Don't Swipe the Small Stuff: A Randomized Evaluation of Rules of Thumb-Based Financial Education

We perform the first rigorous test of a rules of thumb–based approach to financial education on consumer behavior and outcomes. We test two rules of thumb that are targeted at reducing credit card revolving and deliver them in a randomized fashion via email, online banner, and physical mailer. Using monthly administrative data and pre- and postintervention credit data on almost 14,000 consumers, we find that the "Don't swipe the small stuff" rule of thumb reduces participants' targeted credit card balance by an average of 2 percent at a cost of around $0.50 per person. The "Credit keeps charging" rule shows a decline as well but the impact is not significant.

Keywords*:* Financial education, randomized control trial, rules of thumb, credit, debt

JEL Classification: D14

Rules of thumb may be an effective strategy for delivering financial education and improving financial behaviors and outcomes. Rules of thumb are simple heuristics with broad application that are not meant to be strictly accurate in every situation. They are an efficient tool when decision-making is costly and when decision-makers have imperfect information (Baumol and Quandt 1964; Day 1967), as is often the case with financial decisions. Rules of thumb can lead to optimally imperfect decision-making, in which the marginal cost of additional information gathering is equal to its marginal gross yield (Baumol and Quandt 1964).

To test whether rules of thumb can be used to effectively deliver financial education, we perform the first rigorous test of a rules of thumb–based approach to financial education on consumer behaviors and outcomes using a randomized controlled trial (RCT). In collaboration with Arizona Federal Credit Union (Arizona Federal), we send one of two rules of thumb (or none for the control group) to 13,957 credit card revolvers. Credit card revolvers are individuals who carry debt on their credit card from month to month.

The first rule that we test tells participants to not "swipe the small stuff" and to use cash instead of credit for purchases under $20. The second reminds them that "credit keeps charging" — that for the average revolver, purchases cost about 20 percent more than the ticket price when they pay with their credit card (due to compounding interest and fees). Participants receive one (or no) rule via email, online banner, and/or physical mailer throughout the course of a 6-month study period.

Using detailed monthly administrative data on participant purchases, payments, and checking account behaviors, and pre-and postintervention credit data from a large credit reporting firm, we estimate the average treatment effect using a difference-in-difference approach for participants offered access to treatment on key financial behaviors and outcomes. We find that the first rule of thumb causes participant debt on their Arizona Federal credit card to be an average of $104 (or 2

percent) less than it would have been in the absence of treatment. The second rule leads to an average decrease of $58 (1 percent), but the effect is not statistically significant. We do not detect a statistically significant effect on aggregate revolving debt which makes us unable to rule out the possibility of debt reshuffling rather than reductions in debt overall, although the negative point values are promising.

The mechanisms for the decrease in debt are not fully clear, however, leading to some uncertainty about whether the messages are functioning as true rules or whether they are functioning as general reminders. No effect is detected on the number of purchases under $20 or on the number of purchases overall, although the point estimates are all negative and the effect is statistically significantly for some subgroups. Additionally, the fact that the first rule, which included a more direct action request (to use cash instead of credit on small purchases) worked better, implies that messages framed as rules of thumb may be more effective than those framed as general suggestions. We hypothesize that the rules work through a variety of mechanisms (reductions in spending via credit, reductions in spending overall, and increases in payments), which makes it difficult to detect an effect on each individual mechanism.

Our results are smaller in magnitude but similar in sign to those of Drexler, Fischer, and Schoar (2014), who test rules of thumb on microentrepreneurs in the Dominican Republic. They find that their rules of thumb-based training for business owners increases savings by 6 percent. Skimmyhorn et al. (2015), conversely, find that classroom-based rules of thumb education for students at West Point produces no benefit over traditional principles-based financial education in terms of self-reported knowledge or motivation. However, the authors are not able to observe the effects of the intervention on actual behaviors and outcomes. Additionally, they test rules of thumb in a traditional classroom based setting, which may work differently than those sent via messages.

We extend previous work by testing rules of thumb on U.S. consumers using a messaging-based approach. We also directly observe participants' behaviors and financial outcomes rather than their financial knowledge and self-reported behaviors.

Although the effect sizes are modest, the costs of delivering the intervention are trivial: rule delivery via e-mail or online banner was $0.47 per person, and production and delivery of the physical mailer was $0.59 per person. The marginal cost of adding an additional recipient to the email or online banner delivery is close to zero. Additionally, the results are likely attenuated because the sample was not ideal. Preintervention credit card purchase levels were surprisingly low for most participants, so altering their spending behavior was particularly difficult. We were also not able to measure actual uptake of the rules, so while this RCT can estimate the intent to treat effect, we could not estimate the treatment on the treated effects. (In this case, intent to treat is the effect of being offered one of the rules, while treatment on the treated is the effect of actually opening, logging on, and observing the rules as delivered by email, website, and calendar magnet.

These results suggest that rules of thumb can be effective at improving financial behaviors and outcomes for U.S. consumers. And, given that consumers may be less likely to trust rules of thumb that come from a banking institution, such an intervention may be even more effective when the rules come from a trusted nonprofit or public entity. However, not all rules work equally well, so selecting the right rules is key to altering behaviors. Additionally, research is needed to better understand the mechanisms through which rules affect behavior.

1. Background on Rules of Thumb

Rules of thumb are among the more efficient methods for optimal decision-making (Baumol and Quandt 1964), and are also widely used, even if not widely researched. They are likely a more realistic way that people make decisions than rational choice, in which the choice maker weighs

all costs and benefits (Etzioni 1987). In many cases, the full set of matters to weigh into decision-making are unknown or too many. Frank (1987) argues that to gather the information and do the calculations required for a rational choice decision would consume more time and energy than anyone has. Rational choice decision-making, he says, would lead to the ability to make only a few decisions each week, leaving many important decisions unmade. Rules of thumb reduce this cost and allow for quicker decision-making.

Because of these factors, rules of thumb may be an effective way to deliver financial education and improve consumer behaviors and outcomes. Historically, financial education has been didactic, often in a classroom or workshop setting, and focused on sharing general financial knowledge. The empirical evidence is quite limited about the effectiveness of these interventions, and the research that does exist has provided mixed results (Collins 2013; Hastings, Madrian, and Skimmyhorn 2012).

In contrast, rules of thumb are easy to understand and remember, and consumers may grasp the point of the rules of thumb and adopt it more readily than a complex financial calculation. They may also be more likely to implement the rule of thumb correctly, making it more effective than conventional financial education approaches. Rules of thumb are also easy to follow, which increases the probability that consumers will adopt new behaviors and stick with them. But rules of thumb may be most effective for decisions that are frequent and involve learning -- they may not be successful when applied to infrequent financial decisions, such as taking out a mortgage for a house or a loan for a car.

Rules of thumb have been found to be a successful method for optimal decision making in many areas (Fishcher et al. 2010; Baumol and Quandt 1964). They are used in the medical profession to help doctors handle the huge number of decisions they encounter (Andre et al. 2002). They have

also been found to be used by animals to evaluate the benefits and costs associated with a particular behavior (Pyke 1978; Barnard and Brown 1981; Bergelson 1985; Blaustein and O'Hara 1982; Stamps 1988). For example, Bouskila and Blumstein (1992) found that natural selection will not always favor perfect estimates, since there is some cost in acquiring accurate information. The authors found that there is a zone of tolerance where inaccurate perceptions, such as rules of thumb, perform just as well as perfect knowledge for predation. While developing and providing rules of thumb as a financial education tool has promise, they may not be adequate in all settings, as there is likely a trade-off between the simplicity of understanding via rules of thumb and the depth of understanding that might come from a more standard educational approach. Financial decisions can be complex and rules of thumb may be too simplistic. Rules of thumb-based financial education may make it difficult for individuals to adapt the knowledge to different situations and circumstances without an in-depth appreciation of the concepts behind the rules.

Until now, there has been limited empirical evidence to show whether rules of thumb interventions can be successful at improving financial outcomes for consumers. Techniques similar to rules of thumb such as nudges and messaging reminders have been shown to be effective at improving financial outcomes, particularly for those that are close in time to the point of decision making (Agarwal et al. 2015; Bracha and Meier 2014; Karlan et al. 2010; Kast, Meier, and Pomeranz 2010; Stango and Zinman 2011). Messaging reminders are often personalized to the individual, with notes about goals or updates on bank balances and credit scores. For example, Credit Card Accountability, Responsibility, and Disclosure (CARD) Act mandated nudges display the monthly payment required for a credit card holder to pay off their full balance in 36 months; the nudge was found to increase the share of accounts meeting the recommended value (Agarwal et al. 2015), though minimum payment requirements remain salient for many consumers (Keys

and Wang 2016). Similar to rules of thumb, messaging reminders can be sent to recipients close in time to the point of decision-making, and they can be simple and short.

Additionally, two studies directly examine rules of thumb as a method of financial education. Drexler, Fischer, and Schoar (2014) use a randomized controlled trial to compare the effects of standard accounting–based financial business management education to a rules of thumb–based program for microentrepreneurs in the Dominican Republic. They find that the participants who received rules of thumb–based training are significantly more likely to keep accounting records, calculate monthly revenues, and improve their business management practices overall; participants in the traditional training do not change their practices. This study shows that rules of thumb hold promise for improving business practices, but it did not inform how well this approach would translate to consumers, nor (like any single study) was it able to examine a full range of outcomes.

Skimmyhorn et al. (2015) examine the effect of classroom-based rules of thumb financial education on self-reported measures and find that although rules of thumb increase cognitive measures of knowledge and noncognitive measures of self-efficacy, motivation to learn, and willingness to take financial risks, this approach does not produce superior results above those from traditional principles-based financial education. However, the authors are not able to examine impacts on financial behaviors or outcomes, nor are they able to exploit the ability of rules of thumb to be delivered close in time to the desired behavior change.

2. Experimental Design and Data

2.1 Rule Design and Delivery

Our design process for developing rules had 6 phases and was designed and implemented with our research partners at the Consumer Financial Protection Bureau and Commonwealth (formerly D2D Fund). We first reviewed previous research and held discussions with experts to improve our

knowledge of what works when designing a rule of thumb. Second, we did structured consumer listening sessions to gain deeper insight into the behaviors and attitudes of the project's target audience. Third, we brainstormed ideas—generating a wide range of possibilities for further testing. Fourth, we vetted ideas internally and externally to narrow the focus. Fifth, we identified a partner to implement the experiment, described below. We then worked with marketing professionals to refine and design the rules to make them memorable and visually appealing.

The rules are as follows:

- "Don't swipe the small stuff. Use cash when it's under $20." (the cash under $20 rule)

- "Credit keeps charging. It adds approximately 20 percent to the total." (the 20 percent added rule)

Both rules are designed to encourage revolvers to spend less. *The "Credit keeps charging"* rule directly addresses the expenses incurred by using a card—with the logic being that consumers would prefer not to overpay for a good or service. The logic behind the *"Don't swipe the small stuff"* rule is more nuanced. "Pain of payment" refers to the displeasure that people experience immediately after making a purchase, which can undermine the pleasure derived from consumption (Prelec and Loewenstein 1998). This pain varies by the timing of payment, and by method of payment. Whereas paying with cash leaves a vivid memory trace, credit card purchases requires only a signature, and therefore, they have a lower pain of payment. This phenomenon may lead to higher consumption with credit cards than with cash.

Many studies have found that using credit cards as a payment mechanism increases the propensity to spend as compared to cash in otherwise identical purchase situations, a finding typically referred to as the credit card premium (Feinberg 1986; Hirschman 1979; Prelec and Simester 200; Soman 2001; and Soman and Cheema 2002). Chatterjee and Rose (2012) found that

when credit cards as a payment mechanism are more accessible, consumers pay more attention to a product's benefits rather than its costs. Conversely, when cash as a payment mechanism is more accessible, consumers focus more on cost aspects than benefits. Both of these mechanisms lead consumers to spend more with a credit card than they would with cash.

Our implementation partner was Arizona Federal Credit Union (Arizona Federal), a large credit union based in Phoenix, Arizona, to draw the sample of credit card revolvers and deliver the rules. Arizona Federal delivered the rules via one, two, or three delivery channels in a randomized fashion. A control group received no rules. The three delivery modes were e-mail, online web banners at log-in, and a physical mailer (calendar magnet).[1]

Arizona Federal sent participants the rules via e-mail twice each month. They placed the online portal messages on the home page of the participants' online banking site in either a moving banner or a static ad, with variations in type and style throughout the intervention period. The physical mailer incorporated the rules into a magnetic calendar that Arizona Federal had already planned to send to their customers as an end-of-the-year gift. Study participants randomized to receive the rules via physical mail had one of the rules printed at the top of the magnet. Arizona Federal credit card customers randomized to not receive the rules via physical mail received a magnet with the Arizona Federal logo instead of a rule.

2.2 Experimental Design

This study was undertaken via a RCT. To test multiple rules, delivery mechanisms, and timing, we employed a randomized full factorial design. A RCT with factorial design allows for

---

[1] There is a small chance of spillover of the intervention within a household, but this is limited to households that had two separate accounts, both had revolving credit card debt, and where one was randomized to receive the mailer and the other was not. The intervention was undertaken at the account level, so if two family members from the same household shared an account, they counted as one observation in our analysis. If they had two accounts, the only rule that could potentially have been shared between them was the mailer, since that would could have been opened by any family member and/or place on the refrigerator (since it was a magnet). The email went directly to the individual, as did the banner on the online site. Therefore, we believe that the chance of spillover is small and limited to households that had two separate accounts, both had revolving credit card debt, and where one was randomized to receive the mailer and the other was not.

combinations of treatment types, rather than administering just one type of treatment for each participant. We had four factors: online portal, e-mail, physical mail, and rule type (A or B).[2]

Even when assigning individuals at random, it is possible for there to be baseline nonequivalence between the different groups by chance. To help ensure that important subgroups were selected into each treatment group and the control group in proportion to their overall representation in the study sample, we first stratified the sample based on two measures (number of months revolved and age) and then randomized within these groups.

2.3 Data

We draw data for this study from two sources: administrative data from Arizona Federal and demographic and credit data from a large credit reporting firm. The administrative data include information about each revolvers' age and accounts (including credit, checking, and savings). These data include our main variable of interest: amount of revolving credit card debt on participants' Arizona Federal credit cards. Descriptive statistics can be found in Table 1, along with baseline randomization tests. Arizona Federal collected and transferred these data to us monthly for 6 months before the intervention and 6 months during the intervention. These detailed monthly data allow us to account for individual-level trends and baseline characteristics and to examine whether the intervention's effects varied throughout the course of the study period. However, we are not able to observe the extent to which recipients opened the emails, read the online banners, or viewed the physical mailer. As a result, we can only estimate the effect of the offer of treatment, not the effect of actually receiving and participating in the treatment.

---

[2] Consumers were distributed into 15 treatment groups: 7 groups received rule A, 7 received rule B, and the final group received no rule. Other than the control group, each group received the rule via some combination of the three delivery mechanisms with 2,616 consumers receiving the rule via two modes, and 872 consumers receiving the rule via three modes.

The credit data come from a large credit reporting firm and include information on debts and available credit in addition to some demographic information. Because participants may hold accounts at more than one bank and through nonbank channels, it is important to test whether their overall finances changed as a result of the intervention.

To examine both pre- and postintervention credit measures, we pulled credit data from November 2014 (the month before the beginning of the intervention) and August 2015 (two months after the end of the intervention). With only a small number of exceptions, the treatment group means are not statistically different from the control group means at baseline. Baseline characteristics do not jointly predict treatment assignment. We do find recipients of the $20 rule to have purchased, on average, $13 less than the control group at baseline; while this amount is statistically significant at the 5 percent level, the small size of the difference leads us to believe it is not economically significant. Rule recipients also have, on average, $100 lower credit card balances with Arizona Federal. To adjust for these differences at baseline, we use an individual fixed effects model as described further in section 3. These means are based on trimmed data.[3]

[TABLE 1 HERE]

2.4 Study Participants

The participants for this study are credit card revolvers drawn from Arizona Federal's credit card customer base. We define a credit card revolver as someone who carried a credit card balance for at least two of the 6 pre-intervention months, not necessarily consecutively. This definition is based on Arizona Federal's internal definition and was defined to correspond with the definition put forth in the CARD Act (CARD Act Report 2013). Arizona Federal described a problematic revolver as

---

[3] Because extreme outliers existed for several continuous outcome measures, we trimmed the data to ensure that these individuals did not skew the results. To do so, we removed the top and bottom 0.5 percent of each continuous variable when the top or bottom value was not zero. In cases in which the top or bottom values were zero, we trimmed only one side of the variable's distribution by 0.5 percent. All results are robust to untrimmed data and a 0.25 percent trim.

someone who has revolved more than 2 months out of the past 12 months (i.e., the person is revolving in a suboptimal manner). However, because we had only 6 months of preintervention data from Arizona Federal, we designated revolvers as those people who revolved at least 2 months of the 6 months of the preintervention period.

The CARD Act defines a revolver as someone who is not a "transactor" (i.e., someone who pays their balances in full for two consecutive months). Thus "revolvers are consumers who are required to make a payment but do not pay their balance in full for two consecutive months" (CARD Act Report 2013, 95). However, the CARD Act does not define a specific period within which this revolving must take place, and taking their strict definition of someone who revolved just once would place some people in our sample who may not be problematic revolvers—that is, they may be acting in their own best interest. For example, at certain times during the year sales on large items are so great that it is better for someone to purchase the sale item with a credit card and pay interest on that credit card debt rather than wait to purchase the item at full cost. Only including individuals who revolved for at least two of the 6 months of preintervention data helped us reduce the likelihood of including utility-maximizing revolvers in the study. The full sample consists of 13,957 people.[4]

Participants do not typically spend large amounts each month on their Arizona Federal credit card, but they hold high balances. Seventeen percent made no purchases on their Arizona Federal credit card during the preintervention months and 55 percent spent less than $100 on their credit card every month. Thirty-one percent made no purchases under $20 during the preintervention period, and another 21 percent made on average less than one purchase under $20 per month during the preintervention period. Low credit card spending suggests participants may be strategically

---

[4] We find that 0.7 percent of consumers pay off their balance in full for 5 or 6 of the 6 months prior to the intervention.

revolving at Arizona Federal due to the low interest rate this credit union offers on credit card balances. This pattern of credit card use was unknown when the sample was drawn and likely caused the effects to be smaller than they would have been on a more active population. At baseline, the typical revolver in this study made payments sufficient to cover the cost of new purchases, interest, and a small share of the outstanding principal, but he or she was not making quick strides in retiring existing credit card debt.

For the most part, study participants are working-age adults in their forties and fifties who are married with at least one child (Table 2). Most study participants are homeowners. The sample has slightly more males than females.

[TABLE 2 HERE]

3. Results

3.1 Average Treatment Effects

Since we cannot observe whether the participants read and implemented the rules (which would inform an analysis approach referred to as the "treatment on the treated"), we estimate the effect of the opportunity to read and apply the rules (the "intent to treat"). Our primary method for estimating the Intent to Treat (ITT) impact of the delivery of rules of thumb on financial outcomes is a fixed-effects model, which is a form of difference-in-difference. Specifically, we estimate the following equation:

$$Y_{i,t} = c + \beta_A (R_i x Post_t) + \gamma_i + \lambda_t + \varepsilon_{i,t}, \qquad (1)$$

Where $Y_{i,t}$ is the outcome variable for participant i in month t; c is a constant term; $R_{i,t}$ a vector of dummy variables for each of the two rules; β_A is a measure of the effect of the effect of Rule A and Rule B; Post is an indicator of whether the month is pre or post implementation of the rules,

13

γ_i and λ_t are individual and month fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term.[5] We calculate standard errors clustered at the individual level that are robust to heteroskedasticity and arbitrary forms of error correlation within each individual. The analysis includes 6 months prior to the intervention and 6 months during the intervention.[6] The fixed effects allow us to remove any time-invariant unobserved heterogeneity that exists for that individual that may be related to their financial outcomes.

To test how the effects vary by month, we additionally run the following fixed effects model:

$$Y_{i,t} = c + \beta_A (R_i x \lambda_t) + \gamma_i + \lambda_t + \varepsilon_{i,t}, \qquad (2)$$

Where $Y_{i,t}$ is the outcome variable for participant i in month t; c is a constant term; β_A is a measure of the effect of the effect of Rule A and Rule B; $R_{i,t}$ a vector of dummy variables for each of the two rules; γ_i and λ_t are individual and month fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term.[7]

Our primary outcomes of interest are balance on Arizona Federal accounts and aggregate balance for open revolving trades (Table 3). We find that the cash under $20 rule causes participants to reduce their revolving debt by $104 on their Arizona Federal credit card which equates to about 2 percent of their baseline average. We also estimate decreases in aggregate debt across all revolving trades that are approximately equivalent in magnitude to our estimates from the Arizona Federal card itself.[8] However, these estimates are much noisier and are not statistically significant, so we

cannot draw any definitive conclusions regarding overall debt. The 20 percent rule also reduces

balances, but without as large a magnitude or statistical significance.[9]

Collectively, these results suggest that it is possible that the rules are serving to encourage

participants to shift use to other credit cards rather than reducing overall credit card spending.

However, we believe that this is unlikely since Arizona Federal's credit card interest rates are

relatively low so it is unlikely that participants would choose to instead use a different card, and

because we do not precisely estimate a zero effect on overall revolving debt. Additionally, we see

other secondary impacts on credit card usage overall (number of inquiries and aggregate credit

below) which implies that the rules are not causing participants to shift debt to other cards, but are

instead reducing credit card usage overall. However, since our overall debt measure does not

include balances on closed accounts, the effects we see could be due to account closures rather

than true reductions in debt. Therefore, additional research is needed to confirm that such rules of

thumb encourage overall debt reductions rather that reshuffling or account closures.[10]

[TABLE 3 HERE]

We hypothesized three possible mechanisms through which the $20 rules could have an effect

on credit card balance: 1) decreases purchases overall and specifically under $20 on the Arizona

card, 2) decreased purchases on other credit cards, leading to more money for and higher payments

on the Arizona card, and 3) generally more financial awareness leading to fewer overall purchases

on their credit cards and more frequent or larger payments on their credit cards.[11]

---

[9] The effects for the cash under $20 rule and the 20 percent added rule are not statistically different from one another.

[10] We do a robustness test on the 4 percent of individuals with only 1 revolving trade at baseline. We find the point estimates for the $20 rule to be very close to that of the full group and the point estimate for the 20% rule larger (i.e. more negative).

[11] While presumably the effect on the overall balance should come from either fewer purchases or more payments, due to the nature of the regression model the coefficients on the intervention for purchases and payments will not necessarily add up to the coefficient on balance. We hypothesize that it is a combination of these mechanisms that lead to the decrease in balance, although we cannot test this.

We are not able to clearly estimate an effect of the rules directly on any of these potential mechanisms, however (Table 4). We do not detect a significant change in number or value of overall purchases on Arizona Federal accounts for either rule, nor on the number of purchases under $20 on Arizona Federal credit cards. And for payments, we actually estimate a small decrease in the number of payments for recipients of the $20 rule. We hypothesize that this is due to the recipients making fewer purchases on their Arizona Federal cards and therefore requiring fewer payments (as some individuals have autopayment set up after each purchase), since we would normally expect fewer payments to be associated with higher balances. However, we are unable to confirm this since the point estimate on the number and amount of purchases is insignificant for both rules. We also examine the impact on savings, checking, and debit card transactions and find no significant impact of either rule (shown in appendix Table A.1). Therefore, we cannot definitively say what the mechanism is behind the change in Arizona Federal credit card debt, but suspect that it is some combination of the three.

[TABLE 4 HERE]

We do find that the 20 percent added rule reduces the number of credit inquiries[12] and that both rules lead to a decrease in aggregate credit for open revolving trades (by $595 for the $20 rule and $679 for the 20 percent added rule). This decrease in available credit does not lead to a detectable change in overall credit utilization ratios (the ratio of aggregate balance to available credit for open revolving trades), however, and the rules have no detectable effect on overall credit score or other credit measures. We also estimate the impact by type and number of delivery mechanisms (see appendix Table A.2 for results). We find no clear winner in terms of delivery

[12] The effect of the 20 percent added rule is not statistically distinguishable from the effect of the cash under $20 rule.

mode (e-mail, online, or physical mailer) while receiving a rule via fewer channels seems to work better than receiving it via all three.[13]

3.2 Effects Over Time

Because Arizona Federal provided monthly data, we were able to examine how the rules worked over time. For instance, it is possible the rules had the strongest effect toward the beginning months of the intervention before fatigue set in. Alternatively, the rules may have been more effective after they had a chance to soak in for a few months and for new behaviors to become habits.

To test for effects over time, we estimated the following equation:

$$Y_{i,t} = c + \sum_{m=-6}^{6} \beta_t T_{i,t-m} + \gamma_i + \lambda_t + \varepsilon_{i,t},$$

where $T_{i,t-m}$ was a dummy variable for treatment that equaled one for all months for anyone who received treatment (by rule), and

$$\sum_{m=-6}^{6} \beta_t T_{i,t-m}$$

is a series of coefficients and indicators for each month before and during treatment (6 months before treatment began and 6 months during treatment). In the 6 months before the intervention, β_t should be equal to zero. If the rule was effective, β_t should have positive or negative values (depending on the variable) in the 6 months after the treatment began. This equation provided a

---

[13] We estimate the impact by delivery mechanism using the same regression model as equation (1), but allow treatment to have differential effects based on the three delivery mechanisms (online, e-mail, or physical mailer). Specifically, we estimate the following equation: $Y_{i,t} = c + \beta_A (R_{i,t} * M_i) + \gamma_i + \lambda_t + \varepsilon_{i,t}$, where $Y_{i,t}$ is the outcome variable for participant i in month t; c is a constant term; $(R_{i,t} * M_i)$ is the interaction of a vector of dummy variables for each of the two rules and a vector of dummy variables for each of the delivery modes; γ_i and λ_t are individual and month fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term. We also estimate the impact by number of delivery mechanisms using the same regression model as equation (1), allowing treatment to have differential effects based on the number of delivery mechanisms. Specifically, we estimate the following equation: $Y_{i,t} = c + \beta_A (R_{i,t} * F_i) + \gamma_i + \lambda_t + \varepsilon_{i,t}$, where $Y_{i,t}$ is the outcome variable for participant i in month t; c is a constant term; $(R_{i,t} * F_i)$ is the interaction of a vector of dummy variables for each of the two rules and a vector of dummy variables for each of the number of delivery modes; γ_i and λ_t are individual and month fixed effects, respectively; and $\varepsilon_{i,t}$ is the error term. These results can be found in appendix Table A.2.

valuable test of the randomization and allowed us to measure the effect of the intervention in each month.

The effects of the $20 rule increased over time for credit card balance (Figure 1), but the 20 percent rule did not show any increase over time for credit card balance (Figure 2). The rules did not show any discernible pattern for the other outcome measures. It is possible that this observed effect is due to the compounding and cumulative nature of balance rather than to an increase in behavior change over time.

[FIGURE 1 HERE]

[FIGURE 2 HERE]

3.3 Heterogeneous Treatment Effects

The rules of thumb intervention may have heterogeneous effects on distinct groups within the population, such as for those making more frequent versus less frequent purchases at baseline. To assess this possibility, we estimate equation (1) separately for each subgroup within the categories of preintervention number of purchases under $20, total number of preintervention purchases, and initial credit score.[14] The results for each of these subgroups are presented in Table 5 for the primary outcomes and Table 6 for the potential mechanisms.[15]

We chose these specific subgroups for a number of reasons. We test for heterogeneous effects based on number of purchases below $20 which is especially relevant for the $20 rule, and by total number of baseline purchases which gives some sense of the initial 'exposure' to an intervention designed to affect spending behavior. If participants infrequently make purchases under $20, detecting an effect of the first rule is more difficult. Therefore, estimating effects for the subsample

---

[14] The treatment control ratios remain relatively balanced across all of the subgroups, with the control group ranging from 12-13 percent and the treatment groups ranging from 43-45 percent. The minimum detectable effects for each subgroup are all below 0.25, with all but two below 0.20.

[15] Adjusting for multiple outcome tests for these subgroups leads to any significant results no longer being significant.

who had a larger number of purchases under $20 at baseline allows us to gain precision in our estimates. Similarly, if participants infrequently use their credit cards, neither rule is likely to produce much of an effect. In particular, participants who use their cards primarily as a place to hold debt rather than as a means of transacting are likely difficult to influence via the rules.

Finally, previous literature has suggested that rules of thumb may work differently for those of high and low financial standing (Skimmyhorn et al. 2015), so we examine that dimensions as well. In particular, rules of thumb are hypothesized to work better for people who have low levels of financial standing. Therefore, we create subgroups based on initial credit score to test this hypothesis.

Baseline number of purchases.—Because the population may not have been ideal in terms of how frequently many members used their credit cards prior to treatment, we estimate the effects of the intervention for subgroups of participants based on how frequently they made purchases on their credit cards prior to treatment. We separate the sample into participants who made 1 purchase or less on average per month in the 6 months prior to treatment, between 1 and 10 purchases per month, and 10 or more purchases per month. We also divide the sample into groups based on their baseline number of purchases below $20, because those are the purchases targeted by the first rule. The groupings for these are 1 purchase or less under $20 on average per month in the 6 pretreatment months, between 1 and 5 purchases, and 5 or more purchases. Although these subgroups get us a bit closer to an ideal population of card users, the cutoff for the number of purchases in the high purchase buckets are lower than those of an ideal population, so we are unable to fully test how these rules might work on a population of more frequent credit card users.[16]

---

[16] We are unable to draw the buckets with higher cutoffs due to sample size.

Generally, effects are stronger for participants who made a greater number of purchases on their credit card at baseline. Although we detect no clear effect on credit card balance based on the total purchase subgroups, we do find that participants who were at least somewhat frequent users of their credit cards (those who made between 1 and 10 purchases per month pretreatment) have fewer total purchases and fewer purchases under $20 as a result of the rules. This group also has larger reductions in aggregate credit for open revolving trades and a reduction aggregate balance for open trades (at least for those who received the 20 percent added rule). There are less clear patterns of effects for the highest purchase-volume group (those who made more than 10 purchases per month at baseline). For example, the point estimates on (decreasing) credit card balance appear to indicate that the effect is strongest for this group, but the estimates are noisier so we cannot say that these results are statistically different from zero. This may be in part because this "purchase" group is the smallest of the three. We find no discernible pattern for subgroups based on average number of pretreatment purchases under $20. We suspect that this is due to our selected cutoffs for subgroups, which we are unable to make larger due to sample size constraints.

Baseline Credit Score.—To examine whether effects are stronger for participants who began with a higher or lower level of overall financial standing, we estimate effects by subgroup of initial credit score. We divide study participants into three subgroups of close to equal size: one subgroup for participants with credit scores 670 or less, another with scores between 670 and 730, and another with scores of 730 or greater. Most of the effects of the rules appear to be driven by participants with credit scores between 670 and 730, that is, those with fair to good credit. This finding opposes our hypothesis that rules of thumb work better for those with lower financial standing, which has been suggested in previous studies (Skimmyhorn et al. 2015). However, participants in this study have fairly high financial standing overall, so a population with truly low

initial financial status might perform differently. In addition, participants with a baseline credit score of 730 or greater actually have a marginally lower credit score than they would have had in the absence of treatment, which mirrors results found in Bracha and Meier (2014).

[TABLE 5 HERE]

[TABLE 6 HERE]

4. Discussion and Conclusion

More than three quarters of American households hold consumer debt (Brown et al., 2011) and traditional classroom-based approaches to financial education show mixed results at best (Collins 2013; Hastings, Madrian, and Skimmyhorn 2012). Tips, rules, nudges, and reminders are likely to become more prevalent as consumers move to managing more of their financial lives on mobile platforms. Lenders are already making ample use of these strategies, and so too are personal financial management platforms (like Mint.com).

Although the use of these approaches appears to be increasing, the evidence base surrounding them is quite limited. Rules of thumb may be a way to more effectively and efficiently provide financial education and improve financial outcomes, since they are inexpensive to deliver and target behavior change rather than increases in knowledge alone.

In the first rigorous test of a rules of thumb–based approach to financial education on consumer behavior and outcomes, we find that rules of thumb hold promise as a cost-effective method of financial improvement and behavior change. Our first rule of thumb helped to lower participants' credit card debt by $104 or about 2 percent from their baseline average and the other showed a negative but insignificant point estimate. And, the cost of delivering the rules was only about $0.50 per person, with the marginal cost of adding an additional recipient essentially zero. Additionally, the participants in this sample did not use their Arizona Federal credit cards very frequently before

(or during) the intervention, likely attenuating the affects. We were also unable to observe whether participants actually read and applied the rules, making it possible to only calculate intent to treat estimates, which are attenuated by non-participation, rather than treatment on the treated estimates.

However, the mechanisms underlying the reduction in debt are unclear. The first rule was meant to reduce participants' number of purchases under $20, but we do not find this to have occurred. We do not find a statistically significant reduction in total number or amount of purchases nor an increase in number or amount of payments, which could be another mechanism through which the rules work. Therefore, we speculate that the effect is caused by some combination of mechanisms (reductions in purchases and increases in payments), which reduces our ability to precisely estimate an effect on either one individually—though we cannot confirm this. Additionally, neither rule has a significant impact on overall revolving debt, which makes it impossible to rule out debt reshuffling. However, the point estimates on aggregate debt are consistently negative and there is no other evidence to suggest that reshuffling is occurring.

This lack of effect found on the underlying mechanisms for the debt reduction may mean that the rules serve more as a reminder or general awareness boost rather than as a true rule of thumb where the participant undertakes the suggested behavior change–in this case to make fewer purchases under $20. However, even if the participant is not adopting the suggested behavior change, it may still be beneficial to have the rule worded as such. This is supported by the finding that the rule that was worded with more of a direct action or command was more successful ("Don't swipe the small stuff") than the rule worded more as a reminder ("Credit keeps charging"). More research comparing rules of thumb to messaging reminders is needed to confirm or reject this hypothesis.

This study is a test of rules of thumb in concept, rather than a thorough test of rule framing, wording, and topics. The rules we create and test serve only as examples of potential rules, rather than as a conclusive analysis of the best rules to use. We select two rules based on what in-depth consumer interviews and previous literature suggest might have the greatest effect. Future research should test more rules on varying populations to determine which ones work best for whom.

Although it would be difficult for rules of thumb to fully replace other types of financial education and capability supports like financial coaching (Theodos, Stacy, and Daniels 2018), they could be used as complements to these approaches. Additionally, such financial education methods will only work if they are supported by a regulatory environment that protects consumers from predatory practices. Overall, given their low marginal cost of implementation, rules of thumb provide a promising method of delivering financial education and improving financial health.

REFERENCES

Agarwal, Sumit, Souphala Chomsisengphet Neale Mahoney Johannes Stroebel. (2015). Regulating Consumer Financial Products: Evidence from Credit Cards. *The Quarterly Journal of Economics*, 130 (1), 111–164.

Andre, Malin, Lars Borgquist, Mats Foldevi, and Sigvard Molstad. "Asking for 'Rules of Thumb': A Way to Discover Tacit Knowledge in General Practice." *Family Practice* 19, No. 6 (2002): 617-622.

Barnard, Christopher J. and C.A.J. Brown. "Prey Size Selection and Competition in the Common Shrew (*Sorex Araneus*)." *Behavioral Ecology and Sociobiology* 8 (1981): 239.

Baumol, William J., and Richard E. Quandt. (1964). Rules of Thumb and Optimally Imperfect Decisions. *American Economic Review* 54 (2): 24.

Bergelson, Joy M. "A Mechanistic Interpretation of Prey Selection by *Anax Junius* Larvae (Odonata: Aeschnidae)." *Ecology* 66 (1985): 1699-1705.

Blaustein, Andrew R. and Richard K. O'Hara "Kin Recognition in Rana Cascadae Tadpoles: Maternal and Paternal Effects." *Animal Behavior* 30 (1982): 1151.

Bouskila, Amos and Daniel T. Blumstein. "Rules of Thumb for Predation Hazard Assessment: Predictions from a Dynamic Model." *American Naturalist* 139 (1992): 161-176.

Bracha, Anat, and Stephan Meier. (2014). Nudging Credit Scores in the Field: The Effect of Text Reminders on Creditworthiness in the United States. Working Papers 2015 Series No. 15-2.

Brown, Meta, Andrew Haughwout, Donghoon Lee, and Wilbert van der Klaauw. 2011. Do We Know What We Owe? A Comparison of Borrower- and Lender-Reported Consumer Debt. Federal Reserve Bank of New York Staff Report no. 523.

Canner, Glenn B., and Anthony W. Cyrnak. 1985. "Recent Development in Credit Card Holding and Use Patterns among U.S. Families." *Journal of Retail Banking* 8 (3): 63–74.

"CARD Act Report: A Review of the Impact of the CARD Act on the Consumer Credit Card Market." 2013. Washington, DC: Consumer Financial Protection Bureau. http://files.consumerfinance.gov/f/201309_cfpb_card-act-report.pdf.

Chatterjee, Promothesh and Randall L. Rose. "Do Payment Mechanisms Change the Way Consumers Perceive Products?" *Journal of Consumer Research* 38, (2012): 1129-1139.

Choi, Heesun N., and Sharon A. Devaney. 1995. "Determinants of Banks and Retail Credit Card Use." Consumer Interests Annual 41:148–54.

Collins, J. Michael. (2013). The impacts of mandatory financial education: Evidence from a randomized field study. *Journal of Economic Behavior & Organization* 95: 146-158.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. (2014). Keeping It Simple: Financial Literacy and Rules of Thumb. *American Economic Journal: Applied Economics* 6 (2): 1–31.

Etzioni, Amitai. (1987). On Thoughtless Rationality (Rules-of-Thumb). KYKLOS 40 (4): 496–514.

Feinberg, Richard A. "Credit Cards as Spending Facilitating Stimuli: A Conditioning Interpretation." *Journal of Consumer Research* 12 (1986): 248-356.

Fishcher, Greg, Alejandro Drexler, and Antoinette Shoar. "Keep it Simple: Financial Literacy and Rules of Thumb." *CEPR Development Economics Workshop.* Last modified October 8-9, 2010. http://dev3.cepr.org/meets/wkcn/7/784/papers/FischerFinal.pdf

Frank, Robert H. (1987). Shrewdly Irrational. Sociological Forum 2 (1): 21–41.

Hamilton, R., and M. Khan. 2001. "Revolving Credit Card Holders: Who Are They and How Can They Be Identified?" *Service Industries Journal* 21 (3): 37–48.

Hastings, Justine S., Brigitte C. Madrian, and William L. Skimmyhorn. (2013). Financial Literacy, Financial Education, and Economic Outcomes. *Annual Review of Economics* 5: 367.

Hirschman, Elizabeth C. "The Relationship of Credit Card Acceptance to Retail Purchasing." *Journal of Consumer Research* 6 (1979): 58-66.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. (2010). Getting to the Top of Mind: How Reminders Increase Saving. NBER Working Paper No. 16205. http://www.nber.org/papers/w16205.

Karlan, Dean, Martín Valdivia. (2011). Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions. *Review of Economics and Statistics*, May 2011, 93(2): 510–527.

Kast, Felipe, Stephan Meier, and Dina Pomeranz. (2010). Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device. NBER Working Paper No. 18417.

Keys, Benjamin J and Jialan Wang. (2016). Minimum Payments and Debt Paydown in Consumer Credit Cards. NBER Working Paper No. 22742. http://www.nber.org/papers/w22742.

Kling JR, Liebman JB, Katz LF. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica* 75 (1): 83-119.

Skimmyhorn, William L., Brian Mitchell, Evan R. Davies, and David Mun. (2015). Assessing Financial Education Methods: Principles vs. Rules-of-Thumb Approaches.

Prelec, Drazen and George Loewenstein. "The Red and the Black: Mental Accounting of Savings and Debt." *Marketing Science* 17, no.1 (1998): 4–28.

Prelec, Drazen and Duncan Simester (2001), "Always Leave Home without It: A Further Investigation of the Credit-Card Effect on Willingness to Pay." *Marketing Letters* 12, no. 1 (2001): 5–12.

Pyke, Graham H., "Optimal Foraging in Hummingbirds: Testing the Marginal Value Theorem." *American Zoologist* 18, (1978): 739.

Soman, Dilip. "Effects of Payment Mechanisms on Spending Behavior: The Role of Rehearsal and Immediacy of Payments." *Journal of Consumer Research* 27 (2001): 460-474.

Soman, Dilip and Amar Cheema. "The Effect of Credit on Spending Decisions: The Role of the Credit Limit and Credibility." *Marketing Science* 21 (2002): 32-53.

Stamps, Judy A. "Conspecific Attraction and Aggregation in Territorial Species." *American Naturalist* 131 (1988): 329.

Stango, Victor, and Jonathan Zinman. (2011). Fuzzy Math, Disclosure Regulation, and Market
Outcomes: Evidence from Truth-in-Lending Reform. *Review of Financial Studies* 24 (2):
506–34.

Steidle, R. P. 1994. "Determinants of Bank and Retail Credit Card Revolvers: An Application
Using the Life-Cycle Income Hypothesis." *Consumer Interests Annual* 40:170–77.

Theodos, Brett, Christina Stacy, and Rebecca Daniels. 2018. "Client led coaching: A random
assignment evaluation of the impacts of financial coaching programs." *Journal of Economic
Behavior & Organization*, 155 (2018) 140–158.

Traub, Amy, and Catherine Ruetschlin. 2012. "The Plastic Safety Net: 2012." New York:
Demos. http://www.demos.org/publication/plastic-safety-net.

Wasberg, Cynthia A., Tahira K. Hira, and Alyce M. Fanslow. 1992. "Credit Card Usage and
Consumer Debt Burden of Household." *Journal of Consumer Studies and Home Economics*
16 (1): 19–32.