American Economic Journal: Economic Policy 2017, 9(1): 277–307
https://doi.org/10.1257/pol.20140404

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

By Enrique Seira, Alan Elizondo, and Eduardo Laguna-Müggenburg*

Consumer protection in financial markets in the form of information disclosure is high on government agendas, even though there is little evidence of its effectiveness. We implement a randomized control trial in the credit card market for a large population of indebted cardholders and measure the impact of Truth-in-Lending-Act-type disclosures, de-biasing warning messages and social comparison information on default, indebtedness, account closings, and credit scores. We conduct extensive external validity exercises in several banks, with different disclosures, and with actual policy mandates. We find that providing salient interest rate disclosures had no effects, while comparisons and de-biasing messages had only modest effects at best. (*JEL* D14, D83, G21, G28, O16)

The recent financial crisis and the advent of behavioral economics have placed renewed focus on consumer protection in the financial sector. Consumer confusion has often been mentioned as one of the causes of the financial crisis. As a result, many countries—including the United States—have mandated more information disclosures.[1] Many countries have followed suit, requiring financial institutions to report yet more information. However, although evidence has been accumulating especially in the last five years, we still know little about how effective these disclosures are, which ones are better, and for which types of consumers. 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

* Seira: Department of Economics, ITAM, Rio Hondo 1, Cuidad de Mexico 01800, Mexico; Banco de Mexico; and J-PAL (e-mail: enrique.seira@itam.mx); Elizondo: Banco de Mexico (e-mail: aelizondo@banxico.org.mx); Laguna-Müggenburg: Stanford University, 579 Serra Mall Stanford, CA 94305 (e-mail: edu.laguna@stanford.edu). We are grateful to the bank involved in this study and to the Comisión 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 Felix, Karen Olivo, Alonso De Gortari provided outstanding research assistance. Enrique Seira gratefully acknowledges support from Asociación Mexicana de Cultura, A.C. and from Banco de Mexico where most of this work was done. All errors remain our own.

† Go to https://doi.org/10.1257/pol.20140404 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] The United States created the Consumer Protection Bureau and mandated new information disclosures in the Credit Card Accountability Responsibility and Disclosure Act of 2009. Among the 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 and there are no further purchases. The publication of APRs and interest rates has been a requirement since 1968. The Office of Information and Regulatory Affairs (OIRA) under the leadership of Cass Sunstein has emphasized the importance of "smart" information disclosures (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).

matter of debate, and it deserves careful consideration." Indeed, there is a debate. For instance, Julie Williams, acting Comptroller of the Currency in 2005, believed financial disclosure policy had not worked well for consumers and had unnecessarily burdened banks, while R. Krozner, former Chairman of the Federal Reserve System's Committee on Supervision and Regulation of Banking Institutions, was optimistic on their effectiveness (Williams 2005, Krozner 2007).

In order to talk about their effectiveness one needs to understand what these disclosure laws are trying to achieve. The most famous of 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."[2] Although credit cards figure prominently in the letter of the law, after almost half a century of this law there is surprisingly little rigorous evidence on whether disclosures reduce risk, strengthen competition, or modify indebtedness in the credit card market.

A strong case can be made that disclosures could have large effects. Lacko and Pappalardo (2007), for example, find that mortgage cost disclosures in the United States failed to convey key mortgage costs to many consumers, while Woodward and Hall (2012) show that borrowers sacrifice at least $1,000 by shopping from too few mortgage brokers. Stango and Zinman (2016) find that similar consumers pay substantially different interest rates for credit cards, while Ponce, Seira, and Zamarripa (forthcoming) find that Mexican cardholders' debt allocation is insensitive to differences in the interest rates of the cards they already hold, suggesting that they may not know their rates. In fact, in a survey we conducted in Mexico for a subset of this paper's sample only 3 percent of cardholders claim to know the exact interest rate on their card. These studies suggest that high price search cost or inattention may be at play. There is also some evidence of over-borrowing that could potentially 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" with a share of 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 making harder regulatory or policy choices. The Mexican Banking Commission (CNBV) actually encouraged us to pursue this agenda, since

[2] Quoted from the first paragraph of the TILA (1968). 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.

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 at risk of default; we thus focused on this population.

To conduct the experiment, the CNBV paired us with a large Mexican bank that wanted to improve their messages and decrease default. We helped the bank's marketing department to design seven messages. The first two were about hard information, inspired by laws such as the TILA: one included a personalized interest rate and the other a measure of months to pay (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.[3] 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 use a picture to suggest a comparison of risk vis-à-vis 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, as well as by Ausubel's (1991) conjecture of the existence of a large fraction of overconfident consumers.

These messages were sent to a subset of more than 160,000 cardholders that seemed risky, although had not defaulted. Our baseline survey shows that indeed they were highly indebted, unaware of their interest rate, and overconfident as to their ability to pay down their existing debt. The messages were randomized and sent to treatment/control groups so that causal inference could be made 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: interest paying debt, delinquency, voluntary account closures by the client, and 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 overall, or they may strategically lead to less/more default in other cards. We approximate market perceived risk by the credit score in the Credit Bureau, and observe default in other cards also using Credit Bureau information. This set of outcomes is really broad and we view it as one of the contributions of the paper.

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. This result is particularly striking given the low awareness environment among our sample. The other TILA type message—the MTP—actually increased delinquency slightly. On the positive side, we find that non-TILA messages are moderately effective even when the information provided is quite coarse. The "high risk" message was particularly useful in

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

decreasing delinquency, with an effect of 8.2 percent of mean delinquency. 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 relatively small and short-lived, lasting only one or two months. We show that the failure to detect effects is not due to low statistical power.

From these 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 perhaps they are made easier to understand and easily actionable. Nonetheless, even small effects may be worthwhile given that sending messages is very cheap. We do not evaluate the effect of this information on consumer welfare but, given the size 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, Leslie, and Sorensen (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.[4] Ferman (2011), Stango and Zinman (2011), and Agarwal et al. (2015) are 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 (2011) study the effects of the TILA itself on interest rates. Overall, they find no effects on average interest rates, but they do find lower interest rates for borrowers who typically underestimate APRs.[5] Very recently, Agarwal et al. (2015) study several effects of the CARD Act using a differences-in-differences strategy, and try to disentangle and bound the effect of the MTP messages from the other parts of the regulatory package. They find small effects of the MTP message.

Another strand of literature 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 United States and shows that providing APR comparisons has no effect on

[4] See, for example, Soll, Keeney, and Larrick (2013). Early studies by 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.

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

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. Bertrand et al. (2010) show that advertising content can significantly affect micro-loan take-up, as much as a 25 percent change in interest rates.

We capitalize on the above literature and go further in several respects. 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.[6] Second, we complement Bertrand and Morse (2011) and Ferman (2011) by focusing on the intensive margin. 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 who already have cards). Third, we study how information affects outcomes such as indebtedness, default, voluntary account closures, and switching, all of which are not studied in the literature but are closely related to the outcomes that TILA-like regulations seek to affect. Fourth, we examine the credit card market, a huge 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,[7] different message frequencies, and actual regulatory policies.

I. Context and Data

Before we describe the field experiment, let us outline the context. There was an increased awareness by authorities of the need to provide information to cardholders that allowed them to make good financial decisions. The CNBV was thinking about issuing rules that mandated banks to send personalized warnings based on the risk profile of clients and were discussing such rules with banks. The CNBV contacted us with a partner bank that was seeking ways to reduce delinquency in their credit card portfolio and was preparing to send some messages. They agreed to use our advice in the design of these messages and let us measure their effects but allowed us only to work with their riskier clients.[8] We agreed to work with this population since it is precisely them that regulatory authorities are typically concerned about; we had a strong prior that they were overborrowing and that the messages we designed could decrease debt and default for them.[9]

We focused on such a population by drawing a random sample of 167,190 credit cards from the upper tercile of the risk distribution of clients, where 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

[6] Bertrand et al. (2010) and Bertrand and Morse (2011) do this, but not in the credit card market, only with a single lender, and focus on take-up of the loan.

[7] These banks together have more than 70 percent of the credit card market in Mexico.

[8] Which as described below are not that risky when compared to Mexican cardholders.

[9] 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.

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.

A. *Administrative Data*

The data available to us consist of monthly information on the variables that appeared in the monthly statements for the selected 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. From the Credit Bureau 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. This enabled us to look for induced behavioral changes in other loans as a spillover effect. We have limited demographic information. We used administrative data to follow client behavior; this has the virtue of containing virtually no measurement error and is cheap to collect.

Table 1 shows that indeed clients are leveraged and risky. It also shows that interest rates are high and clients somewhat new. Average interest-paying debt in our credit card sample was around 18,000 pesos while mean monthly income—according to our survey implemented in this sample—was close to 9,000 pesos. Clients were also risky: the estimated ex ante probability of default in the subsequent 12 months was 26 percent. 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] Default (more than 90 days past due) reached 9 percent by February 2011. The expected loss per account as calculated by the banks 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 the card has 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 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

[10] The exact formula is given by $\frac{1}{1 + \exp - (2.9704 + 0.6730CD + 0.4696D - 0.0075T - 1.0217MP + 1 - 1513LU)}$.

[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. Interestingly, the PD methodology has good predictive power on average, but the risk classification using PD is volatile: risky individuals—as predicted with the PD model—do not stay risky for long. For instance, 86 percent of the low-risk-message-receiving clients that were in the tenth risk decile in September 2010 transitioned into a lower decile by January 2011. In this environment, a given individual may be high risk in some months and low risk in others.

Table 1—Summary Statistics

	Mean	SD
Dependent variables		
Delinquent (percent)	11	(31)
Debt (MXN)	17,800	(25,297)
Debt[a] (MXN)	18,415	(25,410)
Closed[b] (percent)	3	(16)
Other risk measures		
Probability of default[c] (percent)	26	(16)
Default[b] (percent)	9	(29)
Expected loss[c] (MXN)	2,721	(5,841)
Expected loss[ac] (MXN)	2,767	(5,844)
Credit terms and use		
Credit score*	642	(50)
Number active credit cards*	3.19	(2.82)
Number credit cards in default*	0.293	(0.969)
Number credit cards opened**	0.07	(0.254)
Credit limit (MXN)	27,502	(35,831)
Annual interest rate (percent)	44	(10)
Monthly interest (MXN)	646	(896)
Months to pay	27	(17)
Minimum payment (MXN)	1,490	(3,538)
Utilization	70	(38)
Purchases (MXN)	1,082	(4,365)
Payments (MXN)	1,925	(5,659)
Demographics		
Age (years)	42	(12)
Tenure (months)	43	(26)
Male (percent)	57	(49)
Income[d]	8,563	(7,444)
Observations	3,343,800	

Note: 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.

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 five 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 percentage points in terms of yearly rates) from one month to the next. This means that information does not depreciate quickly. The interest rate had a tiny correlation with risk, which means that clients with similar risk profiles were paying different interest rates even

[12] When we regress a dummy for attrition versus treatment dummies and strata dummies, we cannot reject the joint hypothesis that the treatment dummies are equal to zero (see Table A1 in online Appendix).

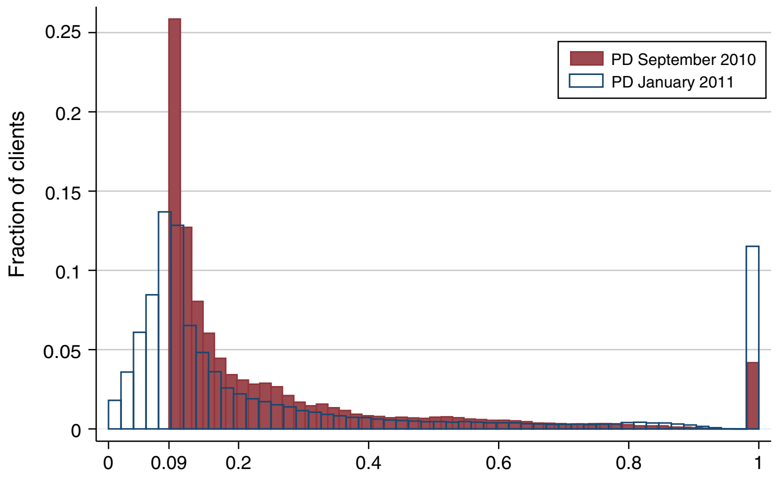


FIGURE 1. CNBV PROBABILITY OF DEFAULT

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 help us formulate hypotheses, we collected two rounds 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 understand how indebted clients were with respect to their income, how satisfied they were with the way information is reported in their monthly statement, their knowledge of the interest rate of their card, and their expectations on future interest payments. We also asked whether they read the monthly statement and elicited their predictions in regards to paying down their debt in the next one, two, and three months. The questions and mean responses are tabulated in the online Appendix (Table 4).

We wish to highlight a couple of lessons learned from this first survey. First, these clients were uninformed: only 3 percent of clients claimed to know their exact interest rate and 34 percent claimed to know it approximately within 5 percentage points. Second, there is overconfidence in paying down their debt. In response to a direct question, 35 percent of the consumers surveyed claimed to have overestimated their ability to pay down their debt in the previous six months, and in fact, when we contrast their prediction about payments in the next one, two, and three months versus their actual payments in those months, around half got it wrong; interestingly, about three-fourths of these erred on the side of overconfidence, thinking that their debt

[13] A regression of interest rates against deciles of the internal probability of default and months dummies yielded an R^2 of 0.01. The bank indeed accepted that they do not do risk-based-pricing at the moment.

would decrease when in fact it increased.[14] Third, 92 percent of clients claimed to read their monthly statement carefully, which is somewhat surprising given how uninformed they are in regards to their interest rate. Part of the explanation may be that the information is hard to understand.[15] For instance, 42 percent said they would prefer a clearer statement and 38 percent opined default happens because people do not realize how fast they are accumulating debt, and not because of strategic default or due to unforeseen 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 of their ability to pay down 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 received, to measure overconfidence in MTP, as well as to have an idea of how dissatisfied they are with their level of indebtedness and what would be their perceived net benefit of defaulting.[16] We found unawareness and overconfidence are very prevalent; 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 data. Additionally, over four-fifths said they would like to decrease their debt even taking into account the sacrifices this would imply, and over nine-tenths said that defaulting would decrease their welfare taking the benefits of defaulting into account. Unfortunately, we do not have enough power to detect reasonably small impacts of different messages on survey measured outcomes ($<$ 0.1 standard deviation).

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 such a 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" (Rate) and the "months-to-pay message" (MTP),

[14] 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. This evidence lends some support to his conjecture.

[15] The online Appendix presents an example of a monthly statement.

[16] The respondents were distributed among the control group (25.17 percent) and those in the High Debt plus Advice (14.99 percent), Months to Pay (15.94 percent), Interest Rate (25.52 percent), and Warning (18.38 percent) 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 and they decided on the final messages sent.

respectively. Both of these disclosures feature prominently in disclosure mandates in many countries, including the United States.

The second set of messages was inspired by literature that stresses that people are influenced by their peers, either through the signal that this behavior provides, such as in rational herding models, or through a conformity channel. We designed four peer-comparison messages: two of these informed the client as to whether his or her card's debt was above the mean of *similar clients*, that is: clients in a cell of the same gender, similar credit limit (as a proxy for income), age, and predicted PD. We call these their "peer group." These two "high debt" messages differed only in terms of whether advice on how to decrease debt was provided or not.[18] The other two comparison messages presented the client with a thermometer that metaphorically showed their *relative* risk of default as a body temperature, again with respect to clients in their cell of characteristics. One thermometer message was sent to clients with a relatively high probability of default within their cell warning them about it, and the other to clients with a relatively low probability of default within their cell; we call these the high thermometer message (HTM) and the low thermometer message (LTM), respectively.[19]

Finally, we included a message that 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 in regards to 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 2, 3, and 4 show some of the messages; the rest can be found in the online Appendix.

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) versus control (90,015 no messages) was performed within each stratum to provide us with an appropriate control group.

[18] Recently, Chen et al. (2010) showed that individuals care about being below or above average in terms of performance on a task.

[19] This later message was not part of our initial design, but our partner bank included it. Some of these clients had estimated default probabilities close to zero but others had estimated default probabilities above 8 percent. But in each case, the clients that received the "low risk" message had predicted default probabilities below their strata median.

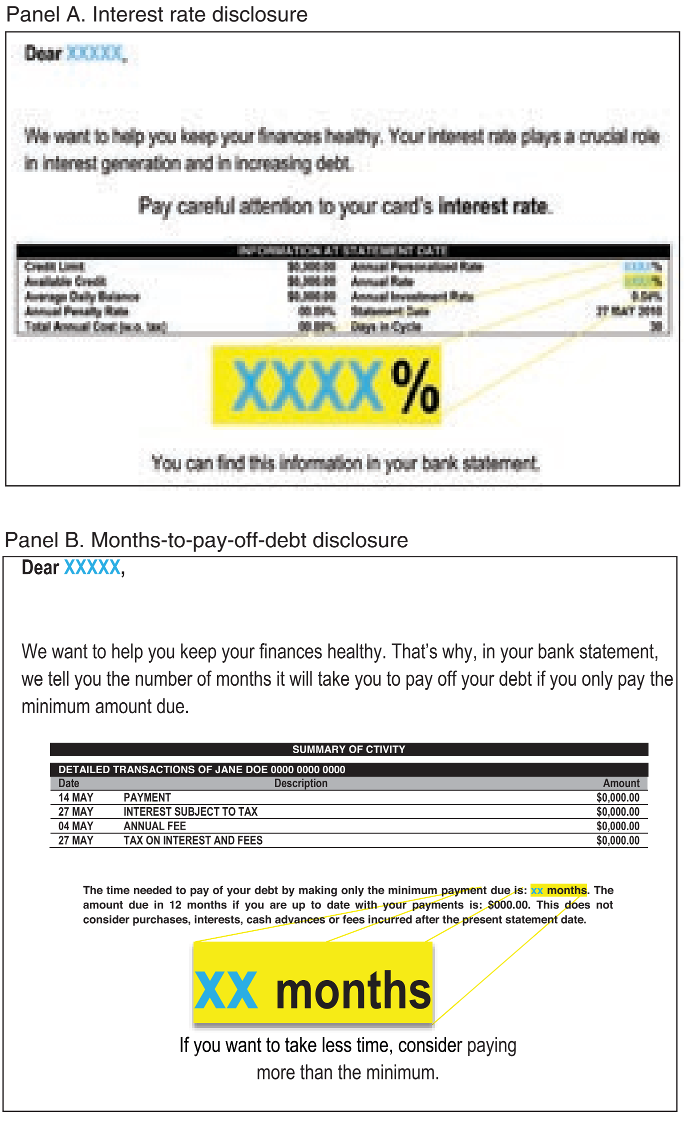


Figure 2. Salient Legally Mandated Messages

Notes: The figures present an English version of the messages sent in the experiment. This is the precise format used (the enlarged boxes were also in the printed message), except that the originals were in Spanish.

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.[20] This meant we had 12,825 clients for each of these messages.[21] Next, within each of the 300 strata, we

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

[21] 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.

Panel A. High debt + advice

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.

Panel B. Warning

Dear XXXXX,

We want to help you keep your finances healthy.

Don't get confident

Paying off a debt is not that easy

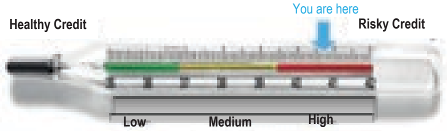
Many studies have found out that consumers overestimate their ability to pay and fail to service their debts.

Don't let it happen to you!

Figure 3. 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 4. High Risk Message

Notes: 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.

identified which unassigned clients had above (below) the median predicted probability of default and randomly allocated these to the "high (low) thermometer" message: 6,444 to the HTM and 6,456 to the LTM. The remaining unassigned treatment clients were randomly allocated within strata as follows: 12,900 for the interest rate and the de-biasing warning message, respectively, and 12,825 for the months-to-pay message; so these three later groups are directly comparable. Table 1 in the online Appendix shows that randomization worked to balance the variables across groups.

At this point, we wish to highlight three 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 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. Third, the LTM is accurate and not deceiving. It is accurate since the message explicitly stated that the information was a relative comparison to people within their cell of similar age, limit, gender, and risk, and this is in fact true.[22] It is not deceiving since the bank defines good credit behavior as not having more than 90 days past due, and by design these were not included in the experiment.[23] LTM recipients compare favorably versus the market in Mexico, as shown in the online Appendix.

The timing of the experiment (shown in Figure 5) 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.

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

[22] *Across* cells it is possible that some clients with larger estimated probability of default than others in other cells received the LTM, but the message did not make across cell comparisons.

[23] It turns out that only 1.4 percent of clients that received the LTM had 90 days or more of payment due at any time during the second semester of 2010; whereas in the paper's entire sample, the analogous number is 8.6 percent. Clients that received the low thermometer message were still profitable for the bank. Unprofitable clients are sent into special loan collection programs, and none of the clients in the experiment were in such programs. We should also say that in the credit card contract the cardholder gives explicit consent for the bank to send messages according to the bank's policies and classifications, and indeed, typically the bank does send information and promotions, often with an RCT design. This experiment carried out by the bank is one such instance. More on this in the online Appendix.

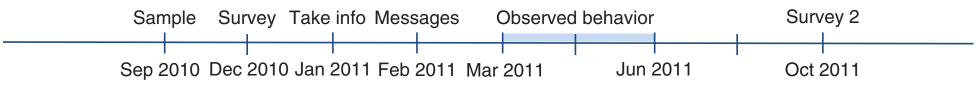


FIGURE 5. EXPERIMENT TIMELINE

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 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, while changes in debt are a proxy for demand responsiveness, both related to competition. Delinquency is related to the stability that the TILA mentions. Later in the paper, we study opening of new credit card accounts for a subsample using system wide Credit Bureau information.

We also report results for a specification that puts both TILA-like messages in one dummy and the five non-TILA messages in another dummy as in equation (2). This is analogous to creating indices as recommended by Anderson (2008). These indices enable us to test the null hypothesis of equality between TILA and non-TILA messages, $\beta_{1t} = \beta_{2t}$ in equation (2) and limit the multiple testing problem.

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

Given that the clients in our sample were highly indebted and at high risk of default, we hypothesized that the "high debt" message, the HTM message, and the warning message would reduce debt and default. We had no strong prior expectation as to their effect on account closures. Since this bank has one of the five 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 estimating the effect of sending the information, which is what TILA-type laws actually mandate, but not necessarily the effect of reading the information, since we do not know if the clients did actually read it. We think this is the treatment effect of the biggest policy interest, as regulations force banks to send information but do not and cannot force cardholders to read it. Having said this, we believe that the information did reach a significant part of the 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. If low salience of the interest rate in monthly statements rationalize the unawareness of the interest rate and the high indebtedness and risk in our population, presumably its salient disclosure could remedy this. Salience in prices has been found to be important in other settings. Chetty, Looney, and Kroft (2009) find that tax elasticities are dependent on how salient taxes are, while Malmendier and Lee (2011) find that overpaying in online auctions depends inversely on the salience of the posted price.

A priori one would like to know first how much could be gained by knowing the card's interest rate, for instance by comparing 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 do not demonstrate) 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, in the spirit of Ponce, Seira, and Zamarripa (forthcoming), we conducted an exercise to measure if consumers debt *allocation* across the cards they have minimizes financing cost. More than 3,000 of our clients had two cards in our bank. On average, the yearly difference in interest rates across these cards was 4 percentage points. 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 10 percentage points 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 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.

Table 2—Baseline Results

	Dependent variables				
	Debt		Delinquent		Closed
	March	April	March	April	June
Panel A					
Mean dependent	17,391	16,541	0.183	0.198	0.043
Standard deviation dependent	(24,425)	(23,964)	(0.387)	(0.398)	(0.204)
Rate	−35	14	0	0	0.001
	(63)	(81)	(0.004)	(0.004)	(0.002)
MTP	43	90	0	0.006	−0.002
	(64)	(83)	(0.004)	(0.004)	(0.002)
HTM	−233	−172	−0.015	−0.006	0.007
	(90)	(118)	(0.005)	(0.005)	(0.003)
LTM	4.4	82	0.014	0.013	0.001
	(84)	(108)	(0.005)	(0.005)	(0.003)
High debt + advice	−29	−127	0.002	0.005	−0.003
	(64)	(83)	(0.004)	(0.004)	(0.00195)
High debt	−104	32.77	−0.002	0	0
	(62)	(81)	(0.004)	(0.004)	(0.002)
Warning	−126	−147	−0.002	−0.002	−0.002
	(62)	(81)	(0.004)	(0.004)	(0.002)
F-test TILA	0.64	0.54	0.99	0.27	0.54
F-test non-TILA	0.04	0.14	0	0.07	0.03
Panel B					
TILA	6	48	−0.001	0.002	−0.001
	(46)	(60)	(0.003)	(0.003)	(0.002)
Non-TILA	−107	−99	−0.004	−0.001	0
	(38)	(50)	(0.003)	(0.002)	(0.001)
F-test	0.03	0.03	0.38	0.31	0.95
Observations	147,634	143,484	167,190	167,190	167,190

Notes: Standard errors are in parentheses. On each panel, each column represents a regression and each row a treatment group dummy. In panel A, each of the variables in 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 five 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 Rate or MTP 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). The previous to last row tests the equality of TILA and non-TILA dummies.

To measure the impact of making the interest rate salient, we sent the interest rate message displayed in Figure 2, panel A, to a randomized treatment group of 12,900 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 statistically 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.

The effect of the interest rate message on debt in March is −35 pesos for March and 14 pesos for April. This is a tiny 0.2 percent of mean debt and is not statistically

Table 3—TILA-Like Disclosures on Subsamples

	Dependent variables					
	Debt		Delinquent		Closed	
	March	April	March	April	June	Observations
Panel A. Interest rate disclosure						
Paid interest[a]	−19	10	−0.002	0	0	52,257
	(70)	(96)	(0.005)	(0.005)	(0.002)	
Interest rate[b]	−32	22	0.003	0.003	0.001	60,947
	(65)	(89)	(0.005)	(0.005)	(0.003)	
Debt[b]	4	161	−0.008	−0.005	0.001	45,744
	(118)	(152)	(0.006)	(0.006)	(0.002)	
Change in absolute value	29	−106	0	−0.005		78,480
	(58)	(76)	(0.004)	(0.004)		
Another credit card[c]	−525	−410	−0.022	−0.024	0.005	7,492
	(194)	(249)	(0.012)	(0.013)	(0.005)	
Panel B. Months to pay off debt disclosure						
Paid interest[a]	−28	−61	0.001	0.011	0	52,436
	(102)	(120)	(0.005)	(0.005)	(0.005)	
Debt[b]	87	185	0.002	0.012	−0.002	45,287
	(174)	(195)	(0.006)	(0.006)	(0.006)	
MTP[b]	26	105	0.001	0.006	−0.003	46,312
	(134)	(141)	(0.004)	(0.004)	(0.004)	

Notes: Standard errors are in parentheses. 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 canceling effects. Panel B shows MTP.

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

[b] Above median in January 2011.

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

different from zero. The effect on delinquency and account closures is also not different from zero and is economically minuscule; it is zero if we only consider the first three decimal points. Given the low interest awareness in our sample, this zero-effect was unexpected. It may be that the average treatment effect masks some heterogeneity. Table 3, panel A, reports estimates for different subsamples where one would expect the message to have effects. The first three rows of Table 3 focuses on clients with high stakes. Row 1 of Table 3 considers clients that typically pay interest (revolved debt for ten consecutive months before receiving the message), while 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 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 performed a very conservative test: 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 cancellations when averaging. If anything, this will overstate any effect. Row 4 presents the results. Again the effect is statistically zero, even when this strategy is biased towards finding an effect. Section IV actually shows that the whole distribution of changes are similar for treatment and control clients.

A third possibility is that clients have limited possibilities to substitute debt across financial products. This does not constitute a methodological problem but an explanation of why there is no effect. The last row of Table 3, panel A, uses only clients that had other credit cards in any other bank, hoping to find larger debt elasticities.[24] 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 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. 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. 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 derived from paying the minimum are worrisome since there is evidence that actual payment anchors on the minimum payments (e.g., Stewart 2009). Figure 6 plots a histogram of actual MTP calculated from the administrative data versus MTP as reported by the clients themselves (two versions of the same histograms can be found in the online Appendix with different conditioning restrictions). Clearly many clients are grossly underestimating the amount of months to pay off their debt.

Due to statistics such as the ones mentioned above, the 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 12,825 messages with the personalized number of months to pay. The exact design is shown Figure 2, panel 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.

[24] We obtained this information from the Credit Bureau. Unfortunately, since the Credit Bureau does not have price information, we do not observe which card is more expensive.

[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 $N = -\frac{\ln\left(1 - \frac{Debt \times MonthlyInterestrate}{MinPay}\right)}{\ln(1 + MonthlyInterestrate)}$. When this formula became indeterminate without having missing inputs or Minimum Payment equal to zero, we concluded the individual would never be able to pay off the debt following that scheme. These clients were told that they would take more than 100 months.

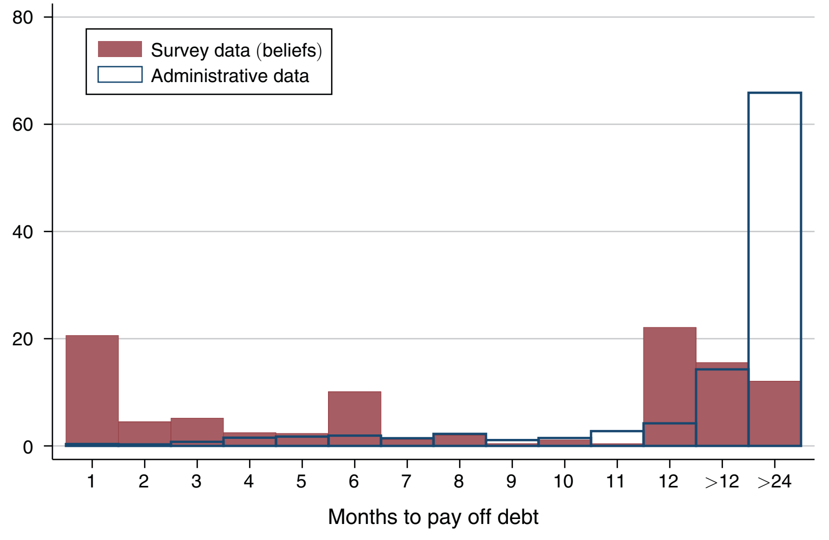


Figure 6. Months to Pay Off Debt (*beliefs versus data*)

Note: The figure presents the distribution of the number of months individuals believe will take them to pay off their debt if they pay only the minimum and make no further purchases against the actual number calculated with the administrative data.

Contrary to our expectation, the message had no effect on debt. 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. Since the message does not affect income or debt, 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 were 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 today.

C. *Peer-Group Comparisons*

While TILA-type laws have concentrated on disclosing information about contract terms, there are many other types of information that may be useful for cardholders. We thought that information that allows the cardholder to situate herself vis-à-vis similar people might be of help.[26] Social comparison messages have been effective in many contexts, inducing participation in elections, encouraging contributions to

[26] There are several channels through which peer comparisons could induce behavioral change. One could imagine an environment in which individuals with similar preferences are subject to common but only partially observable 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 could push the individual toward performing a similar action to that of his peers. Alternatively, there may exist a tendency toward conformity directly in the utility function. We do not attempt to distinguish between these forces.

online public goods, and generating electricity savings. We test their potential for reducing debt and default.

As explained in more detail above, we told some clients if they had more debt or more risk compared to *similar clients*.[27] We conjectured that these messages would decrease debt and delinquency. The bank included a message (the LTM) that congratulated the client for his or her behavior and showed a thermometer in low temperature when they were below the median PD of their cell.

Figures 3 and 4 display five of the messages sent, the other two can be found 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 found significant effects.

Row 3 of Table 2 reports the average impact of the HTM and row 4 the respective estimate for the LTM. The effects across these two messages move in opposite directions and give more confidence to our causal interpretation. The HTM caused a decrease in debt of 233 pesos in March and a decrease in delinquency of 1.5 percentage points, about 8 percent of mean delinquency. The LTM increased delinquency by 7 percent of its respective mean in both March and April. The amount of payments actually went down by about 10 percent of its mean (unreported in Table 2).[28] Voluntary account closures increased for clients receiving the HTM versus their controls by 0.7 percentage points (about 16 percent of mean value).

Rows 5 and 6 of Table 2 show that the high debt message reduced debt although in a small amount (about 0.05 percentage points from the mean at most), which is not statistically distinguishable from zero. 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. Neither of the two messages influenced delinquency. There is a decrease of 0.3 percentage points in voluntary account closures for those that received the advice, significant at the 10 percent level.

[27] One could argue that to really test peer effects one would need a comparison message with no reference to "similar clients" whatsoever so that we could cleanly separate the two contents: the high debt content versus the high debt *with respect to peers* content. The problem with this is that without a reference group it is not clear what we mean by higher debt (higher than what?), and we did not send this comparison message. The message as it was implemented does not allow us to say whether the nature of the response would be different with a different (larger) reference group.

[28] The bank actually lost money from the message they themselves proposed, which they did not expect. Note, however, that consumer welfare could have increased from getting more information.

D. *De-biasing Warning Message*

We also tested the warning message shown in Figure 3, panel 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. The message 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 "de-biasing" individuals by warning them of the bias in answers to hypothetical valuation questions can help them 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, by −126 pesos on average in March and −147 in April. Effects on delinquency have negative signs but are not statistically significant at conventional levels.

E. *TILA-Like versus 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 a joint *F*-test of the bottom of panel A in Table 2, which shows that although we cannot reject all 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 the results of the index regressions (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.

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 17,815 of our accounts from the Credit Bureau to measure these market level spillovers.[29] We study three system-wide outcome variables: openings of new card accounts in other banks (a proxy for switching), the credit score, 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 a small decrease in credit score as a result of the LTM.

The main lesson of the paper is that *all* treatments have zero or tiny effects in *all* outcomes measured. And this contrasts with the huge emphasis that policy has given to disclosures. Even if nonstandard 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, which was the motivation for TILA.

[29] These are distributed across treatments as follows: 603 of HTM; 746 of LTM; 1,467 of 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.

Table 4—Effects on Other Credit Conditions (*subsample*)

	Dependent variable		
	Score June	Opens new card March–June	Number of credit cards in default June
Panel A			
Mean dependent	622	0.008	0.813
Standard deviation dependent	(67)	(0.091)	(1.77)
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)
HTM	−1.09	−0.002	−0.073
	(2.8)	(0.004)	(0.075)
LTM	−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)
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
Observations	17,781	17,781	17,801

Notes: Standard errors are in parentheses. 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. In panel A, each of the variables in 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 five 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).

IV. Multiple Testing, Heterogeneity, Power, and External Validity

A. *Multiple Testing Issues*

Conducting multiple hypothesis tests for each treatment separately and claiming that a treatment worked if its coefficient was statistically significant may be misleading. It is well known that the probability of falsely rejecting a null hypothesis increases with the number of tests performed. We carried out five exercises to deal with this issue.

Table 5—Placebo Tests

	Dependent variables				Dependent variables (subsample)		
	Debt		Delinquent		Credit score	Open cards	Credit cards in default
	September 2010	October 2010	September 2010	October 2010	June 2010	March–June 2010	June 2010
Panel A							
Mean dependent	18,919	18,937	0.135	0.145	642	0.07	0.293
SD dependent	(25,800)	(25,727)	(0.341)	(0.352)	(50)	(0.254)	(0.969)
Rate	54	47	0	−0.003	1.291	0.001	−0.016
	(117)	(120)	(0.003)	(0.003)	(1.44)	(0.007)	(0.029)
MTP	59	11	0	−0.003	−0.151	−0.001	−0.03
	(116)	(120)	(0.003)	(0.003)	(1.45)	(0.007)	(0.028)
HTM	−48	11	0.004	0.002	−0.351	−0.003	−0.025
	(163)	(168)	(0.004)	(0.004)	(2.1)	(0.012)	(0.042)
LTM	139	129	0.001	−0.001	0.238	−0.008	0.009
	(162)	(166)	(0.004)	(0.004)	(1.89)	(0.001)	(0.037)
High debt + advice	22	48	0	0.003	−0.161	−0.003	0.005
	(119)	(122)	(0.003)	(0.003)	(1.38)	(0.007)	(0.028)
High debt	−40	−74	0.002	0.002	3.44	−0.007	−0.051
	(119)	(122)	(0.003)	(0.003)	(1.42)	(0.007)	(0.028)
Warning	70	98	−0.001	−0.001	−0.671	0.006	−0.035
	(116)	(120)	(0.003)	(0.003)	(1.43)	(0.007)	(0.028)
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)	(0.002)	(0.002)	(1.07)	(0.005)	(0.021)
Non-TILA	2	17	0.001	0.001	0.674	−0.001	−0.027
	(71)	(73)	(0.002)	(0.002)	(0.86)	(0.004)	(0.017)
F-test	0.65	0.98	0.74	0.09	0.94	0.8	0.9
Observations	165,042	163,113	167,190	167,190	17,077	17,815	17,815

Notes: Standard errors are in parentheses. This table replicates the estimations on Table 2 for the months of September and October 2010 (pretreatment). Each column represents a regression. In panel A, each of the variables in 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 five 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.

To address the problem, we first implemented a *joint F*-test where the null hypothesis is that *all* TILA or non-TILA coefficients in the regression are zero (reported in panel A of Table 2). The second excercise, in panel B of Table 2, consisted of using an index methodology, as in Kling, Liebman, and Katz (2007) and Anderson (2008), which aggregates different outcome variables in a single score.[30] The index aggregates all TILA treatments into a single dummy variable, and does the analogous

[30] Anderson (2008) remarks that "Summary index tests … have three advantages over individual outcomes: First, they are robust to overtesting because each index represents a single test … Second, they provide a statistical test for whether a program has a "general effect" on a set of outcomes. Finally, they are potentially more powerful than individual-level tests."

for non-TILA. The results point in the same direction: the non-TILA messages are more effective than the TILA messages for debt and delinquency, with the effects being significant for debt in March and April and for delinquency in March. Third, in Table 5, we conduct placebo tests using the same specification and the same data where the placebo treatment is defined as if treatment happened at different placebo months. If the results we found before in Table 2 were solely due to sampling variance, then we should expect a similar number of the placebo treatments to be statistically significant in this exercise. Table 5 shows that this is not the case. That is, there is something specially different across the treatment and control arms *exactly* when the treatments were sent.

As a fourth exercise, in Table A8 of the online Appendix, we computed Bonferroni *p*-values.[31] With these conservative *p*-values, only the non-TILA results in Table 2, panel B, remain significant at the 10 percent level, and only for debt in March. This could well be the result of a substantial decrease in power, however. We ran power simulations using our data and the Bonferroni adjustment, and power decreases by 20 percent to 53 percent. The actual power with Bonferroni tests is much less than the standard 80 percent to 90 percent accepted in the literature. To be less conservative on type I error and reduce type II error, our last exercise uses the False-Discovery-Rate (FDR) Benjamini and Hochberg (1995) procedure, which controls for the expected proportion of falsely rejected hypothesis and is less taxing on statistical power than Bonferroni. Using the FDR *p*-values we reach a similar conclusion as above: the coefficients that were individually significant at a 1 percent level in Table 2 are still significant at a 10 percent level, while none is in the placebo table; some non-TILA effects are still significant while no TILA effect is. All in all, differences between TILA and non-TILA still remain after controlling for multiple testing, in spite of the large decrease in power and its effects being economically modest.

One may be interested in null hypotheses different from the zero effect null. The online Appendix (Figure A8) provides confidence intervals for the equations estimated in panel A of Table 2. We report both intervals accounting and not accounting for multiple testing to inform the interested reader about the effect sizes that can be rejected.[32]

B. *Response Heterogeneity*

Section III reported zero average effects for the interest rate disclosure. Table 3 performed checks that suggested this was not driven by responses canceling

[31] For Table 2, we consider the family to be of size 35 and for panel B of 10. Bonferroni *p*-values have some caveats though. They are overly conservative as the significance level is *at most* α but is generally less and this comes at a high cost in terms of power. They assume tests are independent, which penalizes power even more. In our case, this is potentially a large penalty since we are running the same test for different months, so there should be substantial correlation across tests. In the limit, if the correlation is high enough, then not accounting for the family-wise significance level is fine since they are basically the same test.

[32] The confidence statement provided by the separate confidence intervals should be carefully interpreted; it is important to remember that a confidence interval is just a random set that covers the true parameter with a certain level of *confidence*. The Bonferroni intervals control for the *overall* confidence level so the reader can postulate a joint null hypothesis and judge whether or not it would be rejected in our data.

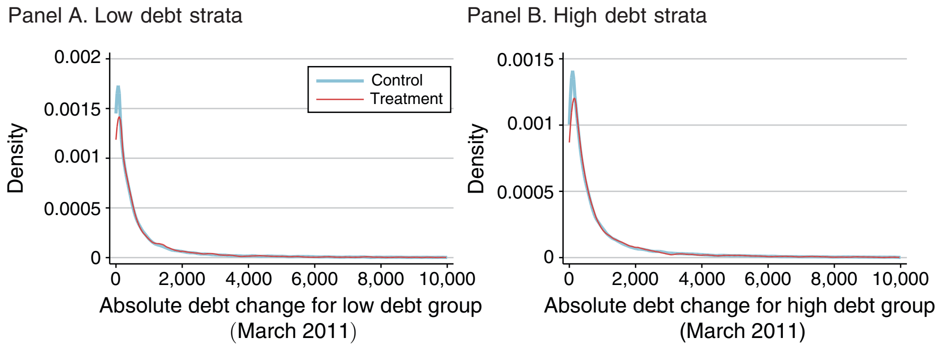


FIGURE 7. ABSOLUTE CHANGE IN DEBT

Notes: These graphs report the density for the absolute change in debt from February to March 2011. We divided the sample between the high and low debt strata to compare the appropriate treatment and control groups.

each other out in the average. Here we plot the distribution of treatment effects non-parametrically. In particular, for each cardholder we calculate the change in debt from February 2011 to March 2011 and calculate the absolute value of those changes. If there was more heterogeneity in the interest rate treatment than in its control group, then the treatment distribution should be shifted to the right. Figure 7 shows that they are not. A Kolmogorov-Smirnoff test cannot reject the equality of distributions.

We did find some heterogeneity related to the predicted riskiness of clients. Recall that the CNBV wanted to issue rules mandating the use of predicted risk as a trigger of messages. The CNBV expected larger responses for risky clients. It turns out that this expectation is correct. 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 8 plots the coefficients of the interactions. It turns out that all the messages had statistically significant effects for the highest PD quintile.

Finally, Tables 10 and 11 in the online Appendix focus on heterogeneity by splitting samples across the number of products with the bank (as a proxy for loyalty or hold-up problems) and categories of income. Some noteworthy results are that the effect on closings when the HTM is received is not present when the cardholder has several products with the bank, and that 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 too much consumption.

Putting the magnitudes of the coefficients into perspective, we interpret this as 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 very different subsubsets of the population.

C. *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

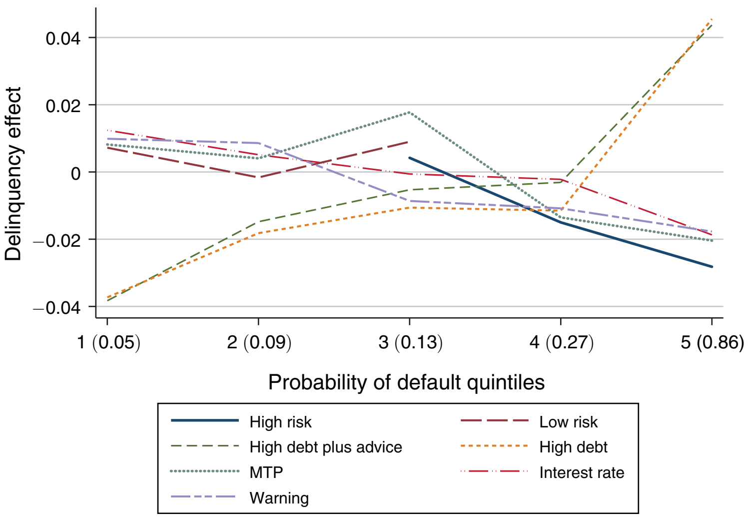


Figure 8. Effects on Delinquency by Probability of Default

Notes: 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: $Y_{ijt} = \alpha_t + \sum_{j=1}^{7} \beta_{tj} T_{ij} + \sum_{r=2}^{5} \gamma_{tr} D_{tir} + \sum_{k=2}^{5} \sum_{j=1}^{7} \delta_{tj} D_{tir} \times T_{ij} + S_{ik} + \epsilon_{ijy}$. The average probability of default of each quintile is expressed in parentheses on the x-axis.

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. That is, we use the same sample and the same methodology as that employed with the real treatments, just two months before treatment.[33] Figure 9 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 0.006 percentage points in delinquency (4 percent of its mean) with 80 percent power. We believe these are relatively small effects and the bank agrees.

As mentioned above, we ran a specification (placebo) test by estimating equation (1) in the same partition of control and treatment cards for the months of September and October 2010, i.e., before treatment. If the equation was misspecified, we would expect more than 5 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.

D. *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

[33] Other months worked similarly.

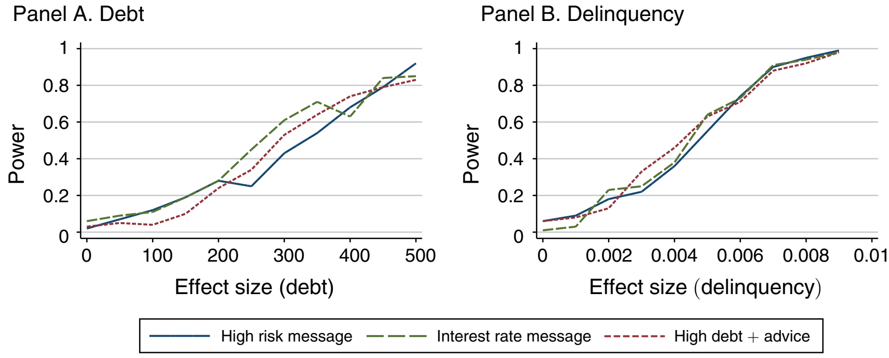


Figure 9. Statistical Power

Notes: 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.

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 such as going to the Internet and checking prices 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). This experiment goes "out of sample" in many ways that address the concerns above: it was carried out at two different banks from the one in this paper;[34] it was representative of all the banks' clients not just the risky ones; the price message was more aggressive as it involved direct price comparisons across banks (see Figure 10);[35] and the frequency of the message was varied randomly. One group of clients in each of two large banks received the price comparison of Figure 10 in April 2012 once, another group (the frequent treatment group) received it *monthly* seven times from April 2012 to October 2012, while a third group acted as a control group and received no messages at all during all of 2012. Group size was approximately 30,000 clients per arm, representative of the banks' overall population of cardholders.

To analyze the resulting behavior, we estimated equation (1) by ordinary least squares separately for each bank and month. Figure 11 plots the estimated βs, scaled down by the average of the dependent variable for ease of interpretation. Panel A

[34] With these two additional banks, the population covered in this paper includes close to 70 percent of Mexico's total cardholders.

[35] Note that this is a very aggressive mandate. It is uncommon to force companies to advertise the prices of their competitors in their own monthly statements.

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 10. APR Comparison Message

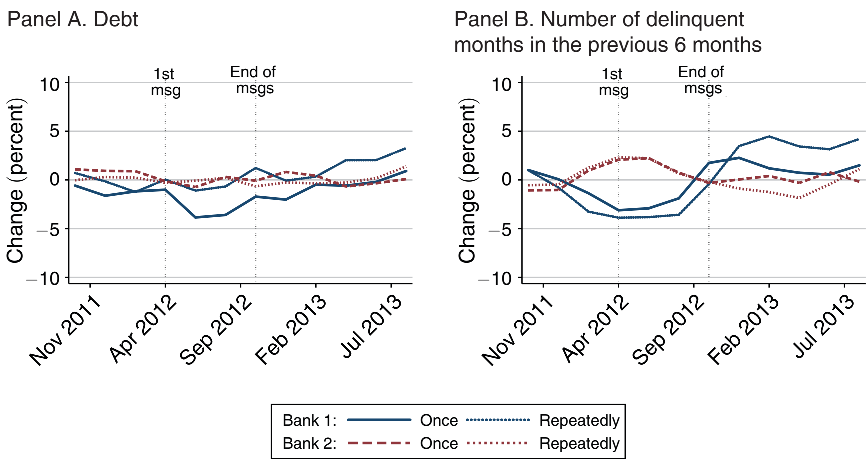


Figure 11. Effects at Two Different Banks

Notes: 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.

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 (always less than 5 percent of the mean and, more often than not, less than 1 percent of the mean of the dependent variable) for all banks, all periods, and all frequencies. None of the coefficients were statistically different from zero at the 10 percent level. Importantly for us, we tested 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 (Section 3), we estimate nonexperimentally the effect of a policy that for the first time forced banks to include in their own monthly statement information of their ranking (in terms of interest rates) vis-à-vis other banks. This is a nice complement since this comparative information was disclosed for the first time then, and it is 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 regulatory response was to disclose more information. Many governments around the world are still mindlessly increasing disclosure requirements as a consumer protection device in the credit card market. The content of the mandated disclosures, the new and the old, have been determined mostly on the basis of introspection, not on the basis of 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 then, only to a small degree.

Small effects do not imply money-losing effects. We estimated equation (1) using profits[36] as the dependent variable. The HTM increased profits by 285 pesos (p-value $= 0.001$), while the high debt message increased profits 93 pesos (p-value 0.09). Given that printing and sending the letter cost 2.5 pesos, these messages were 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.

Finally, this study does not rule out the possibility that other messages could have greater effects, or that effects are the same in all countries. But we conjecture that they won't have transformational effects. Some of the messages that we wanted to try but could not were ones where the client is told how much money other clients with the same credit score and debt are paying at other banks, with an estimate of how much savings *in pesos* they would make per year by switching to the cheapest 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

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. 2015. "Regulating Consumer Financial Products: Evidence from Credit Cards." *Quarterly Journal of Economics* 130 (1): 111–64.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

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

[36] We defined monthly revenue as income from interest and fees in the month and we sum that value for March and April discounting the second month with the interbank rate TIIE. We then deducted the expected losses from default using the bank's definition.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society* 57 (1): 289–300.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. 2010. "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." *Quarterly Journal of Economics* 125 (1): 263–306.

Bertrand, Marianne, and Adair Morse. 2011. "Information Disclosure, Cognitive Biases, and Payday Borrowing." *Journal of Finance* 66 (6): 1865–93.

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

Chen, Yan, F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens." *American Economic Review* 100 (4): 1358–98.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145–77.

Cummings, Ronald G., and Laura O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89 (3): 649–65.

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

Durkin, Thomas A. 1975. "Awareness of Credit Terms: Review and New Evidence." *Journal of Business* 48 (2): 253–63.

Durkin, Thomas A., and Gregory E. Elliehausen. 2011. *Truth in Lending: Theory, History, and a Way Forward*. Oxford: Oxford University Press.

Ferman, Bruno. 2011. "Reading the Fine Print: Credit Demand and Information Disclosure in Brazil." PhD diss. Massachusetts Institute of Technology.

Festinger, Leon. 1954. "A Theory of Social Comparison Processes." *Human Relations* 7 (2): 117–40.

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

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics* 125 (2): 515–48.

Jin, Ginger Zhe, and Phillip Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *Quarterly Journal of Economics* 118 (2): 409–51.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee C. Vermeulen, and Marian V. Wrobel. 2012. "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics* 127 (1): 199–235.

Kroszner, Randall S. 2007. "Creating More Effective Consumer Disclosures." Speech, George Washington University School of Business, Financial Services Research Program Policy Forum, Washington, DC, May 23, 2007. https://www.federalreserve.gov/newsevents/speech/kroszner20070523a.htm.

Lacko, James M., and Janis K. Pappalardo. 2007. *Improving Consumer Mortgage Disclosures: An Empirical Assessment of Current and Prototype Disclosure Forms*. Federal Trade Commission Bureau of Economics. Washington, DC, June.

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

Malmendier, Ulrike, and Young Han Lee. 2011. "The Bidder's Curse." *American Economic Review* 101 (2): 749–87.

Melzer, Brian T. 2011. "The Real Costs of Credit Access: Evidence from the Payday Lending Market." *Quarterly Journal of Economics* 126 (1): 517–55.

Negrín, Jose Luis, and Enrique Seira. 2014. "The effects of price comparison messages." Unpublished.

Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa. Forthcoming. "Borrowing on the Wrong Credit Card? Evidence from Mexico." *American Economic Review*.

Seira, Enrique, Alan Elizondo, and Eduardo Laguna-Müggenberg. 2017. "Are Information Disclosures Effective? Evidence from the Credit Card Market: Dataset." *American Economic Journal: Economic Policy*. https://doi.org/10.1257/pol.20140404.

Shay, Robert P., and Milton W. Schober. 1973. *Consumer Awareness of Annual Percentage Rates of Charge in Consumer Installment Credit: Before and After Truth-in-Lending Became Effective*. US

National Commission on Consumer Finance Technical Studies. Washington, DC: Government Printing Office.

Soll, Jack B., Ralph L. Keeney, and Richard P. Larrick. 2013. "Consumer Misunderstanding of Credit Card Use, Payments and Debt: Causes and Solutions." *Journal of Public Policy and Marketing* 32 (1): 66–81.

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.

Stango, Victor, and Jonathan Zinman. 2013. "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees." https://www.dartmouth.edu/~jzinman/Papers/SZ_LimitedAttention_Overdrafts_2013_4.pdf.

Stango, Victor, and Jonathan Zinman. 2016. "Borrowing High vs. Borrowing Higher: Price Dispersion and Shopping Behavior in the U.S. Credit Card Market." *Review of Financial Studies* 29 (4): 979–1006.

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

Truth in Lending Act. 1968. *15 U.S. Code § 1601* (*a*). https://www.law.cornell.edu/uscode/text/15/1601.

White, Michelle J. 2007. "Bankruptcy Reform and Credit Cards." *Journal of Economic Perspectives* 21 (4): 175–200.

Williams, Julie L. 2005. "Remarks by Julie L. Williams Acting Comptroller of the Currency Before Women in Housing and Finance and The Exchequer Club." Speech, Washington, DC, January 12, 2005. https://www.occ.gov/news-issuances/speeches/2005/pub-speech-2005-1.pdf.

Woodward, Susan E., and Robert E. Hall. 2012. "Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence." *American Economic Review* 102 (7): 3249–76.