Are Information Disclosures Effective? Evidence from the Credit Card Market*

Enrique Seira† — *ITAM/J-PAL*

Alan Elizondo — *Banco de México*

Eduardo Laguna-Müggenburg — *ITAM*

The recent financial crisis and the advent of behavioral economics have placed renewed focus on consumer protection in the financial sector. The US recently created the Consumer Protection Bureau and mandated new information disclosures in the Credit Card Accountability Responsibility and Disclosure Act (known as the Credit CARD Act) of 2009.[1] Many countries –including Mexico– have followed suit, requiring financial institutions to report more information. However, important questions remain unanswered; for

*We are grateful to the bank involved in this study and to the Comision Nacional Bancaria y de Valores for the support to this project. We thank Marianne Betrand, Yan Chen, Alejandro Ponce and Radovan Vadovic for initial comments, and Cesar Martinelli, Nicholas Melissas, Andrei Gomberg, and Emilio Gutierrez, and seminar participants at ITAM for comments on this draft. Edith Félix, Karen Olivo, Alonso De Gortari provided outstanding research assistance. All errors remain our own.

†Corresponding author: enrique.seira@itam.mx

[1]In the US, Title X of the Dodd-Frank Act created the Consumer Protection Bureau, the agency responsible for promoting fairness and transparency for financial products and services. In 2009, the Obama Administration produced the Credit CARD Act which, among various requirements, mandates more information disclosure, in monthly statements in particular. Among a wide range of disclosures, it requires card companies to specify the time it would take to pay off existing debt if only the minimum required payment is made. The publication of APRs and interest rates has been a requirement since 1968. The Office of Information and Regulatory Affairs (OIRA) recently emphasized the importance of 'smart' information disclosure (see "Disclosure and Simplification as Regulatory Tools" (June 18, 2010), downloaded in May 2012 from http://www.whitehouse.gov/sites/default/files/omb/assets/inforeg/disclosure_principles.pdf)

instance, how effective these disclosures are, which are better, and for which types of consumers.[2]

In order to be able to talk about effectiveness, one needs to understand what these disclosure laws are trying to achieve. The precursor to these laws, the Truth in Lending Act (TILA) of 1968, was motivated by a desire to standardize how the price of a loan was quoted. It was thought that facilitating price comparisons would "protect the consumer against inaccurate and unfair credit billing and *credit card* practices..." "[enhance] economic stabilization [i.e. reduce risk]... by the informed use of credit," as well as "[strengthen] competition among the various financial institutions."[3]

But is there evidence that contract terms awareness is low and that mandating information disclosure could potentially help? Our read of the evidence is that there is some promise for disclosure mandates, although few rigorous studies actually measure their effects, especially in a developing country setting. Regarding awareness and understanding of mortgage terms, Lacko and Pappalardo (2007) find that mortgage cost disclosures in the US failed to convey key mortgage costs to many consumers. In a survey we conducted in Mexico for a subset of this paper's sample –described below– only 3 percent of cardholders claim to know the exact interest rate on their card.

This low cost awareness may impact consumer shopping behavior. Stango and Zinman (2011a) find that similar consumers pay substantially different interest rates for credit cards. Hall and Woodward (2013) show that borrowers sacrifice at least $1,000 by shopping from too few mortgage brokers. Ponce et al. (2012) find that Mexican cardholder's debt allocation is insensitive to

[2]Surprisingly, these are still unsolved questions. Referring to Truth-in-Lending-Act (TILA) disclosures, Durkin and Elliehausen (2011) recently wrote: "The degree to which such disclosures can protect consumers is still a matter of debate, and it deserves careful consideration." On the one hand Julie Willams, acting Comptroller of the Currency in 2005, believed financial disclosure policy had not worked well for consumers and had unnecessarily burdened banks (see her 2005 speech "Remarks before Women in Housing and Finance and The Exchequer Club"). In contrast, former Chairman of the Federal Reserve System's Committee on Supervision and Regulation of Banking Institutions, R. Krozner, was optimistic on their effectiveness (see his 2007 speech at George Washington University).

[3]Quoted from the first paragraph of the TILA. The use of italics is the authors' own to emphasize that credit cards were the only type of credit mentioned explicitly in the Act's stated purpose.

differences in the interest rates of the cards they already hold. These studies suggest that high search or attention cost may be at play. There is also some mild evidence of over-borrowing that could be mitigated by better disclosures. Melzer (2011) shows that improving credit access for some low-income households inhibits their ability to pay important bills. In a similar vein, White (2007) cites that in the Panel Study of Income Dynamics the most common reason that households gave for filing for bankruptcy was "high debt/misuse of credit cards" (33 percent).

This paper seeks to provide a rigorous answer to two questions. First, do TILA-type disclosures have an effect on credit card risk, indebtedness, and switching? How large are these effects? Second, are nonstandard disclosures, such as warnings and social comparisons, more effective at inducing changes in behavior? The answers to these questions are important since disclosure requirements are often mandated and amended without evidence, imposing costs on financial institutions and diverting the attention of policymakers. Mandating disclosures may be a way for politicians to avoid conducting more rigorous analysis of market failures and making hard choices.

We were particularly interested in running a horse race between TILA-type disclosures and more innovative disclosures, such as warnings and peer comparisons. The Mexican Banking Commission (CNBV) actually encouraged us to pursue this agenda, since it was contemplating sending personalized warnings triggered by risky consumer behavior; they acknowledged that rigorous evidence was needed. The CNBV was particularly concerned with a fraction of consumers that seemed highly indebted and on risk of default. We thus focus on this population.

To conduct the experiment the CNBV paired us with a large Mexican bank. Together with the bank's marketing department we designed seven messages. The first two were inspired by laws such as the TILA: one included a personalized interest rate and the other a measure of time to pay, reported as the number of months it would take the client to pay off his or her debt if he/she paid only the minimum amount due (MTP). Both of these disclosures feature prominently in the US disclosure mandates, the latter being added in the Credit CARD Act of 2009.

A second set of messages, not present in TILA-type regulations, were inspired by the psychology literature on peer comparisons[4] and by a recent paper in economics by Chen et al. (2010) showing that individuals care about being below or above average in terms of performance on a task. To this aim we designed four "social comparison" messages, two of which inform the client that his or her credit card debt is above the mean for similar clients, one of these provides broad advice while the other does not. The other two comparison messages told the consumer whether he or she has a high risk or a low risk of default, respectively, compared to similar clients. A final nonconventional message gave an explicit warning against overconfidence in paying down debt. It was inspired by current labeling on food, tobacco, and drug products, but also by Ausubel (1991)'s conjecture of the existence of a large fraction of overconfident consumers. Initially, the bank's personnel believed that none of the messages would have any impact.

To test client responses to these messages, we implemented an extensive blind field experiment in conjunction with the bank. We selected a random sample of 167,190 clients whose credit card payments were not more than 89 days past due in September 2010 and belonged to the upper tercile of the bank's risk distribution.[5] These costumers turned out to be highly indebted, unaware of their interest rate, and overconfident as to their ability to pay down their existing debt as we document below. As will be described later in the paper, the messages were randomized and sent to treatment/control groups so that causal inference could be made more straightforwardly. The messages came in an envelope that was indistinguishable from a monthly statement, but instead of containing a monthly statement it only had our message.

We measured the effect of these messages on four prominent outcomes TILA intended to affect: (a) interest paying debt, (b) delinquency, measured as a dummy variable that turns on if the client has an overdue payment of 30, 60

[4]For an early account see Festinger (1954) and the vast literature after this paper.

[5]Mexican Law mandates that banks hold reserves as a function of the probability of default based on an official formula described later in the paper. The upper tercile was calculated using this official formula.

or 90 days, (c) voluntary account closures by the client, and (d) openings of new card accounts in other banks (a proxy for switching). It is possible that messages have effects beyond the particular card they refer to. For instance they may reduce risk taking by the consumer over all, or they may strategically lead to less/more default in other cards. We approximate market perceived risk by (e) the credit score in the Credit Bureau, and observe (f) default in other cards also using Credit Bureau information.

In light of the strong policy emphasis on disclosures, some of the findings are surprising. We find that even when disclosed saliently, the interest rate *does not* change levels of debt, delinquency, or account closing/switching. This zero effect is quite precise and robust across subsamples.[6] This result is particularly striking given the low awareness environment among our sample. The other TILA type message –the MTP– actually *increased* delinquency by 6 percent of the mean for the population that pays interest regularly. This was also unexpected.

On the positive side, we find that non-TILA messages *are* effective, even when the information provided is quite coarse. The "high risk" message was particularly useful in decreasing delinquency, with an effect of 8.1 percent on mean delinquency, and caused closures by clients. The "low risk" message actually *increased* delinquency by 7.6 percent. The "high debt" peer comparison had no effects on debt, and a small barely significant decrease in account closings. We found no evidence that providing a call to specific actions in the form of general advice increases the efficacy of the message. Finally, the warning message reduced debt by about 0.7 percent but had no incidence on delinquency. When present effects were short-lived, lasting only one or two months. We show that the failure to detect effects is not due to low statistical power.

From our results, we believe that information disclosures in this market are unlikely to induce any significant change in competition, indebtedness or risk, contrary to the expectations of proponents of TILA, unless these are made easier to understand and actionable. Nonetheless, even small effects may be

[6] With one exception discussed below.

worthwhile given that sending messages is very cheap. In fact, messages are cost effective even if we only consider the effects on a bank's balance sheets for one month: the cost of sending a message is around 2.5 pesos, whereas the benefit to the bank is in the order of 245 pesos in expected loss reductions. The bank's personnel were surprised at these results and said they will begin sending the effective messages to their population at risk of default. We do not evaluate the effect of this information on consumer welfare but, given the smallness of the responses induced, it is likely to be small.

There is a growing literature that studies the effect of information provision on choices in many settings. For example, Jensen (2010) provides information on the returns to a college degree; Hastings and Weinstein (2008) on school test scores; Jin and Leslie (2003) on restaurant hygiene grades; Bollinger et al. (2011) on food calories, and Bertrand and Morse (2011) on payday lending. Surprisingly, few studies have been carried out on disclosures in the credit card market, despite the fact that credit cards command considerable attention in policy circles and were mentioned explicitly in TILA.

Most studies on credit card information disclosure look at the effect of different mock statements on an individual's understanding or awareness, rather than on their actual behavior.[7] Ferman (2011) and Stango and Zinman (2011b) are two important exceptions. Ferman (2011) randomizes interest rate disclosure and the interest rate itself in credit card fixed-repayment plan offers in Brazil. In line with our results, he finds small effects of information on payment plan take-up and take-up/interest elasticities. Stango and Zinman (2011b) study the effects of the TILA itself on interest rates. Overall, they find no effects on average interest rates, but do find lower interest rates for borrowers who typically underestimate APRs.[8] Another strand of literature

[7] See for example Soll et al. (forthcoming). Early studies: Shay and Schober (1973), Day and Brandt (1974), and Durkin (1975) look at the effect of the TILA itself on awareness, though do not include a control group.

[8] Using a triple-difference design around a 1981 regulatory change that decreased TILA enforcement for financial companies as compared to banks, Stango and Zinman (2011b) find that, after the regulatory change, borrowers who underestimate APRs are more likely to pay more on instalment loans taken out with financial companies than those who taken out with banks.

studies consumer responses to information that is not standard in TILA disclosures. However, neither of these studies the credit card market. One highlight is Bertrand and Morse (2011) which studies payday loans in the US and shows that providing APR comparisons has no effect on subsequent borrowing –in line with our results– but that providing information on cumulative dollar cost does reduce by 5.4 percentage points the likelihood of payday borrowing in subsequent cycles.[9]

We were able to capitalize on the above literature and, in several respects, go further. First, by having a unified multi-arm experiment, we were able to run an internally valid horse race between TILA-type messages and other types of information. Second, we complement Bertrand and Morse (2011) and Ferman (2011) by focusing on the intensive margin, whereas they focus on loan take-up. This is relevant since large segments of the population already have credit cards and laws such as the TILA emphasize disclosures in monthly statements (i.e., for people that already have cards). Third, we study how information affects outcomes such as indebtedness, default, voluntary account closures and switching, all of which have been unstudied in the literature and are closely related to the outcomes that TILA-like regulations seek to affect. Fourth, we examine the credit card market, an extensive market that has been blamed for the increases in US bankruptcy filings in the late 90s (White (2007)) and one that has been a primary focus of TILA disclosures. Fifth, we conduct an extensive external validity exercise including several banks that together have more than 70 percent of the credit card market in Mexico.

[9]In a public good context relating to movie ratings online, Chen et al. (2010) show that telling users that they rate fewer movies than the average user leads to a fivefold increase in the number of movies they rate subsequently. Allcott (2011) finds that, in the US, sending letters comparing the energy use of households to that of similar neighbors leads to a fall of 1.1 percent to 2.8 percent in energy consumption.

I. Context and Data

Before we describe the field experiment, we outline the context in which it was carried out. The CNBV was thinking about issuing rules that mandated banks to send personalized warnings based on the risk profile of clients, and contacted us with a partner bank that was seeking ways to reduce delinquency in their credit card portfolio. They agreed to work with us to design, send information messages, and measure their impact, but only for their riskier clients. We accepted to work with this population since it is precisely them that regulatory authorities were concerned about, and since we had a strong prior that they were overborrowing and a clear hypothesis that the messages we designed could decrease debt.

We focused on such population by drawing a random sample from the upper tercile of the risk distribution of clients, were risk is measured using the CNBV methodology to predict probability of default. This methodology assigns a predicted probability of default (PD) in the next 12 months to each credit card based on card use according to a logistic function with five regressors: number of consecutive months delinquent (CD), number of total months delinquent in the last six months (D), tenure of the card (T), last month's payment as a proportion of the minimum payment due (MP), and last month's percentage of credit line used (LU).[10] When drawn in September 2010 the PD ranged from 9 percent to 100 percent with a mean of 26 percent. As described in Section II, a random subset of this sample received messages in February 2011.

Obviously this sample is not representative of the bank's entire clientele, but we argue that it is the type of population at which consumer protection laws are directed. Indeed, the patterns we found in the two surveys we implemented in this sample display low contract-term awareness and high indebtedness. Section V reports additional results for representative populations based on samples from two other large banks.

[10] The exact formula is given by $\frac{1}{1+exp-(2.9704+0.6730CD+0.4696D-0.0075T-1.0217MP+1-1513LU)}$. The model fit and predictive power are high: ROC curves of 86 percent and prediction error in backtesting of less than 5 percent of mean default.

A. Administrative Data

The data available to us consists of monthly information on the variables that appeared in the monthly statements for the selected 167,190 credit cards in the period from September 2010 to June 2011. These variables include interest-paying debt, account closings by clients, delinquency (30, 60 or 90 days overdue indicators), payments, purchases, interest rate, credit limit, fees, etc. We were also able to obtain data on credit score, default status and opening dates for all the loans of a random sample of 17,815 our clients from the Credit Bureau. This enables us to look for induced behavioral changes in other loans, a kind of spillover effect of information. We have limited demographic information. We used administrative data to follow client behavior as this has the virtue of containing virtually no measurement error and is cheap to collect.

Table 1 shows that indeed clients are highly leveraged and risky. It also shows that interest rates are high and clients somewhat new. Average interest-paying debt was around 18,000 pesos while mean income –according to our survey implemented in this sample– was close to 9,000 pesos, so clients seemed highly leveraged. Mean card utilization was 70 percent of the credit line. Clients were also risky: the estimated ex-ante probability of their defaulting in the subsequent 12 months using the CNBV formula was 26 percent, and default (more than 90 days past due) was already 9 percent by February 2011. Figure 1 plots a histogram of this probability measured in September 2010, when the sample was taken, and in January 2011, just before the messages were sent.[11] The expected loss per account as calculated by the bank proprietary formula was 2,721 pesos on average in September 2010, which explains why the bank wanted to induce a behavioral change in these clients.

We created a dummy variable called "Delinquent" which takes the value of one in a month if there are payments that are 30, 60 or 90 days overdue. This is our main measure of ex-post risk realization. In an average month 11 percent of accounts are classified as delinquent by this standard. Clients do not close

[11] We wish to emphasize that we are covering a broad domain of default probability and that this study is not just about clients with an extreme likelihood of default. In fact, some clients had zero default probability in January 2011.

their accounts often. Only 2.6 percent of accounts had been voluntarily closed by the client five months after the sample was selected. Counting also the closing of accounts by the bank the number increases to 4.4 percent. By April 2011 we have an attrition rate of just above 9 percent but we do not think it is a serious problem to our analysis since it is balanced across treatment and control groups (see Figure 1 and Table 2 in the Online Appendix).[12]

Interest rates are high and stable at 44 percent per year. In fact, our bank is persistently among the top 5 in terms of highest interest rate charged. There was little variation in interest rates over time: only for 4 percent of observations did the interest rate change by more than 0.4 monthly percentage points (5 pp in terms of yearly rates) from one month to the next. This means that being informed as regards interest rates and trying to obtain a lower rate may pay off. Moreover, the interest rate was not highly correlated with risk, which means that clients with similar risk profiles were paying very different interest rates even within this same bank.[13] The average number of months to pay current debt with no further purchases and making the minimum payment due was 27. As we will show below, this quantity is much greater than what the people in our sample expected.

B. Survey Data

To get an understanding of the context that would help formulate hypothesis, we collected two round of surveys for random subsamples of clients in our administrative data. The first random sample of 800 clients was collected by phone in December 2010, before sending the messages. Its main purpose was to help us understand how indebted clients were with respect to their income, how satisfied with the way information is reported in their monthly statement, their knowledge of the interest rate of their card and their expectations on MTP. We also asked whether they read the monthly statement and

[12] When we regress a dummy for attrition vs. treatment dummies and strata dummies, we cannot reject the joint hypothesis that the treatment dummies are equal to zero.

[13] A regression of interest rates against deciles of the internal probability of default and months dummies yielded an R-square of 0.01.

elicited their predictions as regards their paying down their debt in next 1, 2 and 3 months. The questions and mean responses are tabulated in Table 7 at the Online Appendix.

We wish to highlight three lessons learned from this first survey. First, these clients were uninformed as regards their interest rate: only 3 percent of clients claimed to know their exact interest rate and 34 percent claimed to know it approximately within 5pp. This happens in spite of Mexican Law mandating interest rate disclosure in the bank statements.[14] Second, clients underestimated the number of months needed to pay off their debt, on average believing it to be 13 months rather than the 27 months indicated by the real data. This overconfidence is also captured more directly: in response to a direct question (see Online Appendix) 35 percent of the consumers surveyed claimed to have overestimated their ability to pay down their debt in the previous six months.[15] Third, 92 percent of clients claimed to read their monthly statement carefully, which is somewhat surprising given how uninformed they were as regards interest rates. Part of the explanation may be that the information is hard to read. For instance, 42 percent said they would prefer a clearer statement and 38 percent claimed default happens because people do not realize how fast they are accumulating debt, and not because of strategic default or due to unforseen shocks. Overall, a substantial proportion of the clients were unaware of interest rates, unsatisfied with the clarity of their monthly statement, and displayed signs of overconfidence regarding their paying down their debt.

We also implemented an ex-post survey of about 2,300 clients in October 2011 with the objective of testing whether there seemed to be different attitudes as a function of the messages, as well as to have an idea of how dissatisfied

[14]See an example of a monthly statement in the Online Appendix.

[15]Ausubel (1991) conjectured that people care little about interest rates in the credit card market because they wrongly believe they will not incur any interest. To test his conjecture we asked clients to make a prediction on whether or not they would have more or less debt in the following two months compared to their present debt and verified whether the prediction was ex-post correct using the administrative data. Around half got it wrong; interestingly, about 3/4 of these erred on the side of overconfidence, thinking that their debt would decrease when in fact it increased, lending support to Ausubel.

they are with their level of indebtedness and the net benefit of defaulting.[16] Over 4/5 said they would like to decrease their debt even taking into account the sacrifices this would imply, and over 9/10 said that defaulting would decrease their welfare taking the benefits of defaulting into account. Unfortunately, we do not have enough power to detect impacts of different messages (see Figure 2 in the Online Appendix shows power calculations and Table 8 in the Online Appendix shows regression output estimating the effects).

II. Experiment Design and Model Specification

A. Experiment Design

The aim of the field experiment was first to test whether information and warning messages indeed induced a change in behavior, and secondly to test which message was more effective in inducing behavioral change. We compared TILA-type disclosures to more innovative disclosures, such as warnings and peer comparisons. As we have previously stated, we teamed up with a bank to design and send seven messages.[17]

The first two messages were inspired by the disclosures that are typically mandated by laws such as the TILA. In particular, we sent a message disclosing the personalized interest rate very saliently, and a second message displaying the number of months it would take a consumer to pay off his or her debt if making only the minimum payment due without further purchases. Let us call these messages the "interest rate message" and the "months-to-pay message" (MTP), respectively. Both of these disclosures feature prominently in the US law.

The second set of messages was inspired by literature that stresses that people are influenced by what their peers do, either through the signal that this

[16]The respondents were distributed among the control group (25.17 %) and those in the High Debt plus Advice (14.99 %), Months to Pay (15.94 %), Interest Rate (25.52 %) and Warning (18.38 %) groups.

[17]Working with the bank offered the advantage that it enabled us to use its experience in marketing, though we had to adhere to the bank's communication protocols.

behavior provides, such as in rational herding models, or through a conformity channel. We designed four of these peer-comparison messages: two of these informed the client as to whether his or her card's debt was above the mean of clients of the same gender, similar credit limit (as a proxy for income), age, and risk: their "peer group". These two "high debt" messages differed only in terms of whether broad advice was provided or not. The other two comparison messages told the client their *relative* risk of default, again compared to similar clients. One message was sent to clients with a high probability of default warning them about it, and the other message to clients with a low probability of default as a congratulatory note; we call these "high risk" and "low risk" messages respectively.[18]

Finally, we included a message which did not contain any direct comparison but rather an explicit warning against overconfidence in paying down debt. We call this the "warning message". We thought this message was interesting because the clients in our survey seemed overconfident as regards paying down their debt and because these types of warning messages are common in health disclosures (e.g., "smoking kills") but have been understudied. This message could increase attention even when no hard information is provided. In a recent paper, Stango and Zinman (2013) show that surveying people about their card overdraft fees seems to cause them to pay less in fees, even when the survey does not contain much information in this regard. They interpret this as evidence of inattention.

Figures 3, 4 and 5 show some of these messages, the rest can be found at the Online Appendix. At this point, we wish to highlight two facts: first, with the help of the marketing department at the bank we made these messages very salient. Font sizes of around 50 points were used for the relevant amounts and the language used was as simple as bank communication protocols would allow. We believe the salience of the messages is an upper bound on the salience that TILA-like laws typically mandate. Second, unlike the "interest

[18]This later message was not part of our initial design, but our partner bank suggested we include it. Some of these clients indeed had default probabilities close to zero but others had default probabilities above 8 percent. But in each case the clients that received the "low risk" message had low *relative* default probabilities, i.e. below their strata median.

rate" and "months-to-pay" messages, the peer comparisons were coarse in the sense that they were not tailored to particular individuals. We could have told each individual exactly where he or she was in the distribution of risk for example. We did not do this as it was simpler for the bank; but as we shall see we still found some impacts for these messages. We discuss our hypothesized effects for these messages in subsection B below.

The allocation of messages to clients was random within their strata; thus we have treatment and control clients within strata. The randomization design was done as follows: since some messages involved comparisons among "similar" clients, we had to create an operational definition of what it meant to be similar. To this end, we stratified the sample into cells by crossing four variables –gender, quintiles of age, quintiles of credit limit, and terciles of predicted default probability– to produce 150 cells in total. Next, within each cell we identified clients who had debt that was above the cell mean. These were candidates for receiving the "high debt" message. When we take into account the high-debt stratification, we effectively have 300 strata. Clients within a stratum constitute a peer group. Randomization into some treatment (77,175 messages) vs control (90,015 no-messages) was performed within each stratum to provide us with an appropriate control group.

We started by allocating the "high debt" message to 85 clients in each of the 150 high debt strata for each of the two high debt messages.[19] This meant we had 12,825 clients for each of these messages.[20] Next within each of the 300 strata we identified which unassigned clients had above (below) the median predicted probability of default and randomly allocated these to the "high (low) risk" message: 6,444 to the high risk message and 6,456 to the low risk one. The remaining unassigned treatment clients were randomly allocated within strata as follows: 12,900 for the interest rate and the debiasing warning message, respectively, and 12,825 for the months-to-pay message, so these

[19]Some cells randomly included 86 clients rather than 85 to be able to distribute all the sample.

[20]Note that this does not exhaust all the high-debt clients in the sample but does leave fewer high-debt clients for the remaining messages and control group. To take this into account, all regressions included a high-debt strata dummy.

three later groups are directly comparable. Table 1 in the Online Appendix shows that randomization worked to balance the variables across groups.

The timing of the experiment (shown in Figure 6) was as follows: the selection of the sample and the randomization into treatments was carried out in September 2010. The baseline survey was carried out in December 2010. Messages were printed and sent out in February 2011, using administrative information from January 2011 for personalized messages. From the outside, the envelope was indistinguishable from a monthly statement, but inside it contained only our message, no monthly statement came with it. We were told that the delivery service of the bank is of a very high quality and that more than 95 percent of the clients should have received the message. Although we cannot be sure that they did actually read the message, in the survey 78.3 percent claimed to read their monthly statements and 67 percent claimed to read them very carefully.

B. Model Specification

Since messages were conditionally randomized, we can estimate the average causal effect from the difference in conditional means. We estimate the average treatment effect of message T_j on outcome variable Y in month t by estimating equation (1). Since the sample size is large we decided to estimate specifications separately for each month t as reflected in equation (1) although we did pull all treatments in the same equation.

$$Y_{ijt} = \alpha_t + \sum_{j=1}^{7} \beta_{tj} T_{ij} + S_{ik} + \epsilon_{ijt} \tag{1}$$

α_t estimates the mean on the control group of the respective message in month t and β_{tj} is the average treatment effect of message j in month t, while S_k are stratification indicators for the k strata described above. We also estimated two related models: one regression for each treatment separately, and a differences-in-differences specification. The results were similar and are not reported here.

The main outcome variables are interest-paying debt, a delinquency dummy variable, and a dummy for the client closing the credit card, described in the data section. We focus on these variables because we believe they are important on their own, but also because they are close to the outcomes that the TILA sought to influence. Closing the account is a proxy for switching[21] while changes in debt are a proxy for demand responsiveness, both related to competition. Delinquency is related to the stability that the TILA mentions.

We also report results for a specification that pulls both TILA-like messages in one dummy and the five non-TILA messages in another dummy as in equation 2. This may afford more statistical power under the assumption that effects push in the same direction, and also enables us to test the null hypothesis of equality between TILA and non-TILA messages, $\beta_{1t} = \beta_{2t}$ in equation 2.

$$Y_{ijt} = \alpha_t + \beta_{1t} TILA_{ij} + \beta_{2t} NONTILA_{ij} + S_{ik} + \epsilon_{ijt} \tag{2}$$

Given that the clients in our sample were highly indebted and at high risk of default, we hypothesized that the "high debt" message, the "high risk" message, and the warning message would reduce debt and default. We had no strong prior expectation as to their effect on account closures. Regarding the salient interest rate message, since this bank has one of the 5 highest interest rates in the market, we expected that revealing the interest rate would decrease debt and increase account closures. Finally, we expected that the MTP message would make clients realize they were underestimating their months to pay and induce larger payments, lower debt, and less delinquency on their part. We test these hypotheses in the next Section.

Before we proceed to the results, it is important to state that we are not necessarily estimating the effect of *reading* the information –since we do not know if the clients did actually read it– but we are estimating the effect of sending the information, which is what TILA-type laws mandate. Having said this, we believe that the information did reach a significant part of the

[21] Later in the text, for a subsample, we study opening of new credit card accounts, another proxy for switching.

sample: mail delivery accuracy is about 95 percent according to bank staff, and according to our survey 92 percent of clients claimed to read their monthly statement carefully. Note also that some messages did have an effect, so the messages did arrive.

III. The causal effects of TILA-type and non-TILA messages

A. Personalized Interest Rate Message

Probably the most prominent disclosure in TILA-type laws is the price of credit, as reflected in either the interest rate or the APR. In spite of its importance, to our knowledge there are no randomized control trials that measure the impact of increasing the salience of this information on the use of credit cards and their risk of default. In an interesting study, Malmendier and Lee (2011) found that online auction bidders pay, on average, prices above the posted price for the same good. They find, however, that the extent to which this happens is inversely related to the salience of the posted price. Chetty and Kroft (2004) find that tax elasticities are dependent on how salient taxes are. Low salience of the interest rate in monthly statements could rationalize the high indebtedness and risk in our population, as well as the low interest rate awareness we found in the survey. Presumably its salient disclosure could remedy this.

Before we proceed to the analysis of the effects of the message, it is useful to provide a brief discussion on how much money is potentially at stake. Ideally one would like to compare consumer's indirect utility with and without knowledge of interest rates. Unfortunately we have no way of doing this. Instead we report some statistics that suggest (although it doesn't demonstrate for sure) some clients may be leaving money on the table. First, note that the level of incurred interest is high. On average our clients paid 7,752 pesos a year in interest, which is more than half of their average monthly reported income. Second we conducted an exercise to measure if consumers debt *allocation* across the cards they have minimizes financing cost, assessing if consumers that have two cards with different interest rates in our cooperating bank actually allocated debt to the cheaper card when feasible.[22] For 44 percent of clients it is the case that more than half the time they could save on interest by reallocating debt from the expensive card to the cheaper one. Actual financing cost was 18 percent higher than the minimum feasible one. Third, the yearly interest rate on credit cards at our bank is almost 10pp higher than that of the cheapest of Mexico's five largest banks; hence the mean consumer could probably save around 1,800 pesos per year from having this debt in the cheapest of these five banks. So our individuals seem not only unaware of interest rates but could also potentially profit from knowing them. Our conjecture was that reporting saliently the personalized interest rate would lead to a decrease in debt and an associated decrease in delinquency, perhaps through the clients' substitution towards cheaper cards, or from just decreasing their total debt. The message could also cause an increase in voluntary closures as clients switch to cheaper cards.

To measure the impact of making the interest rate salient we sent the interest rate message displayed in Figure 3 (a) to a randomized treatment group of 12,825 clients, as described in Section II. We estimate its impact using the specification in equation (1) for the months of March and April 2011 separately. We do not show any results for the months of May and June 2011 in this paper as these were economically small and not different from zero for any of the variables or treatments. Each column in Table 2 represents an estimation of equation (1). Dependent variables are displayed in columns and treatment messages in rows.

From the first column fifth row, we can see that the effect of the interest rate message on debt in March is -35 pesos for March and 14 for April. This is a tiny 0.2 percent of mean debt and is not statistically different from zero. The effect on delinquency and account closures is also not different from zero and is economically minuscule. Given the low interest awareness in our sample

[22] Actually more than 3000 of our clients had two cards in our bank. On average, the yearly difference in interest rates across these cards was 4pp.

this zero-effect was unexpected.

We now test how robust this result is by estimating a regression similar to equation (1) but focusing only on the interest rate messages and using different subsamples. Results are reported in Table 3 Panel A. One potential explanation of the null effect is that stakes are low, and that we should find responses in subsamples with higher stakes. The first row of Table 3 considers clients that typically pay interest (revolved debt for 10 consecutive months before receiving the message). Rows 2 and 3 consider populations that have above-median interest and above-median debt, respectively. We again find a zero effect.

A second alternative is that responses are heterogeneous. Some clients may perceive the disclosed interest rate as "good news" leading them to increase debt, while other clients may perceive it as an expensive interest, leading to a decrease in debt. The effects may cancel out on average. To investigate this, we classify any change –either positive or negative– as a positive change by using the absolute value of the change in debt as a dependent variable, therefore avoiding such cancelations when averaging. Row 4 presents the results. Again the effect is statistically zero, even in with this strategy with is biased towards finding an effect. Section IV. formally models treatment effect heterogeneity allowing for a random coefficient in the treatment dummy and shows that the distribution of the response is tightly concentrated close to zero. Thus, the evidence does not support much heterogeneity in the response arising from differing ex-ante beliefs about interest rates.

A third possibility is that clients have limited possibilities to substitute debt across financial products. Row 5 uses only clients that had other credit cards in any other bank, hoping to find larger debt elasticities.[23] We do find a negative coefficient of -525 pesos in March, significant at the 1 percent confidence level. Although this is only 3 percent of mean debt, it suggests that ability to

[23]We now obtained this information from the Credit Bureau. A prior version of the paper used information from the bank as of the date of application which presumably is less accurate. Unfortunately, since the Credit Bureau does not have price information, we do not observe which card is more expensive. However as we mentioned before, our bank is in the top 5 in terms of highest credit card interest rates and it is likely that this card is more expensive.

substitute across cards may mediate the effects of information.

B. Months to pay outstanding debt

For many consumers, paying their card's debt is not an easy task and many pay close to their required minimum.[24] Given that such minimum payments are approximately 5 percent of debt, this implies that clients take a long time to pay. In our data, for 12.9 percent of clients, making the minimum payment due implies never paying off their debt, even if they make no further purchases. For the remaining observations, the mean number of months to pay (MTP) is 27 and the 99th percentile is 83 months.[25] Such long payment periods are worrisome since there is evidence that actual payment anchors on the minimum payments (e.g. Stewart (2009)). Figure 2 plots a histogram of actual MTP calculated from the administrative data vs MTP as reported by the clients themselves. This is a concern, clearly many clients are grossly underestimating the amount of months to pay off their debt.

Due to concerns such the one mentioned above, policymakers that enacted the Credit CARD Act (2009) required card companies to disclose the number of months consumers will take to pay off debt if they stop purchasing and only pay the minimum amount due. The logic was that giving consumers information on the time burden for paying their debt would make them more debt-conscious and lead to faster debt decreases. We expected that when consumers were made aware of their overconfidence, they would indeed decrease their debt and pay more.

To measure the impact of this disclosure, we randomly sent 12,900 messages with the personalized number of months to pay. The exact design is shown Figure 3 (b). It informed cardholders of their personalized MTP, explicitly advising them to pay more than the minimum amount due. The results are reported in Table 2. Contrary to our expectation, the message had no effect on

[24]In our sample, 6 percent of those clients that pay above the required minimum pay within 1 percent of the minimum and 20 percent pay within 10 percent of the minimum.

[25]The number of months to pay off the current debt balance if the minimum is paid and if no further purchases are made is given by the following formula $N = -\frac{ln(1-\frac{Debt*MonthlyInterestrate}{MinPay})}{ln(1+MonthlyInterestrate)}$.

debt on average. It did however cause delinquency to *increase* for those that often paid interest or had high debt (see Panel B of Table 3). For these clients, the effect was an increase in delinquency of 0.011 points in April, equivalent to 7.4 percent of its mean value. The "low risk" message increased delinquency mean value. The amount of payments actually went down by about 10 percent of its mean (unreported in the Table). This response was not what we expected and, we suspect, contrary to the aims of policymakers. Although the bank lost money from this, we have no way of telling how it affected consumer welfare.

Since the message does not affect income or debt, we believe that the delinquency it induces must be strategic in the sense of not being forced by circumstances. One interpretation of the increase in delinquency is that some clients are discouraged by finding out that there are still too many months in which to pay interest, which may seem unfair or unfeasible, and therefore decided to stop paying.

C. Peer-Group Comparisons

While TILA-type laws have concentrated on disclosing information on contract terms, we now measure the effect of messages directly related to cardholder's behavior and comparisons with the behavior of their similar others. Peer comparisons have been shown to be effective in many contexts, inducing participation in elections, encouraging contributions to online public goods, and increasing savings on energy. We test their potential for reducing debt. For instance, in a recent paper, Chen et al. (2010) showed that movie raters respond sharply to peer comparisons. When individuals were told that they rated fewer movies than the median rater, the number of movies they rated increased fivefold.

Inspired by these results, we sent a similar message informing clients if they had above-average debt (or were higher-risk). We conjectured that these messages would decrease debt and delinquency. The bank suggested that we also tested a message that congratulated the client for his or her low risk, displaying a thermometer indicating a "low risk" reading. [26]

[26] We are agnostic about which are the channels through which peer comparisons induce

Figures 4 and 5 display five of the messages sent, the other two can be found as Figures 7 and 8 in the Online Appendix. The main lines of the "High Debt + Advice" message said "...with respect to this group, your debt is HIGHER than the average of people similar to you." A footnote explained that the group was composed of people of a similar age, income, and the same gender, but no further details were provided. It then gave three broad pieces of advice: analyze your ability to pay, pay at least twice the minimum payment due, and decrease your debt. We also sent another message identical to this one except that we omitted the explicit advice in order to enable us to measure the effect of the advice per se. We measure the treatment response against a control group of clients in the same cell who also had above-mean debt by including the strata dummies. The messages are admittedly coarse, yet in spite of this we still find significant effects. We considered telling clients what their exact location in the distribution of debt and risk was but this required personalization.

Presenting information on the probability of default was harder, the marketing department of the bank argued that their average client would not grasp the concept. Therefore, it decided to present the information graphically in the form of a thermometer. The thermometer was in the high temperatures when the client was above the median probability of default. The bank also decided to congratulate clients who were below the median probability of default and show the thermometer in the low temperatures.

Row 1 of Table 2 reports the average impact of the "high risk" message and Row 2 the respective estimate for the "low risk" message. As expected, the effects across these two messages move in opposite directions and give more confidence to our causal interpretation. The "high risk" message caused a decrease in debt of 233 pesos in march and a decrease in delinquency of

behavioral change. It may reflect sophisticated inferences and rational behavior. One could imagine an environment in which individuals with similar preferences are subject to common but only partially observable income shocks, where each peer observes a signal of the shock. In such a context, observing the actions of others would convey information about the state of the shock and push the individual toward performing a similar action than his peers. Alternatively, there may exist a tendency toward conformity directly in the utility function. We do not attempt to distinguish between these forces.

1.5pp, about 8 percent of mean delinquency. The "low risk" message increased delinquency by 7 percent of its respective mean in both March and April. Note also that voluntary account closures increased for clients receiving the high risk message vs. their controls by 0.7pp (about 16 percent of mean value).

Rows 3 and 4 show that the "high debt" message reduced debt although in a small amount (about 0.05pp from the mean at most) which is not statistically distinguishable from zero.[27] Neither of the two messages influenced delinquency. There is a decrease of 0.3pp in voluntary account closures for those that received the advice, significant at the 10 percent level.

D. Debiasing Warning Message

We also test a warning message, shown in Figure 4 (b), aimed at increasing awareness that paying down debt is hard and de-biasing consumers by explicitly telling them that people are typically overconfident in their ability to pay down debt. This is pertinent given that as documented above our consumers seem to display overconfidence. The only papers we are aware of that measure response to warnings against biases are those by Cummings and Taylor (1999) and List (2001). They show that "debiasing" individuals by warning them of the bias in answers to hypothetical valuation questions can help individuals to approximate true valuations.

The bank sent this message to a randomized treatment group of 12,900 consumers. Results in row 7 of Table 2 show that the message did decrease debt, although again, to a very limited degree, -126 pesos on average on March and -147 on April. Effects of delinquency have negative signs but are not statistically significant at conventional levels.

E. TILA-like vs non-TILA messages, and other outcomes

All in all, the nonstandard disclosures were more effective at reducing delinquency than TILA-type disclosures. This is confirmed in the F-test of the bottom of Panel in A Table 2 which shows that although we cannot reject

[27]We estimated regressions comparing the high debt messages with and without the advice, testing if clients followed the advice of paying close to twice the minimum amount; there was no evidence that they did.

TILA messages having a zero effect, we can reject non-TILA messages having zero effect on debt and delinquency. Panel B of Table 2 reports estimates of equation 2 and rejects equality of effect of TILA and non-TILA on debt, although we can not reject equality for other outcomes.

Even if non-standard disclosures seem more effective, their average effects are still small and in our opinion unlikely to have any major impact on consumer interest payments, financial sector stability, or competition.

We mentioned earlier that not finding behavioral changes in this card does not imply that there is no change in other cards. For example the consumer may open a new card and slowly start to shift activity to this new card, or the consumer may realize that he has too much debt and default on other cards.

We were able to obtain information for a random sample of 17,801 of our accounts from the Credit Bureau to measure these market level spillovers[28]. We study three outcome variables: (a) openings of new card accounts in other banks (a proxy for switching), (b) the credit score, (c) and number of other credit cards in default. Table 4 presents the estimates of equations 1 and 2 with these outcome variables. There are no statistically significant spillover effects, except for the credit score decrease as a result of the low risk message.

IV. Heterogeneity, Specification Checks and Power

A. Response Heterogeneity

Section III. reported zero average effects for the interest rate disclosure. We performed some checks that suggested this was not driven by responses canceling each other out in the average (the good news vs bad news interpretation). In this section we explore treatment effect heterogeneity explicitly by allowing the response to depend on i. We estimate the following equation with a random intercept and a random treatment effect coefficient using data from

[28] These are distributed across treatments as follows 603 of High Risk, 746 of Low Risk, 1,467of High Debt plus Advice, 1,392 High Debt, 1,323 Months to Pay, 1,311 Interest Rate, 1,431 Warning and 9,616 for the control group. We verified that there is balance across treatments in this subsample

March to June for each consumer:[29]

$$Debt_{it} = \alpha + \nu_1 i + (\beta + \nu_{2i})T_j + \gamma S_k + \lambda_t + \epsilon_{it} \tag{3}$$

The results are reported in the Online Appendix (Table 3). One highlight is that the mean of β_i is -36 and its 95 percent confidence interval is (-486,414). ν_{2i} has a small standard deviation, its 95 percent confidence interval is (5.35e-24,1.86e+18). This means that there is very little heterogeneity in the response to treatment, casting doubt on the good news vs bad news interpretation for the small estimated average effect. Figure 8 shows graphically the extent of heterogeneity by plotting plots the empirical Bayes prediction of the random coefficients for the treatment ν_{2i}. The distribution has most of its mass on the negative part of the support, suggesting that most clients decrease their debt as a response to the treatment, but the whole support of the distribution is economically tiny and concentrated around zero.

Tables 4 and 5 in the Online Appendix focus on heterogeneity by splitting samples across some pre-specified dimensions: number of products with the bank (as a proxy for loyalty) and categories of income as reported on the card's application. Here we just want to highlight some results. First, note that the effect on closings when the "high risk" message is received is not present when the card holder has several products with the bank, maybe owing to higher switching costs (Table 4, Online Appendix). Second, income also seems to matter: high income individuals reduce debt more strongly (3 percent of *their* mean debt) when receiving the "high debt" message (perhaps because they can afford to do this without foregoing much consumption); while low income individuals respond mainly through less default, with no detectable effect in debt (Table 5, Online Appendix).

We perform one last exercise here, raised by one important policy question that the CNBV had. The CNBV wanted to issue rules that mandate the use

[29] Month dummies (λ_t) were included to control for omitted time effects. S_k represent the stratification controls. We used Stata xtmixed command which assumes joint normality of the error terms ν_{1i} and ν_{2i} below. Individuals i (second level) group monthly observations t, which run from 1 to 5 (first level).

of predicted risk as a trigger of messages. That is, send messages only to those most likely to default in the near future. The CNBV expected larger responses for these individuals, but is it really the case? We estimated a version of equation 1 with delinquency as the outcome variable, where we additionally include quintiles of CNBV's predicted probability of default (PD) by themselves and interacted with the treatment messages. Figure 9 plots the coefficients in the interactions.[30] It turns out that all the messages had statistically significant effects for the highest PD quintile, which supports CNBV's conjecture. Interestingly (although hard to rationalize) the "high debt" messages appear to be highly effective to decrease risk for low risk individuals while they *increased* risk for high PD clients.

Putting the magnitudes of the coefficients into perspective, we interpret that there is little heterogeneity, except perhaps at the very top of the risk distribution. We therefore think that the main message of the paper –small effects of messages– is robust across a very different subsubsets of the population.

B. Power and Specification check

Statistical power is crucial in studies that cannot reject the null hypothesis of a zero effect. To estimate the power of our test we did Montecarlo exercises. We simulated placebo treatments of different sizes for January 2011 (i.e., just before treatment) and used the regression specification in equation (1) to estimate the faction of time we were able to reject the null of zero effect. That is, we use the same sample and the same methodology as that employed with the real treatments, just two months before treatment.[31] Figure 10 shows that the design/sample has substantial power: we can detect an effect size of 455 pesos in debt (1.6 percent of mean debt) and 0.006pp in delinquency (4 percent of its mean) with 80 percent power. We believe these are relatively small effects, the bank agrees.

We also ran an specification (placebo) test by estimating equation (1) in the same partition of control and treatment cards for the months of September

[30]For the "high risk" and "low risk" message we only plot the last three quintiles and the first three quintiles respectively because these messages were sent to these subpopulations.

[31]Other months worked similarly.

and October 2010, i.e. before treatment. If the equation was misspecified, we would expect more than 95 percent of coefficients to be significant at the 5 percent level. Table 5 shows this was not the case: only one coefficient out of 28 was significant at the 5 percent level or less. This increases our confidence that the significant coefficients we found in Tables 2 and 3 are not due to sampling variance.

Since we are estimating the effects for several treatments and several dependent variables, we may be finding spurious statistically significant effects by failing to account for multiple testing. Underestimation of p-values are not a major concern for the paper as we are arguing that the effects are zero or very small and inconsistent with policymakers' large emphasis on the importance of disclosures. Note also that the statistically significant effects have the expected sign and are present only one or two months after the treatment and not before or later.

C. External validity

In the introduction we mentioned that the main motivation behind the TILA was to enable interest rates to be compared more easily. One could argue that the information provided in our interest rate message caused no response because this information was not useful as it gave no benchmark for comparing the interest rate to that of other banks. Indeed Kling et al. (2012) show that small comparison frictions can have significant effects. Another argument against our findings is that external validity is limited in three ways. First, we studied only one bank. Second, we considered only a risky population. Finally, and potentially more importantly, we sent the message only once. A higher frequency of messages could have had a greater impact.

External validity questions are hard to address since by definition they are beyond the scope of the study sample. However one of the authors was able to run more experiments at other banks for another paper (Negrín and Seira (2014)). In this section we report some of those results. The experiment reported in this subsection goes out of sample in many ways that address the concerns above: it was carried out at two different banks from the one referenced in this paper and was representative of all their clients, not just the risky;[32] the price message was more aggressive as it involved direct price comparisons across bank; and the frequency of the message was varied randomly.

The experiment was motivated by the TILA's emphasis on price comparison and a new disclosure mandate from Mexico's central bank. In 2011, the Central Bank mandated disclosing the interest rates and APRs of *competitor banks* for similar cards in monthly statements every 6 months, separately for classic, gold and platinum cards. Figure 7 contains an example of the price comparison printed with the monthly statement. One could argue that this is a very aggressive mandate, it is uncommon to force companies to advertise the prices of their competitors when many of these prices are actually lower. Actually we know of no other country which requires this.

Since we were interested in varying the frequency of the message one of the authors designed an experiment in conjunction with the central bank. One group of clients in each of two large banks received the price comparison of Figure 7 in April 2012 once, another group (the frequent treatment group) received it *monthly* 7 times from April 2012 to October 2012, while a third group acted as a control group and received no messages at all during all 2012. Group size was approximately 30,000 clients per arm, representative of the banks' overall population of cardholders (no selection on risk was made).

To analyze the resulting behavior we estimated equation (1) by ordinary least squares (there were no strata). We estimated the regression separately for each bank and month. Figure 11 plots the estimated $\beta's$, scaled down by the average of the dependent variable, for ease of interpretation. Panel A reports the results when debt is the dependent variable and Panel B when it is delinquency. As can be seen, the effects are very small economically speaking for all banks, all messages and all periods, i.e., always less than 5 percent of the mean and, more often than not, less than 1 percent of the mean of the dependent variable. None of the coefficients were statistically different from zero at the 10 percent level. Importantly for us, we tested

[32]With these two additional banks the population of this paper would easily include more than half of Mexico's total cardholders.

and found that message frequency made no difference. These results are a powerful demonstration that our main results seem to be more general than just a specific interest message in a specific bank sent once. In the Online Appendix we estimate non-experimentally with matching techniques the effect of *the first time* comparative message of April 2011, using consumers in our main. This is a nice complement since tas his comparative information was disclosed for the first time it it hard to argue that it was not new. We also found zero effect.

V. Conclusion

One lesson from the recent financial crisis is that consumers do not fully understand the financial products they buy, and that this can translate into defaults and bankruptcies. One obvious regulatory response to disclose more information and increase financial education. But which information should we disclose and what should we expect? The content of current legally mandated disclosures has been determined mostly on the basis of introspection rather than rigorous evidence. This paper shows that currently mandated disclosures are likely to have zero effect and that alternative messages that include peer comparisons and warnings are probably more effective, though even so, only to a small degree.

Small effects do not imply money-losing effects. We estimated the same specification with the expected loss in pesos (calculated by the bank) as the dependent variable for the "high risk" message. The treatment group had a 245 pesos lower realized losses in March 2011 (and 295 in April 2011) compared to its corresponding control group. Given that printing and sending the letter cost 2.5 pesos, this message was clearly beneficial to the bank. A sign of the usefulness of the message is that the bank now intends to use it and the authorities aim to mandate such a disclosure.

Many governments around the world are still mindlessly increasing disclosure requirements as a consumer protection device in the credit card market. Given the status quo this negative result is a positive and useful result that must be publicized. This study certainly does not rule out the possibility that other messages could have greater effects, but then they should be rigorously tested.

If we could convince another bank, we would try using a message in which the client is told how much clients with the same credit score and type of card as them are paying at the cheapest bank and an estimate of how much savings *in pesos* they would make in a year by switching to that bank. Bertrand and Morse (2011) show that adding up several months and putting quantities in money terms has worked in other contexts, while Kling et al. (2012) show that direct and personalized comparisons among providers have been effective. Facilitating direct comparisons and easing procedures to switch banks may be important complementary policies to information disclosures.

References

Allcott, Hunt, "Social Norms and Energy Conservation," *Journal of Public Economics*, October 2011, *95* (9-10), 1082–1095.

Ausubel, Lawrence M., "The Failure of Competition in the Credit Card Market," *American Economic Review*, March 1991, *81* (1), 50–81.

Bertrand, Marianne and Adair Morse, "Information Disclosure, Cognitive Biases and Payday Borrowing," *The Journal of Finance*, December 2011, *66* (6), 1865–1893.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen, "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 2011, *3* (1), 91–128.

Chen, Yan, F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *American Economic Review*, September 2010, *100* (5), 1358–1398.

Chetty, Raj and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2004, *99* (4).

Cummings, Ronald G. and Laura O. Taylor, "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method," *American Economic Review*, June 1999, *89* (3), 649–665.

Day, George S. and William K. Brandt, "Research and the Evaluation of Information Disclosure Requirements: The Case of Truth in Lending," *Journal of Consumer Research*, June 1974, *1* (1), 21–32.

Durkin, Thomas A., "Awareness of Credit Terms: Review and New Evidence," *The Journal of Business*, April 1975, *48* (2), 253–263.

_ **and Gregory Elliehausen**, *Truth in Lending*, Oxford University Press, 2011.

Ferman, Bruno, "Reading the Fine Print: Credit Demand and Information Disclosure in Brazil." PhD dissertation, MIT 2011. Job Market Paper.

Festinger, Leon, "A Theory of Social Comparison Processes," *Human Relations*, 1954, *7.*

Hall, R. and S. Woodward, "Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence," *American Economic Review*, 2013, *102.*

Hastings, Justine S. and Jeffrey M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence form Two Experiments," *Quarterly Journal of Economics*, 2008, *123* (4), 1373–1414.

Jensen, Robert, "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.

Jin, Ginger Zhe and Phillip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 2003, *118* (2), 409–451.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel, "Comparison Friction: Experimental Evidence from Medicare Drug Plans," *Quarterly Journal of Economics*, January 2012, pp. 1–37.

Lacko, J. and J. Pappalardo, "Improving Consumer Mortgage Disclosures," *Federal Trade Commission Bureau of Economics Staff Report*, 2007.

List, John A., "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards," *American Economic Review*, 2001, *91* (5), 1498–1507.

Malmendier, Ulrike and Young Han Lee, "The Bidder's Curse," *American Economic Review*, 2011, *101* (2), 749–787.

Melzer, B., "The Real Costs of Credit Access: Evidence from the Payday Lending Market," *Quarterly Journal of Economics*, 2011, *126.*

Negrín, José Luis and Enrique Seira, "The effects of price comparison messages," 2014. Work in progress.

Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa, "Do consumers borrow on their cheapest credit card? Evidence from Mexico," 2012. Stanford University.

Shay, Robert P. and Milton W. Schober, *Consumer Awareness of Annual Percentage Rates of Charge in Consumer Installment Credit: Before and After Truth-in-Lending Became Effective*, Vol. 1, Washington, DC: U.S. Government Printing Office, 1973.

Soll, Jack B., Ralph L. Keeney, and Richard P. Larrick, "Consumer Misunderstanding of Credit Card Use, Payments and Debt: Causes and Solutions," *Journal of Public Policy and Marketing*, forthcoming.

Stango, Victor and Jonathan Zinman, "Borrowing High vs. Borrowing Higher: Sources and Consequences of Dispersion in Individual Borrowing Costs," November 2011.

_ **and** _ , "Fuzzy Math, Disclosure Regulation, and Market Outcomes: Evidence from Truth-in-Lending Reform," *The Review of Financial Studies*, 2011, *24* (2), 506–534.

_ **and** _ , "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees," April 2013.

Stewart, Neil, "The Cost of Anchoring on Credit-Card Minimum Repayments," *Psychological Science*, 2009, *20* (1), 39–41.

White, M., "Bankruptcy Reform and Credit Cards," *Journal of Economic Perspectives*, 2007, *21.*

VI. Tables and Figures

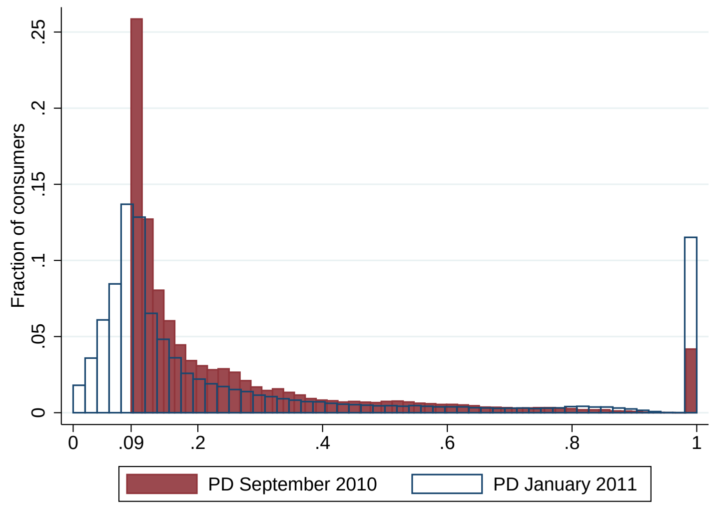


Figure 1: CNBV Probability of Default

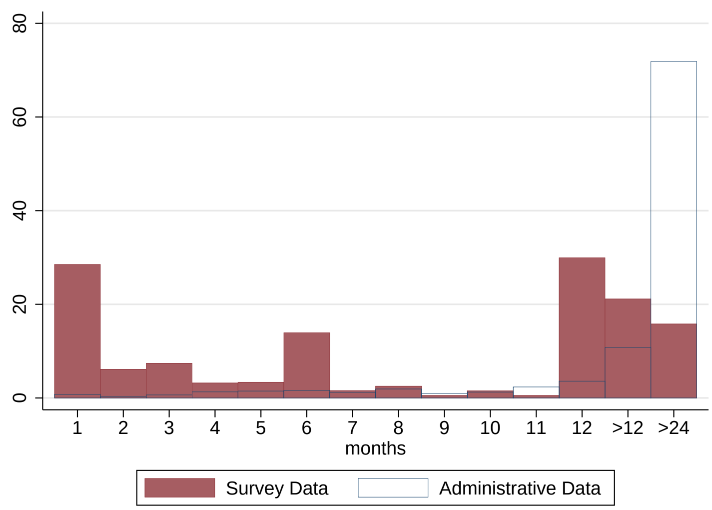


Figure 2: Beliefs on months to pay debt against actual administrative data

Dear XXXXX,

We want to help you keep your finances healthy. Your interest rate plays a crucial role in interest generation and in increasing debt.

Pay careful attention to your card's **interest rate**.

INFORMATION AT STATEMENT DATE			
Credit Limit	$0,000.00	Annual Personalized Rate	XXXX%
Available Credit	$0,000.00	Annual Rate	XXXX%
Average Daily Balance	$0,000.00	Annual Investment Rate	0.00%
Annual Penalty Rate	00.00%	Statement Date	27 MAY 2010
Total Annual Cost (w.o. tax)	00.00%	Days in Cycle	30

XXXX %

You can find this information in your bank statement.

(a) Interest Rate Disclosure

Dear XXXXX,

We want to help you keep your finances healthy. That's why, in your bank statement, we tell you the number of months it will take you to pay off your debt if you only pay the minimum amount due.

SUMMARY OF CTIVITY		
DETAILED TRANSACTIONS OF JANE DOE 0000 0000 0000		
Date	Description	Amount
14 MAY	PAYMENT	$0,000.00
27 MAY	INTEREST SUBJECT TO TAX	$0,000.00
04 MAY	ANNUAL FEE	$0,000.00
27 MAY	TAX ON INTEREST AND FEES	$0,000.00

The time needed to pay of your debt by making only the minimum payment due is: xx months. The amount due in 12 months if you are up to date with your payments is: $000.00. This does not consider purchases, interests, cash advances or fees incurred after the present statement date.

xx months

If you want to take less time, consider paying more than the minimum.

(b) Months-To-Pay-Off-Debt Disclosure

Figure 3: Salient Legally Mandated Messages

The figures present an English version of the messages sent in the experiment. This is the precise format used, except that the originals were in Spanish.

Dear XXXXX**,**

We want our clients to have healthy finances. That's why we have analyzed the credit behavior of a group of cardholders.

With respect to this group **your debt** is:

HIGHER

than the average of
people similar to
yourself*

To reduce this risk, we recommend you do the following:

Analyze your ability to pay and budget your monthly expenses.

Pay at least twice the minimum amount due in order to reduce the time it will take you to pay off your debt.

Maintain your debt well below your credit limit.

(a) High Debt + Advice

Dear XXXXX**,**

We want to help you keep your finances healthy.

Don't get confident
Paying off a debt is
not that easy

Many studies have found out
that consumers overestimate their
ability to pay and
fail to service their debts.

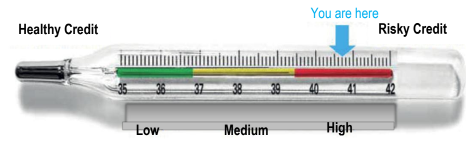
Don't let it happen to you!

(b) Warning

Figure 4: High Debt and Warning Messages

Dear XXXXX,

Based on your credit behavior, we have detected that your credit card has the following **probability of default:**



To reduce this risk, we recommend you do the following:

Analyze your ability to pay and budget your monthly expenses.

Pay at least twice the minimum amount due in order to reduce the time it will take you to pay off your debt.

Maintain your debt well below your credit limit.

Figure 5: High Risk Message

The “low risk” message is analogous but the arrow is placed over the lower legend on the thermometer and the client is congratulated. See the Online Appendix.

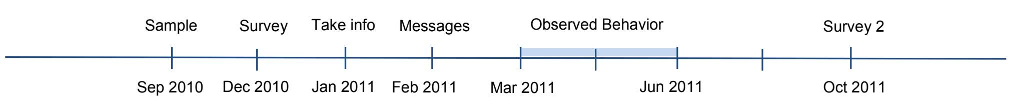


Figure 6: Experiment Timeline

Institution	Product	CAT (%)	Weighted Average Effective Rate (%)	Annual Fee (pesos)	Credit Limit (median value in pesos)*
Santander	Santander Light	31.4	24.4	430	17,821
BBVA Bancomer	Azul Bancomer	34.9	26.8	460	11,500
Banco Inbursa	Clásica Inbursa	41.4	35.1	0	4,300
Scotiabank	Tasa Baja Clásica	44.2	34.0	395	15,000
Banamex	Clásica Internacional	44.7	33.6	500	44,000
HSBC	Clásica HSBC	45.4	34.3	480	13,300
Banorte	Clásica	46.6	35.5	430	15,000
BNP Paribas	Comercial Mexicana	78.3	57.1	250	6,500
BanCoppel	BanCoppel	88.3	65.0	0	4.200
Products that account for less than 0.5% but more than 0.1% of the total number of “classic” cards					
Banco Walmart	Super Tarjeta de Crédito	56.3	43.9	200	3,200
American Express	Blue	56.4	41.8	459	12,000
SF Soriana	Soriana - Banamex	56.7	42.4	420	16,800
Ixe Tarjetas	Ixe Clásica	64.5	47.2	440	5,000
Globalcard	Globalcard	90.5	60.4	684	7,500

Figure 7: APR Comparison Message

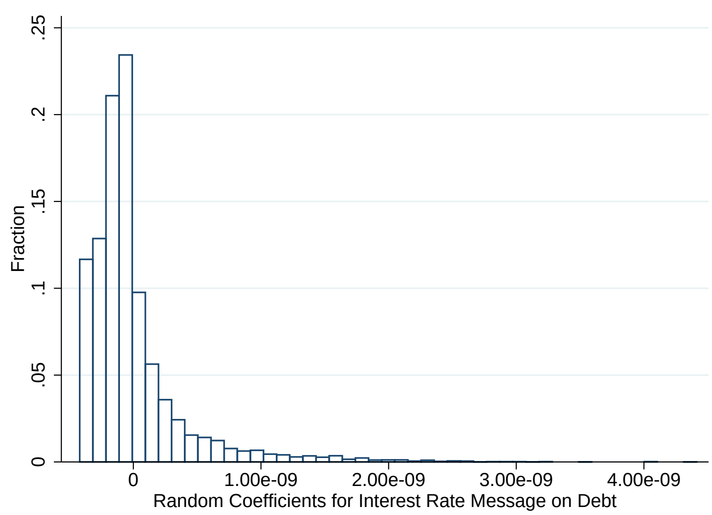


Figure 8: Empirical Bayes Prediction for Interest Rate Message

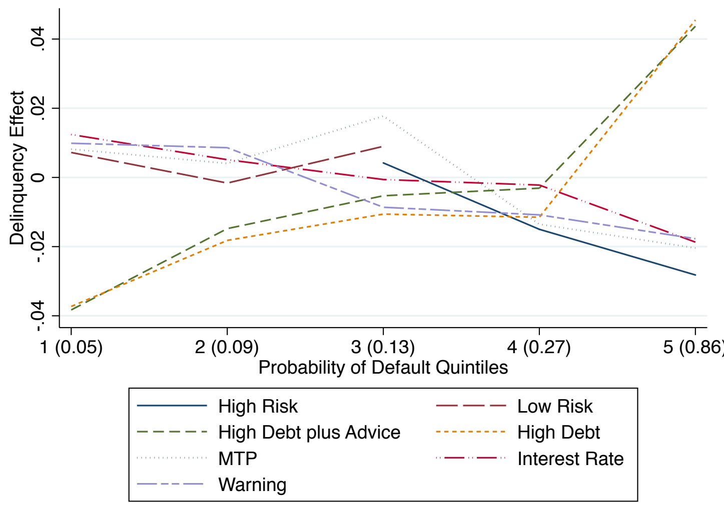
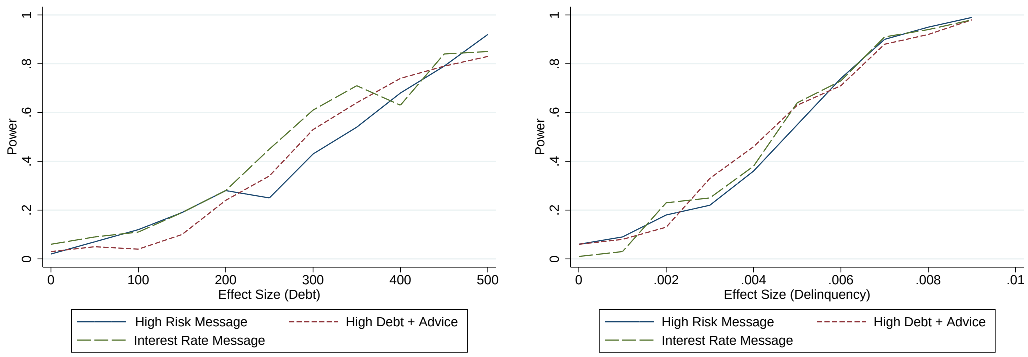


Figure 9: Effects on Delinquency by Probability of Default

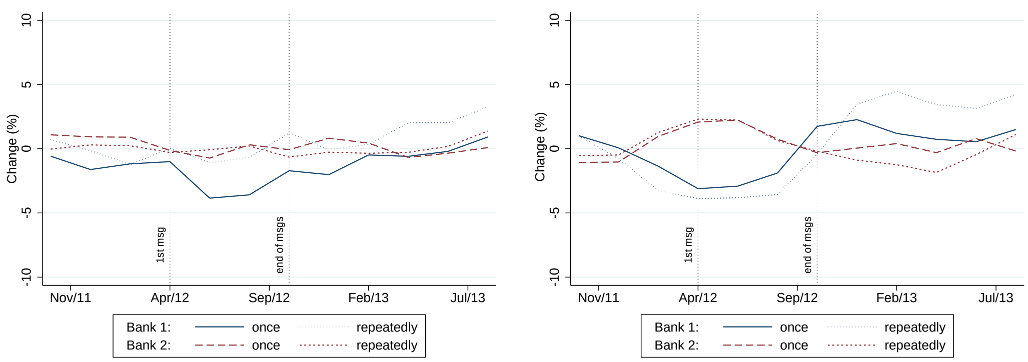
This graph plots the effect of each treatment on delinquency by quintiles of the probability of default as measured in January 2011. The model is analogous to equation (1) but we introduced dummies for each quintile of the probability of default and its interaction with the treatment dummies: $Y_{ijt} = \alpha_t + \sum_{j=1}^{7} \beta_{tj} T_{ij} + \sum_{r=2}^{5} \gamma_{tr} D_{tir} + \sum_{k=2}^{5} \sum_{j=1}^{7} \delta_{tj} D_{tir} * T_{ij} + S_{ik} + \epsilon_{ijt}$. The average probability of default of each quintile is expressed in parentheses on the x axis.



(a) Debt (b) Delinquency

Figure 10: Statistical Power

These graphs report the statistical power to identify effects for selected treatments. We simulated placebo treatments of different sizes for January 2011 (i.e., just before treatment) and used the regression specification in equation (1) to estimate the fraction of time we were able to reject the null of zero effect.



(a) Debt (b) Number of Delinquent Months in the Previous 6 Months

Figure 11: Effects at Two Different Banks

These graphs report the average treatment effect divided by the control group mean for two different banks and two different treatments in each bank . A group of individuals was sent a unique message on April 2012 and another received monthly messages from April to October 2012.

Table 1: Summary Statistics

	Mean	St. Deviation
Dependent Variables		
Delinquent (%)	11	(31)
Debt (MXN)	17800	(25297)
Debt ^{a} (MXN)	18415	(25410)
Closed^{b} (%)	3	(16)
Other Risk Measures		
Probability of Default^{c} (%)	26	(16)
Default^{b} (%)	9	(29)
Expected Loss^{c} (MXN)	2721	(5841)
Expected Loss^{ac} (MXN)	2767	(5844)
Credit Terms and Use		
Credit Score*	642	(50)
# Active CC*	3.19	(2.82)
# CC in Default*	0.293	(0.969)
# CC Opened***	0.07	(0.254)
Credit Limit (MXN)	27502	(35831)
Annual Interest Rate (%)	44	(10)
Monthly Interest (MXN)	646	(896)
Months to Pay	27	(17)
Minimum Payment (MXN)	1490	(3538)
Utilization	70	(38)
Purchases (MXN)	1082	(4365)
Payments (MXN)	1925	(5659)
Demographics		
Age (Years)	42	(12)
Tenure (Months)	43	(26)
Male (%)	57	(49)
Income^{d}	8563	(7444)
Observations	3343800	

Credit card variables are expressed in monthly terms.

* Obtained from a subsample of 17,815 individuals for whom we have a snapshot of Credit Bureau information from June 2010.

*** We count the number of credit cards opened during March, April, May and June 2010, as reported by the Credit Bureau.

[a] Conditional on being positive.

[b] Measured in February 2011.

[c] Measured in September 2010.

[d] Proxied by expenditures. Self-reported in the survey. After trimming the top 5 percent

Table 2: Baseline Results

	Dependent Variables				
	Debt		Delinquent		Closed
	March	April	March	April	June
	Panel A				
Mean Dep.	17391	16541	0.183	0.198	0.043
S.D. Dep.	(24425)	(23964)	(0.387)	(0.398)	(0.204)
High Risk	-233***	-172	-0.015***	-0.006	0.007***
	(90)	(118)	(0.00495)	(0.00511)	(0.00267)
Low Risk	4.4	82	0.014***	0.013***	0.001
	(84)	(108)	(0.00494)	-0.0051	(0.00267)
High Debt + Advice	-29	-127	0.002	0.005	-0.003*
	(64)	(83)	(0.00361)	(0.00373)	-0.00195
High Debt	-104	32.77	-0.002	0	0
	(62)	(81)	(0.00355)	(0.00366)	(0.00192)
Rate	-35	14	0	0	0.001
	(63)	(81)	(0.00356)	(0.00367)	(0.00192)
MTP	43	90	0	0.006	-0.002
	(64)	(82.74)	(0.00361)	(0.00373)	(0.00195)
Warning	-126**	-147*	-0.002	-0.002	-0.002
	(62)	(81)	(0.00355)	(0.00366)	(0.00192)
F-test TILA	0.635	0.535	0.995	0.270	0.536
F-test Non-TILA	0.0400	0.143	0.00272	0.0739	0.0254
	Panel B				
TILA	6	48	-0.001	0.002	-0.001
	(46)	(60)	(0.00265)	(0.00274)	(0.00143)
Non-TILA	-107***	-99**	-0.004*	-0.001	0
	(38)	(50)	(0.00217)	(0.00224)	(0.00118)
F-test	0.0312	0.0312	0.381	0.312	0.948
N	147634	143484	167190	167190	167190

Significance level: * 10 percent ** 5 percent *** 1 percent. Standard errors in parenthesis. On each panel, each column represents a regression and each row a treatment group dummy. On panel A each of the variables on the first row is regressed on dummies for all treatments and stratification indicators; at the bottom of the panel we report the p-values of testing whether the coefficients of Rate and MTP (TILA) are jointly different from zero and whether the other 5 treatments have jointly different from zero results. Panel B reports the coefficients of regressing the same outcome variables on two dummies, the first one takes the value of one when the cardholder is in the interest rate or months-to-pay treatment groups and the other one when the individual was on any other (non-TILA) treatment group with the exception of the Low Risk message (because the effect intended of this message goes in the opposite direction).

Table 3: TILA-like Disclosures on Subsamples

	Dependent Variables					
	Debt		Delinquent		Closed Account	
	March	April	March	April	June	Observations
	Panel A:Interest Rate Disclosure					
Paid Interest[a]	-19	10	-0.002	0	0	52257
	(70)	(96)	(0.0047)	(0.0050)	(0.0017)	
Interest Rate[b]	-32	22	0.003	0.003	0.001	60947
	(65)	(89)	(0.0046)	(0.0047)	(0.0026)	
Debt[b]	4	161	-0.008	-0.005	0.001	45744
	(118)	(152)	(0.0058)	(0.0060)	(0.0015)	
Change in	29	-106	0	-0.005		78480
Absolute Value	(58)	(76)	(0.0041)	(0.0042)		
Another Credit Card[d]	-525***	-410	-0.022*	-0.024*	0.005	7492
	(194)	(249)	(0.0122)	(0.0127)	(0.0053)	
	Panel B: Months to Pay Off Debt Disclosure					
Paid Interest[a]	-28	-61	0.001	0.011**	0	52436
	(102)	(120)	(0.0047)	(0.0050)	(0.0050)	
Debt[b]	87	185	0.002	0.012*	-0.002	45287
	(174)	(195)	(0.0058)	(0.0060)	(0.0060)	
MTP[b]	26	105	0.001	0.006	-0.003	46312
	(134)	(141)	(0.0035)	(0.0038)	(0.0038)	

Significance level: * 10 percent ** 5 percent *** 1 percent. Standard errors in parenthesis.

In this table each coefficient comes from a different regression of the outcome variables of each month on the treatment dummy and the stratification indicators. Panel A shows the effects of the interest rate message on different subsamples to disentangle any cancelling effects:

[a] Paid interest in the 10 months prior to March 2011.

[b] Above median in January 2011.

[c] Temporary low interest rate offer in January 2011.

[d] Cardholder has an active card from another bank. We ran this regressions on the individuals (in the control or interest-rate message groups), from our Credit Bureau sample, that had an active credit card form a different bank at each specific month.

Table 4: Effects on Other Credit Conditions (subsample)

	Dependent Vars		
	Score June	Opens New Card March-June	# CC in default June
		Panel A	
Mean Dep.	622	0.008	0.813
S.D. Dep.	(67)	(0.091)	(1.77)
High Risk	-1.09	-0.002	-0.073
	(2.8)	(0.004)	(0.075)
Low Risk	-4.32*	0.001	0.037
	(2.52)	(0.004)	(0.067)
High Debt + Advice	-1.365	-0.002	0.038
	(1.85)	(0.003)	(0.05)
High Debt	1.18	0.002	-0.068
	(1.9)	(0.003)	(0.051)
Rate	0.18	-0.004	-0.007
	(1.92)	(0.003)	(0.051)
MTP	1.69	-0.003	-0.07
	(1.91)	(0.003)	(0.051)
Warning	0.58	-0.002	-0.039
	(1.9)	(0.003)	(0.051)
F-test TILA	0.677	0.135	0.386
F-test Non-TILA	0.493	0.936	0.474
		Panel B	
TILA	1.26	-0.003	-0.0409
	(1.42)	(0.002)	(0.038)
Non-TILA	0.22	-0.001	-0.0303
	(1.15)	(0.002)	(0.031)
F-test	0.517	0.322	0.804
N	17781	17781	17801

Significance level: * 10 percent ** 5 percent *** 1 percent. Standard errors in parenthesis.

This table shows the beta coefficients of estimating equation (1) on a subsample of individuals for whom we had a June 2011 snapshot from the Credit Bureau. Here we can observe their credit score by June and whether they opened new credit cards in different banks. As in Table 2, each column represents a regression. On Panel A, each of the variables on the first row is regressed on dummies for all treatments and stratification indicators; at the bottom of the panel we report the p-values of testing whether the coefficients of Rate and MTP (TILA) are jointly different from zero and whether the other 5 treatments have jointly different from zero results. Panel B reports the coefficients of regressing the same outcome variables on two dummies, the first one takes the value of one when the cardholder is in the interest rate or months-to-pay treatment groups and the other one when the individual was on any other treatment group with the exception of the Low Risk message (because the effect intended of this message goes in the opposite direction).

Table 5: Placebo tests

	Dependent Variables				Dependent Variables (subsample)		
	Debt		Delinquent		Credit Score	Open Cards	#CC in Default
	September 2010	October 2010	September 2010	October 2010	June 2010	March-June 2010	June 2010
			Panel A				
Mean Dep.	18919	18937	0.135	0.145	642	0.07	0.293
S.D. Dep.	(25800)	(25727)	(0.341)	(0.352)	(50)	(0.254)	(0.969)
High Risk	-48	11	0.004	0.002	-0.351	-0.003	-0.025
	(163)	(168)	(0.004)	(0.00429)	(2.098)	(0.0107)	(0.0416)
Low Risk	139	129	0.001	-0.001	0.238	-0.008	0.009
	(162)	(166)	(0.00391)	(0.00429)	(1.888)	(0.00965)	(0.0375)
High Debt + Advice	22	48	0	0.003	-0.161	-0.003	0.005
	(119)	(122)	(0.00285)	(0.00313)	(1.383)	(0.00707)	(0.0275)
High Debt	-40	-74	0.002	0.002	3.442**	-0.007	-0.051*
	(119)	(122)	(0.00285)	(0.00313)	(1.418)	(0.00722)	(0.0281)
Rate	54	47	0	-0.003	1.291	0.001	-0.016
	(117)	(120)	(0.00281)	(0.00308)	(1.441)	(0.00731)	(0.0285)
MTP	59	11	0	-0.003	-0.151	-0.001	-0.03
	(116)	(120)	(0.00280)	(0.00308)	(1.448)	(0.00734)	(0.0284)
Warning	70	98	-0.001	-0.001	-0.671	0.006	-0.035
	(116)	(120)	(0.00280)	(0.00307)	(1.425)	(0.00727)	(0.0282)
F-test TILA	0.81	0.93	0.99	0.41	0.65	0.99	0.52
F-test Non-TILA	0.93	0.87	0.87	0.9	0.24	0.76	0.41
			Panel B				
TILA	47	19	0	-0.002	0.588	0	-0.024
	(87)	(89)	(0.00209)	(0.00230)	(1.074)	(0.00545)	(0.0212)
Non-TILA	2	17	0.001	0.001	0.674	-0.001	-0.027
	(71)	(73)	(0.00171)	(0.00188)	(0.862)	(0.00439)	(0.0171)
F-test	0.65	0.98	0.74	0.09	0.94	0.8	0.9
N	165042	163113	167190	167190	17077	17815	17815

Significance level: * 10 percent ** 5 percent *** 1 percent. Standard errors in parentheses.
This table replicates the estimations on table 2 for the months of September and October 2010 (pre-treatment). Each column represents a regression. On panel A each of the variables on the first row is regressed on dummies for all treatments and stratification indicators; at the bottom of the panel we report the p-values of testing whether the coefficients of Rate and MTP (TILA) are jointly different from zero and whether the other 5 treatments have jointly different from zero results. Panel B reports the coefficients of regressing the same outcome variables on two dummies, the first one takes the value of one when the cardholder is in the interest rate or months-to-pay treatment groups and the other one when the individual was on any other treatment group with the exception of the Low Risk message (because the effect intended of this message goes in the opposite direction). The last three columns are estimated on our subsample of individuals for whom we have Credit Bureau data. In this table we use a snapshot of June 2010 and estimate the same model as in Table 4