.0343**	-0.0510**	-0.0450*
	(0.105)	(0.0524)	(0.0278)	(0.0268)	(0.0350)	(0.0392)	(0.0587)	(0.0309)	(0.0162)	(0.0164)	(0.0211)	(0.0245)
First-timers	-0.745***	-0.508***	-0.674***	-0.771***	-0.900***	-1.109***	-0.185***	-0.164***	-0.243***	-0.250***	-0.312***	-0.360***
	(0.115)	(0.0634)	(0.0361)	(0.0367)	(0.0470)	(0.0524)	(0.0629)	(0.0348)	(0.0193)	(0.0203)	(0.0247)	(0.0288)
Reminder × First-timers	0.134	-0.00725	0.0654	-0.0101	0.0443	0.0747	0.0635	-0.00524	0.0524**	0.00498	0.0550*	0.0358
	(0.132)	(0.0732)	(0.0412)	(0.0414)	(0.0532)	(0.0593)	(0.0715)	(0.0397)	(0.0220)	(0.0228)	(0.0281)	(0.0325)
N	1,030	3,436	1,2586	1,4198	9,092	8,086	1,030	3,436	12,586	14,198	9,092	8,086

Notes: The table shows the tobit regression analysis of number of non-sufficient funds (NSF) fees one week after the first reminder and four months of experiment for the six income groups. Marginal effects are presented and delta-method standard errors are in parentheses. Reminder refers to any kind of reminder and the underlying variable is the Control condition (no reminders). First-timers is a dummy variable for new Overdraft Prediction users. Reminder × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12. Probit Regression Analysis of Open Rates: Simplified Avoid Versus Simplified Save , by Income Group

Variable	Income group					
	Up to \$10,000	**\$10,000–20,000 (up to 25th percentile)**	**\$20,000–40,000 (up to 50th percentile)**	**\$40,000–70,000 (up to 75th percentile)**	**\$70,000–115,000 (up to 90th percentile)**	**\$115,000 and above (above 90th percentile)**
Simplified Avoid	-0.0347	0.0490	-0.00511	-0.00742	-0.00114	-0.0549***
	(0.0670)	(0.0341)	(0.0172)	(0.0161)	(0.0203)	(0.0213)
First-timers	-0.0671	-0.0457	-0.0377**	-0.0533***	-0.0417*	-0.0959***
	(0.0625)	(0.0354)	(0.0186)	(0.0179)	(0.0224)	(0.0238)
Simplified Avoid × First-timers	0.0279	-0.00143	0.0185	0.0202	0.000879	0.103***
	(0.0891)	(0.0490)	(0.0263)	(0.0251)	(0.0317)	(0.0330)
N	387	1,337	4,901	5,624	3,616	3,213

Notes: The table shows the probit regression analysis for the open rates for the six income groups. Marginal effects are presented, and delta-method standard errors are in parentheses. Simplified Avoid is the condition and the underlying variable is the Simplified Save condition. First-timers is a dummy variable for new Overdraft Prediction users. Simplified Avoid × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13. Tobit Regressions for Number of NSF Fees One Week After the First Reminder and Four Months of the Experiment: The Simplified Avoid Group Versus the Simplified Save Group, by Income Group

Variable	four months of experiment						one week after reminder					
	Up to $10,000	$10,000–20,000 (up to 25th percentile)	$20,000–40,000 (up to 50th percentile)	$40,000–70,000 (up to 75th percentile)	$70,000–115,000 (up to 90th percentile)	$115,000 and above (above 90th percentile)	Up to $10,000	$10,000–20,000 (up to 25th percentile)	$20,000–40,000 (up to 50th percentile)	$40,000–70,000 (up to 75th percentile)	$70,000–115,000 (up to 90th percentile)	$115,000 and above (above 90th percentile)
Simplified Avoid	-0.0890	0.120	-0.0945**	-0.0327	-0.0354	-0.0853	0.0206	0.0763	-0.0605	0.000898	0.0101	0.00279
	(0.156)	(0.0748)	(0.0377)	(0.0362)	(0.0462)	(0.0540)	(0.169)	(0.0990)	(0.0516)	(0.0576)	(0.0773)	(0.105)
First-timers	-0.734***	-0.497***	-0.633***	-0.794***	-0.882***	-1.021***	-0.0975	-0.338***	-0.423***	-0.571***	-0.595***	-0.791***
	(0.136)	(0.0748)	(0.0390)	(0.0382)	(0.0494)	(0.0554)	(0.158)	(0.0858)	(0.0502)	(0.0514)	(0.0659)	(0.0906)
Simplified Avoid ×First-timers	0.0279	-0.0332	0.0485	0.0239	0.0534	0.0158	-0.169	-0.0408	0.0662	0.0373	-0.0775	0.0309
	(0.194)	(0.103)	(0.0555)	(0.0536)	(0.0693)	(0.0769)	(0.210)	(0.127)	(0.0695)	(0.0733)	(0.0952)	(0.125)
N	387	1,337	4,901	5,624	3,616	3,213	387	1,337	4,901	5,624	3,616	3,213

Notes: The table shows the tobit regression analysis of the number of non-sufficient funds (NSF) fees one week after the first reminder sent and four months of the experiment for the six income groups. Marginal effects are presented, and delta-method standard errors are in parentheses. Simplified Avoid is the condition and the underlying variable is the Simplified Save condition. First-timers is a dummy variable for new Mint Overdraft Prediction users. Simplified Avoid × First-timers is the interaction variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1

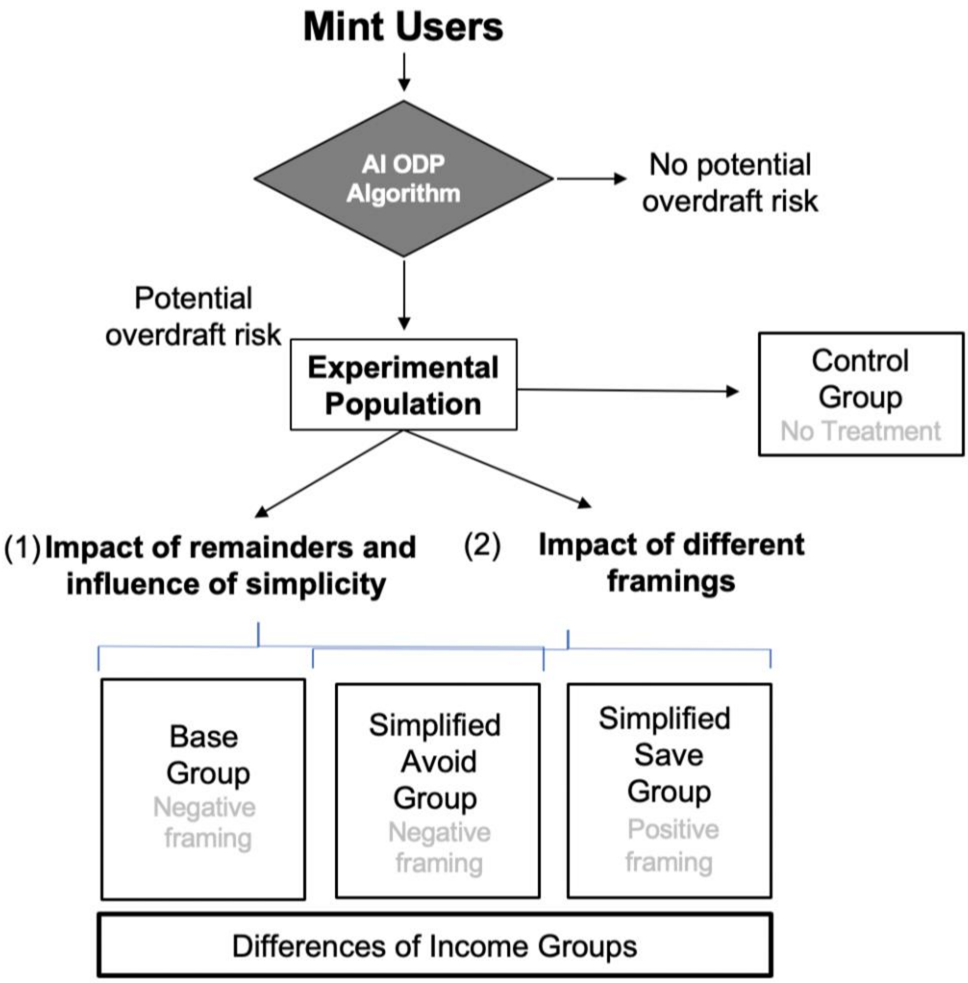

Figure 1. Experimental design and overview of participant allocation to the different conditions in the experiment. AI = artificial intelligence; ODP = Overdraft Prediction.

Figure 2

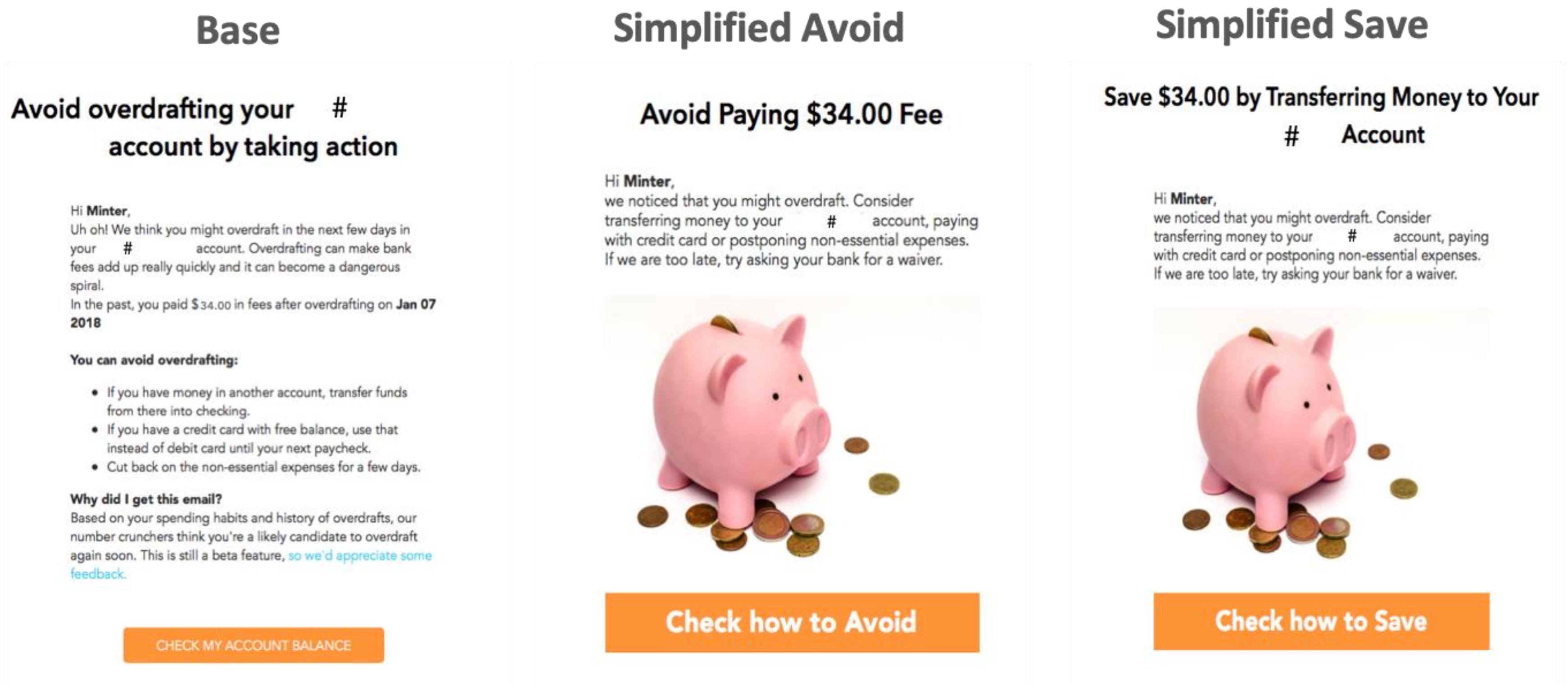

Figure 2. The Base screen represents the original reminder that was sent to Overdraft Prediction algorithm users. The Simplified Avoid and Simplified Save versions were constructed in the experiment.

Figure 3

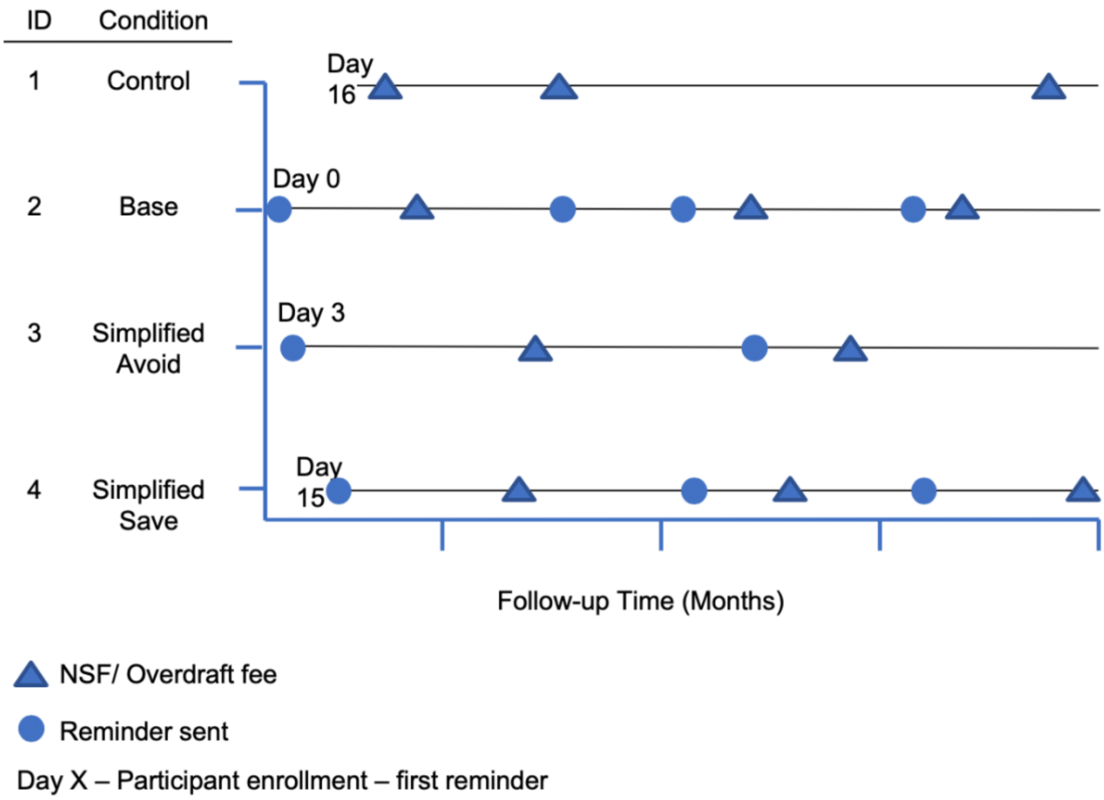

Figure 3. Illustration of the follow-up period in months for the participants in different conditions. Over the first month, new participants could be assigned, and we subsequently followed them to the end of the experiment and for at least three more months.

Figure 4

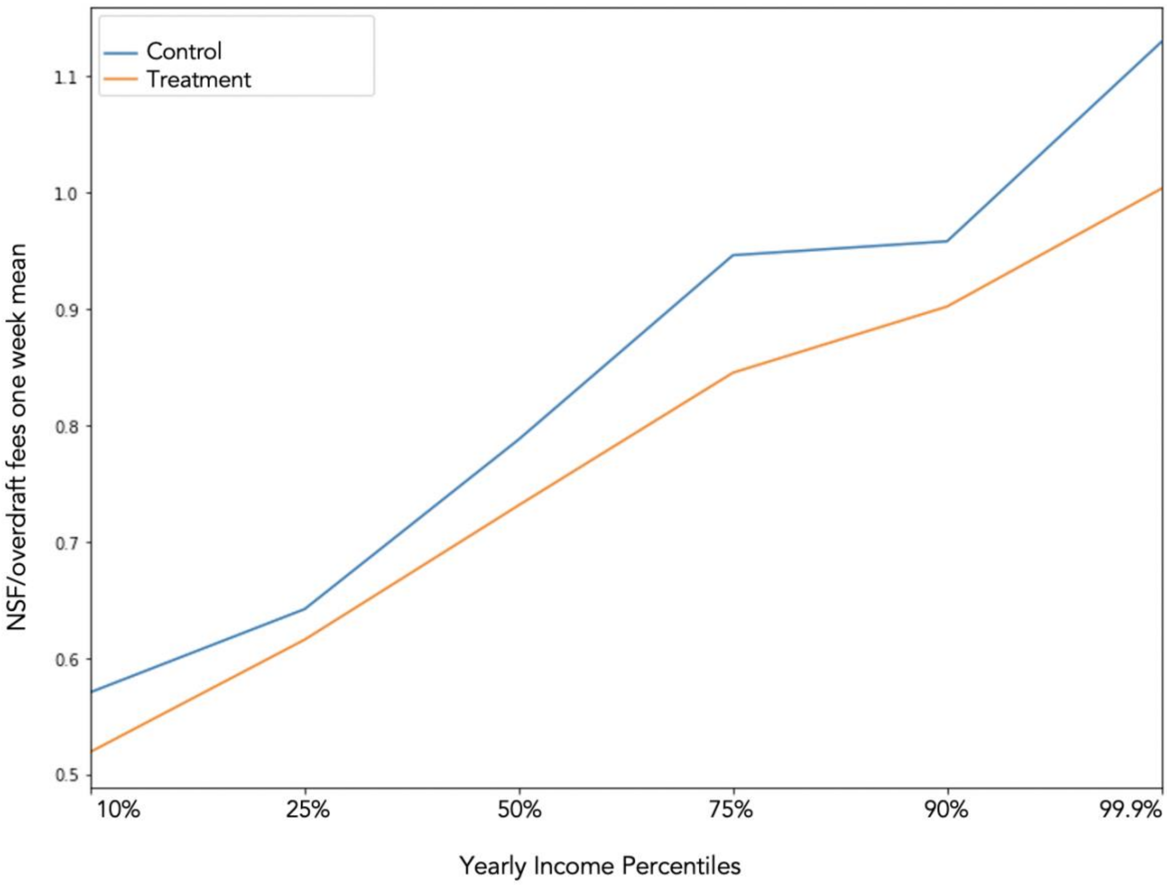

Figure 4. A plotline showing the distribution of the average number of fees for non-sufficient funds one week after the first reminder for the treatment versus control (no treatment) populations with respect to the six income levels examined: Annual income less than $10,000; (2) up to 25th percentile, $10,000–$20,000; (3) up to 50th percentile, $20,000–$40,000; (4) up to 75th percentile, $40,000–$70,000; (5) up to 90th percentile, $70,000–$115,000; (6) above 90th percentile, $115,000 and above).

Appendix: Simplification Manipulation Check

Experimental methodology

We used a computerized questionnaire that contained three items. The survey was administered to 107 U.S. participants via Amazon's Mturk, a widely used internet platform. Every participant received the same three questions and had 2 min to answer them. Participants earned $1 USD for answering the three questions.

Experimental design

Before participants received the questions, we showed them two screens, one with the Base alternative and the other with what we have called the Simplified Avoid alternative (see Figure 2). The questions were as follows: (1) Which one of the two alternatives do you think is simpler? (2) Is the alternative you chose easier to comprehend? (3) Do you think the alternative you chose is more vivid? The results corroborate that simplification was perceived as intended: 76.6% stated that the new reminder is simpler, 74.75% that it is easier to comprehend, and 71% that it is more vivid.