



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Cash or Card? Evidence from a Large-Scale Field Experiment to Increase Consumers' Debit Card Use

Ray Charles Howard, Michael King, Paolina C. Medina, Roland Umanan



TEP Working Paper No. 1826

August 2026

Cash or Card? Evidence from a Large-Scale Field Experiment to Increase Consumers' Debit Card Use*

Ray Charles Howard[†] Michael King[‡] Paolina C. Medina[§]
Roland Umanan[¶]

Abstract

Cash remains a dominant payment method in many markets despite the widespread availability of digital alternatives such as debit cards, creating costs for consumers, firms, and governments. This research examines which interventions most effectively increase debit-card use, how effectiveness varies with incentive form and amount, and whether average effects conceal meaningful differences across customers. A large-scale randomized field experiment with approximately 1.5 million customers of a Mexican bank tests cash and gift-card incentives of 60 and 300 MXN, repeated benefit-focused nudges, and combinations of nudges with cash incentives. Estimated treatment effects on debit-card transaction frequency and spending range from approximately 0% to 5%, with larger incentives producing less-than-proportional gains in transaction frequency and incentive-form effects varying across outcomes and incentive amounts. Causal-forest estimates show substantial treatment-effect heterogeneity: for example, the effect of the 300 MXN cash incentive is concentrated among customers with relatively high baseline debit-card use, while some near-zero average effects mask offsetting positive and negative subgroup responses. Nudges have small nonsignificant effects alone and no detectable incremental effect when combined with cash incentives. These findings show that effective payment-activation strategies depend not only on intervention design but also on which customers receive them.

JEL Codes: C93, D12, D14, D18, G21, G51

Keywords: Financial Inclusion, Debit Cards, Cash and In-kind Incentives, Nudges, Spillover Effects

*For helpful comments and valuable feedback, we thank participants at the Trinity College Dublin Economics internal seminar series. We also thank Jose Antonio Murillo for invaluable support to run the experiment, as well as Sofia Terrazas and Pedro Guerrero for outstanding research assistance on preliminary analysis.

[†]University of Virginia.

[‡]Trinity College Dublin, University of Dublin.

[§]University of Houston.

[¶]Trinity College Dublin, University of Dublin.

1 Introduction

Digital payment solutions such as debit cards have proliferated in recent years (Statista, 2023). However, cash is still the dominant payment method for billions of people around the world. Globally, consumer cash payments are estimated to total approximately \$26 trillion annually (Bruno et al., 2024), and cash remains the most widely used point-of-sale payment method in dozens of countries, even where digital payment options are widely accepted (Chaudhuri et al., 2022). This continued reliance on cash creates financial and social costs for businesses and governments because cash is expensive to transport and store securely, necessitates in-person transactions, and can facilitate criminal activity (Perez, 2020). Shifting consumer spending toward digital payment alternatives such as debit cards can also have meaningful benefits for consumers (Bachas et al., 2021), retailers (Higgins, 2024), and the macroeconomy (World Bank, 2012; Zandi et al., 2013), with particular relevance for low- and middle-income countries.

Banks, merchants, and governments therefore have strong incentives to increase consumer use of digital payments, and debit cards are a common starting point because of their simplicity. Financial incentives to increase debit card use are one common strategy, but managers face consequential choices about how those incentives should be designed. Cash rewards are intuitively attractive because they offer consumers flexibility, but a gift card tied to a popular brand may be especially salient or attractive. Managers must also determine how large an incentive needs to be to generate meaningful behavior change without making the program unnecessarily costly. Lower-cost behavioral interventions provide yet another option: firms can use SMS nudges to communicate the security, convenience, or budgeting benefits of digital payments rather than directly paying consumers to change their behavior. Finally, even an intervention that works well on average may be substantially more effective for some consumers than others, creating an opportunity to improve effectiveness through targeting.

These considerations raise three related questions. First, what type of financial incentive should managers offer, and how much should they offer? Standard economic models imply that a more fungible incentive such as cash should be preferred to a less fungible gift card, whereas research on mental accounting shows that consumers do not always treat economically equivalent resources as interchangeable (Reinholtz et al., 2015; Shaffer and Arkes, 2009; Thaler, 1985, 1999). Similarly, increasing incentive amount should generally increase its motivational value, but diminishing sensitivity implies that each additional unit of reward may provide less incremental psychological value than the previous one (Kahneman and Tversky, 1979). Second, can lower-cost behavioral nudges meaningfully change payment behavior, either on their own or when combined with financial incentives? Third, how much does responsiveness to these interventions vary across consumers? Most marketing interventions are evaluated based on average treatment effects, but substantial treatment-effect heterogeneity can make these averages misleading (Chen et al., 2020; Greenberg et al., 2023; Mrkva et al.,

2021; Shah et al., 2023).

We examine these questions in a large-scale randomized field experiment conducted in partnership with a Mexican bank. Approximately 1.5 million customers were randomly assigned to receive financial incentives for increasing their debit-card use, repeated benefit-focused nudges, combinations of nudges with cash incentives, or neither. The incentive treatments vary both incentive form—cash versus gift card—and incentive amount—60 versus 300 MXN—allowing us to compare a fully fungible cash reward with a less fungible gift-card reward and to examine how treatment effects scale with incentive size. (At the time of the study, 300 MXN was worth roughly as much in Mexico as \$30 USD was worth in the USA, and 300 MXN was approximately 3% of the median participant’s monthly income.) The nudges emphasize the security, convenience, or budgeting benefits of debit cards relative to cash and were delivered repeatedly over several weeks. Our primary outcomes are debit-card transaction frequency and amount spent using a debit card, and we also observe a broader measure of spending for a large subset of participants. In addition, the bank provides approximately 120 pretreatment covariates for each customer, allowing us to use causal forests to estimate how treatment effects vary across customers (Athey and Wager, 2019).

The average treatment effects provide several insights. Estimated effects on debit-card transaction frequency and spending range from approximately 0% to 5%, but their magnitude depends on both intervention design and the behavioral outcome. At the higher incentive amount, the 300 MXN cash condition produces larger estimated effects than the 300 MXN gift-card condition for both transaction frequency and amount spent. At the lower amount, however, the pattern differs: the 60 MXN gift-card condition produces a larger estimated effect than the 60 MXN cash condition for amount spent, while the effects on transaction frequency are similar in magnitude. Incentive amount also matters differently across outcomes. Although the 300 MXN incentives are five times as large as the 60 MXN incentives, their effects on transaction frequency are only about two to three times as large. For amount spent, the 300 MXN cash effect is substantially larger than the 60 MXN cash effect, whereas the 300 MXN and 60 MXN gift-card effects are nearly identical. Finally, the repeated nudges have nonsignificant effects when delivered alone and no detectable incremental effect when combined with cash incentives.

These average effects, however, conceal substantial treatment-effect heterogeneity. Causal-forest estimates indicate that consumers differ meaningfully in their responsiveness to the interventions, sometimes in ways that change the interpretation of the average effect. For example, the positive average effect of the 300 MXN cash incentive is concentrated among customers with relatively high baseline debit-card use. In other conditions, positive estimated effects among some customers are offset by negative estimated effects among others, producing average effects close to zero. Thus, an intervention that appears weak when evaluated only by its average effect may still be effective for an identifiable subset of consumers.

The patterns across treatment conditions also provide insight into how incentive design affects behavior. The stronger performance of cash at the higher incentive amount is consistent with the fungibility principle in classical models of consumer choice: a cash reward can be allocated across uses, whereas a gift card is more restricted. The absence of a consistent cash advantage at the lower amount is consistent with behavioral research documenting violations of fungibility and suggests that the relative effectiveness of reward forms can depend on both reward magnitude and the behavioral outcome being changed (Reinholtz et al., 2015; Shaffer and Arkes, 2009; Thaler, 1985, 1999). Likewise, the less-than-proportional increase in transaction frequency as incentive size increases is consistent with diminishing sensitivity to reward value (Kahneman and Tversky, 1979), although this pattern does not extend uniformly to amount spent. The weak effects of repeated nudges further suggest that communicating the benefits of debit cards may be insufficient to change an established payment behavior, particularly when compared with direct financial incentives. Finally, although prior research provides reasons to expect that moving consumers from cash toward digital payments could increase overall spending (Agarwal et al., 2024; Hirschman, 1979; Prelec and Loewenstein, 1998), we find no detectable evidence of such an increase in our broader spending measure.

These findings make three primary contributions to research on consumer behavior and marketing interventions. First, we contribute to research on incentive and reward design by showing that the effectiveness of financial incentives depends jointly on incentive form, incentive amount, and the behavioral outcome marketers seek to change. More fungible and larger rewards do not uniformly generate proportionally larger behavioral effects, underscoring the importance of considering reward architecture rather than incentive value alone. Second, we contribute to research comparing financial incentives with behavioral interventions. Repeated benefit-focused nudges produce limited behavior change in this setting and do not measurably augment cash incentives, identifying an important context in which inexpensive informational interventions are not substitutes for financial rewards. Third, we contribute to research on treatment-effect heterogeneity and personalized marketing by showing that modest average effects can mask substantial differences in responsiveness across consumers (Chen et al., 2020; Greenberg et al., 2023; Mrkva et al., 2021; Shah et al., 2023). Taken together, the findings show that effective payment-activation strategies depend not only on which intervention managers choose and how it is designed, but also on which consumers receive it.

2 Institutional Setting

To conduct our study we partnered with one of the largest commercial banks in Mexico. An important feature of our study is that the target customers are people who already have a bank account and debit card, but their debit card usage rate is low. (See ‘Experimental Design and Descriptive Statistics’ for detailed information on baseline debit card usage rates.)

Following [Soman \(2015\)](#), these individuals can be conceptualized as ‘last mile’ customers: they are already aware of and in possession of the focal product, but their use of the product remains limited. This segment represented over 15% of the bank’s approximately 9 million customers at the time of the study, and the available evidence suggests the meaningful size of this segment is not unique to the bank: although the share of adults in developing economies making or receiving digital payments increased from 35% to 57% in 2021, 1.6 billion adults with an account continued to make merchant payments exclusively in cash ([World Bank, 2022](#)).

The goal of our study was to increase customer debit card use, which provides several important benefits for financial institutions. First, debit cards generate fee-based revenue through interchange fees—per-transaction charges paid by the acquiring bank to the issuing bank. Importantly, a bank’s fee-based revenue rises with both the frequency of debit card transactions and, up to a point, with the transaction amounts. (In Mexico, interchange fees for debit card transactions at points of sale can be as high as 1.15% of the purchase amount. At the time of the study, the maximum interchange fee per transaction was 13.50 MXN.) Thus, our study examines two dependent variables: (1) the number of debit card transactions a customer makes, and (2) the total amount they spend using their debit card.

Second, debit card usage is a critical growth driver for banks because it serves as an entry point to other banking products and services. Customer debit card use is central to several cross-selling opportunities, ranging from short-term credit to long-term savings and retirement products ([Visa, 2022](#)). More broadly, encouraging customers to use their debit cards fosters everyday engagement with the bank, increasing the likelihood that they will obtain additional financial products (e.g., mortgages or personal loans), thereby enhancing overall revenue.

Third, greater reliance on debit cards reduces customers’ need to withdraw cash from ATMs, lowering the operational costs of servicing and replenishing those machines. Fourth, higher debit card usage tends to raise customers’ average account balances, since they are less likely to withdraw cash in advance for transactions. Larger balances provide banks with a more stable, low-cost funding base to support lending activities. Finally, debit card transactions also create float deposits due to the settlement lag between purchase authorization and final clearing. These short-term, no-cost balances can be invested by the bank in low-risk instruments, generating incremental income and improving liquidity management.

3 Experiment Design

We partnered with one of the largest commercial banks in Mexico, to conduct a large-scale field experiment testing incentive- and information-based interventions designed to increase debit card use. The study involved all bank customers who had: (i) a valid mobile phone

number to receive SMS messages, (ii) an average daily checking account balance of at least 50 MXN over the two months prior to the study; and (iii) had a debit card but exhibited low debit card use as identified by the bank based on a proprietary measure provided by Mastercard, which takes into account the number of merchant categories transacted with and the total number of transactions. In total, 1,526,177 bank customers met these criteria and were included in the experiment. We conducted the main experiment during November 2019, with a preliminary notification message sent to incentive-based treatment participants in late October. We also selected a random subset of treatment participants to continue receiving their assigned treatment during December so we could examine persistence effects. The full study timeline is illustrated in Figure 1.

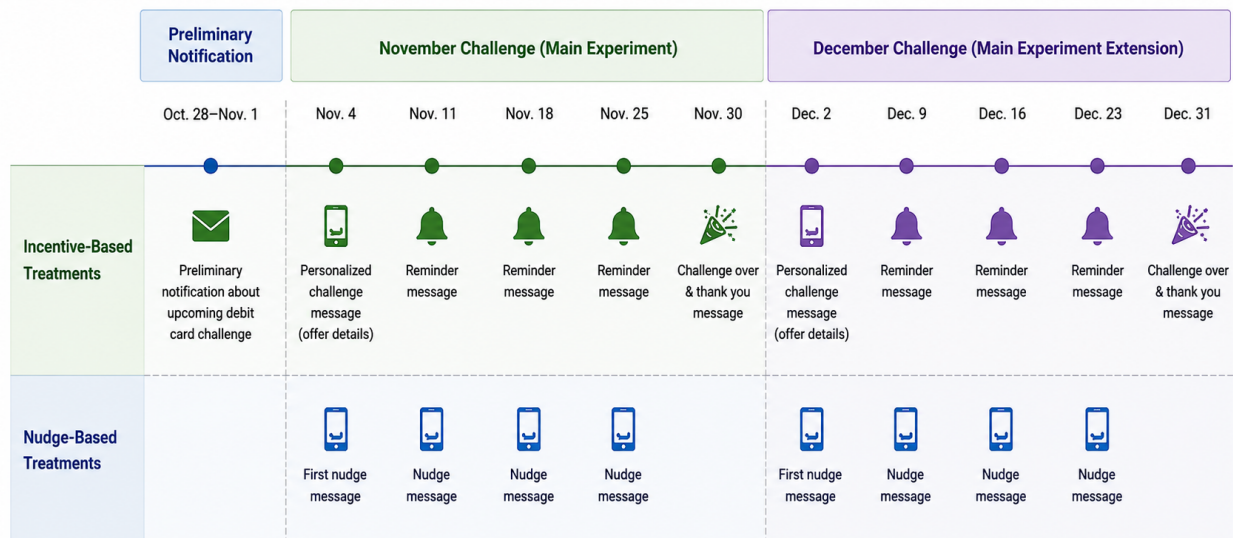


Figure 1: Study timeline.

As illustrated in Figure 2, the experiment had four branches: control, incentives, nudges, and incentives + nudges. We used stratified random assignment to improve covariate balance across experimental conditions. Strata were defined by the full cross-classification of gender, credit-card ownership, payroll-account status, bank tenure above versus below the median, baseline debit-card transaction level, age quartile, and income quartile, yielding 759 populated experimental strata. Customers were randomized to treatment conditions within each stratum. The decision to over-sample the four incentive-based treatment conditions relative to the other treatment conditions reflects their high priority to the bank. We next describe each branch of the experiment in detail.

Control Group. Following the logic of randomized waitlist designs used in resource-constrained field experiments (Hirshleifer et al., 2016; McKenzie, 2025), customers assigned to the control condition were randomly assigned to one of two subgroups: a “pure control group” ($n = 208,205$), which was guaranteed to remain untreated, and a “reserve control group” ($n = 718,339$), from which customers would have been randomly selected for treat-

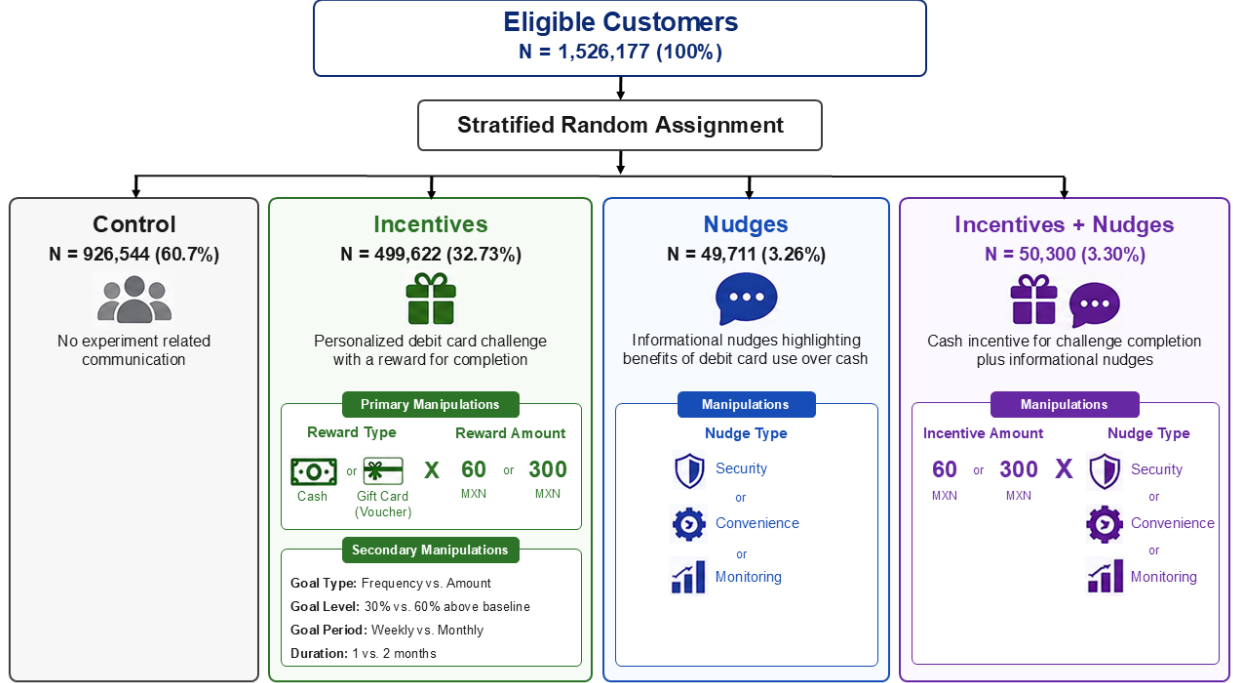


Figure 2: Experimental design.

ment if additional funds became available. Because no additional funds were secured, all customers in the reserve control group remained untreated. An unused randomized waitlist effectively constitutes an additional control stratum (McKenzie, 2022); accordingly, we pool the two control subgroups in our main analysis to increase statistical power given our cost constraints (McConnell and Vera-Hernández, 2015; Duflo et al., 2007). Neither subgroup received any experiment-related communications during the study.

Incentive-Based Treatments. Participants in the incentives branch were offered a reward for completing a personalized debit card challenge. The two primary manipulations in this branch were incentive value (300 MXN vs 60 MXN) and incentive type (cash vs gift card). Cross-randomizing these variables resulted in four incentive-based treatment conditions of primary interest, in which the rewards were 300 MXN cash, a 300 MXN gift card to Liverpool (a high-end department store chain), 60 MXN cash, or a 60 MXN 2-for-1 movie ticket gift certificate to Cinépolis (a large movie theatre chain). At the time of the study, 300 MXN [60 MXN] was worth roughly as much in Mexico as \$30 USD [\$6 USD] was worth in the USA.

We manipulated incentive *amount* to test whether a relatively large incentive (300 MXN) is necessary to induce behavior change, or if a more modest incentive (60 MXN) is sufficient. (300 MXN was approximately 3% of the median participant’s monthly income, and 60 MXN was approximately 0.6%.) We manipulated incentive *type* for two reasons. First, from a theoretical perspective, consumers should prefer cash over vouchers because the former is more fungible than the latter. Thus, if behavior reflects preferences, cash rewards should

create more behavior change than gift cards of equivalent value. However, there are notable examples of consumers violating the fungibility principle (Heath and Soll, 1996; Reinhardt et al., 2015; Thaler, 1985, 1999), and hedonic non-cash incentives like the gift cards in our study are sometimes preferred to cash incentives (Shaffer and Arkes, 2009). Thus, it is an open question as to which incentive type will have a larger effect in this context. Second, from a practical perspective, gift cards are cheaper for the bank than equivalent cash rewards because they can be purchased in bulk, at discounted prices.

The debit card challenge was communicated to participants via SMS messages that informed them they would receive their reward if they hit a personalized debit card spending goal based on their debit card activity over the 12 months before the study began. The challenge itself involved four secondary manipulations related to the spending goal: goal type (number of debit card transactions vs total amount spent), goal level (30% vs 60% increase above baseline),¹ goal period (weekly vs monthly), and challenge duration (November only vs November and December). We manipulated goal type because a transaction-based goal requires consumers to repeat the behavior multiple times, receiving frequent feedback and cumulative progress markers, whereas an amount-based goal can often be met with a single large purchase. Thus, it is possible that the former leads to larger behavior change by creating a habit (Wood and Neal, 2009), even though the latter is commonly used by practitioners. We manipulated goal level because goals motivate behavior (Bagozzi and Dholakia, 1999), and we were interested in understanding whether a higher or lower goal leads to greater behavior change. We manipulated goal period because breaking a goal down into smaller chunks can motivate progress toward the goal (Sharif and Woolley, 2020), so it could be the case that four weekly goals create more behavior change than one equivalent monthly goal. We manipulated challenge duration to be able to examine persistence effects.

All six variables in this branch of the experiment (incentive value, incentive type, goal type, goal level, goal period, and challenge duration) were cross-randomized. Because each variable has two levels, this resulted in $2^6 = 64$ total conditions. However, we had no *a priori* hypotheses regarding interactions between our primary and secondary manipulations, so we treat this branch of the experiment as having 12 conditions in our analysis: the four primary treatment conditions related to incentive amount and type, and the eight secondary conditions related to goals.

Information-Based Treatments. Participants in the information branch of the experiment were randomly assigned to one of three ‘nudge’ conditions: security, convenience, or expense monitoring benefits. The nudges were sent via SMS messages every Monday for the duration of the study. English translations of the messages in each condition include:

¹For example, if a participant who spent 500MXN per month over the 12 months preceding the study was assigned to the 30% (60%) increase condition, their goal would be to spend 650MXN (800MXN) in the next month. The minimum goal in the goal amount conditions was 400MXN per month in the 30% condition and 800MXN per month in the 60% condition. The minimum goal in the number of transaction conditions was 4 transactions per month in the 30% condition and 8 transactions per month in the 60% condition.

- Security condition: Thinking about going to the ATM? It is better to pay for your purchases today with your [bank name] debit card and stay safer without carrying cash.
- Convenience condition: Thinking about going to the ATM? It is better to avoid payday lines by paying for your purchases today with your [bank name] debit card.
- Monitoring condition: Thinking about going to the ATM? It is better to track your spending on your statement by paying for your purchases with your [bank name] debit card.

We chose these specific themes because security concerns regarding cash can shift consumer payment behavior toward digital options (Kosse, 2010; Schuh and Stavins, 2010), convenience is a catalyst for behavior change (Thaler and Sunstein, 2008) that can affect which payment method people use (Brown et al., 2022), and expense monitoring can help consumers manage their spending (Heath and Soll, 1996; Lukas and Howard, 2023) and people value payment methods that facilitate such monitoring (Borzekowski et al., 2008; Schuh and Stavins, 2010).

Information + Incentives. Finally, we cross-randomized the three information nudges (security vs convenience vs monitoring) with the two levels of cash incentive value (300 MXN vs 60 MXN) to generate the six information + cash incentive conditions illustrated in Figure 2. (Note: Incentive type was not manipulated in this arm of the experiment—all incentives were cash. Goals were also not manipulated—for all participants in this branch, goal type was transactions, goal level was 30%, goal period was weekly, and challenge duration was November and December.) We did this to determine if the potential benefits of incentives and nudges might be greater than the sum of their parts. Thus, participants in this condition received SMS messages containing their personalized debit card challenge and nudges emphasizing the benefits of debit card use as compared to cash.

Table 1 presents the distribution of participants across our focal conditions.

Table 1: Distribution of control and treatment groups

	N	Percentage
Control	926,544	60.71
300MXN cash	125,008	8.19
60MXN cash	124,614	8.17
300MXN Liverpool voucher	125,004	8.19
60MXN 2-for-1 movie ticket	124,996	8.19
Security SMS	16,572	1.09
Convenience SMS	16,567	1.09
Monitoring SMS	16,572	1.09
Security SMS + 300MXN cash	8,482	0.56
Security SMS + 60MXN cash	8,290	0.54
Convenience SMS + 300MXN cash	8,480	0.56
Convenience SMS + 60MXN cash	8,277	0.54
Monitoring SMS + 300MXN cash	8,486	0.56
Monitoring SMS + 60MXN cash	8,285	0.54
Total	1,526,177	100.00

This table presents the distribution of participants across control and treatment groups.

4 Summary Statistics

The bank provided account-level pre-treatment data for each customer in the sample covering the six months before the experiment. There are 120 pretreatment variables in total, including socio-demographic characteristics, debit card and credit card use, and variables indicating the customers’ relationship with the bank. Table 2 reports summary statistics for an illustrative subset of these variables. Baseline debit card use is low, confirming the sample consists of ‘last mile’ customers. Credit card use is also low, implying the sample predominately uses cash. The average participant has been a customer of the bank’s for over six years, suggesting a relatively high degree of familiarity. Finally, most participants have a ‘payroll’ account with the bank. In Mexico, a payroll account is a type of deposit account offered by banks in partnership with companies, specifically designed for employees to receive their salary payments. Since the bank has direct access to payroll information, this makes the data on salaries in the payroll group very accurate. We over-sampled *non*-payroll clients at the bank’s request. To obtain estimates representative of the bank’s debit-card customer population, we weight observations by the inverse of their sampling probabilities (Solon et al., 2015).

We formally test if the randomization process was successful in balancing these variables across the control and treatment groups. Appendix table B.1 shows that only gender is unbalanced. However, the normalized differences for gender across experimental arms range from -.002 to .001, which indicates adequate balance because they are well below the .25

threshold suggested in the literature (McKenzie, 2017). Nonetheless, we control for the gender imbalance in our analysis.

Table 2: Summary statistics

	Mean	Median	Std. Dev.	Min	Max
<i>Socio-demographic characteristics</i>					
Age (years)	45.89	44.00	15.65	20.00	85.00
Female (=1)	0.42	0.00	0.49	0.00	1.00
At least college-educated (=1)	0.10	0.00	0.30	0.00	1.00
Monthly income (MXN)	14,805.46	9,924.88	16,550.97	2,333.88	113,996.51
<i>Debit card use</i>					
Amount of debit card transactions (MXN)	505.03	34.08	1,530.58	0.00	11,479.53
Number of debit card transactions	1.56	0.00	3.19	0.00	20.00
Amount of withdrawals (MXN)	44,059.83	9,397.98	158,923.13	0.00	1,269,981.88
Number of withdrawals	6.88	5.00	7.07	0.00	42.00
<i>Credit card use</i>					
With partner-bank credit card (=1)	0.21	0.00	0.40	0.00	1.00
Partner-bank credit card limit (MXN)	9,880.10	0.00	35,193.50	0.00	244,000.00
Partner-bank credit card tenure (months)	7.67	0.00	25.42	0.00	157.00
Partner-bank credit card closing balance (MXN)	2,853.76	0.00	11,279.48	0.00	80,359.06
Partner-bank credit card interest paid (MXN)	45.55	0.00	221.63	0.00	1,665.05
Number of partner-bank credit card transactions	0.71	0.00	3.28	0.00	24.00
<i>Relationship with bank</i>					
Tenure with partner-bank (months)	91.58	74.60	73.71	1.83	282.50
Non-payroll account (=1)	0.21	0.00	0.41	0.00	1.00
Observations	1,526,177				

This table reports the summary statistics. Variables with (=1) are dummy variables. All monetary variables are in Mexican Peso (MXN). All continuous variables are winsorized at 1th and 99th percentiles. The experimental pool of 1,526,177 customers is a random sample drawn from the universe of the bank's customers satisfying three requirements: had a valid mobile phone number to receive SMS messages; with at least 50 MXN average daily balance in the checking account over the two months prior to the intervention; and identified by the bank as having low debit card use based on a proprietary measure provided by Mastercard, which takes into account the number of merchant categories transacted with and the total number of transactions. The statistics are computed using monthly data at the customer level covering six months prior to the intervention. Monthly income are salary transfers flagged as such in the payroll accounts. Partner-bank credit card tenure is the number of months since the customer's first credit card was activated. Partner-bank credit card closing balance is the statement balance for a given billing cycle. Partner-bank credit card interest paid is the amount of interest paid for a given billing cycle. Tenure with bank is the number of months since the first checking or payroll account was opened. Non-payroll account is a dummy variable equal to 1 if the customer does not have a payroll account with partner-bank, and equal to zero if the customer has a payroll account. Sampling weights were applied due to the oversampling of customers from the non-payroll group. In 2019, 1 USD = 9.67 MXN (PPP adjusted).

5 Methodology

We analyze the experiment in five ways. First, to document the impact of all experimental treatments on the number and amount of debit card transactions during the main experiment (i.e., November 2019), we estimate

$$Y_i = \alpha_s + \sum_{k=1}^{10} \beta_k \times T_{k,i} + X_i + \epsilon_i \quad (1)$$

where Y is the outcome variable for customer i at month t , α_s and ω_t denote strata and month fixed effects, $T_{k,i}$ represents the indicator variable for whether customer i was assigned to level (or category) k within a given treatment, β_k represents the average treatment effect of treatment k , X_i refers to the baseline value of the outcome variable for customer i , and $\epsilon_{i,t}$ is the error term. We also control for gender because it is unbalanced across conditions.

Standard errors are heteroskedasticity-robust. We correct for multiple hypothesis testing and report both sharpened q-values (Anderson, 2008) and Romano-Wolf stepdown adjusted p-values (Clarke, 2021). The adjusted p-values from the Romano-Wolf method are based on 1,000 bootstrap replications.

Second, we use causal forests (Athey and Wager, 2019) to examine treatment effect heterogeneity in the main experiment. Causal forests use machine learning to predict individual level treatment effects based on observable pre-treatment characteristics. We train causal forests for each of the core treatments (incentive type, incentive amount, and nudges) and, for each treatment, we split individuals into quartiles of predicted treatment effects (similar to the approach of Chernozhukov et al., 2018). We then calculate the average treatment effects on each quartile and present the characteristics of the average individual in each quartile.

Third, we examine whether increasing debit card use during the main experiment affected other types of spending. Fourth, we combine the main experiment data from November with the additional data from December to investigate whether treatment effects persist from one month to the next. Finally, we conduct a simulation that explores optimal targeting strategies given the firm’s budget constraint.

6 Results

6.1 Main Experiment Results

Figure 3 presents the average treatment effects for the main experiment, estimated using equation 1. Panel A shows results for the number of debit card transactions made, and Panel B shows results for the total amount spent using a debit card.

The effect of incentives on number of debit card transactions. As compared to control, the number of debit card transactions is 4.0% higher in the 300 MXN cash condition ($\beta = 0.040$, $SE = .004$, $p < .010$), 2.1% higher in the 300 MXN Liverpool voucher condition ($\beta = 0.021$, $SE = .005$, $p < .010$), 1.2% higher in the 60 MXN cash condition ($\beta = 0.012$, $SE = .005$, $p < .010$), and 0.9% higher in the 60 MXN Cinopolis voucher condition ($\beta = 0.009$, $SE = .005$, $p < .100$). The mean number of debit card transactions in the control group is 1.82; thus, the most successful intervention—300 MXN cash—increased the number of transaction by an average of approximately 0.07 transactions per customer. If scaled up across the entire sample of approximately 1.5 million customers, this would net roughly 105,000 additional debit card transactions for the bank in one month.

With respect to the effect of *incentive type* on number of debit card transactions, the treatment effect for 300 MXN cash is approximately twice as large as for the 300 MXN voucher, but the treatment effects for 60 MXN cash and 60 MXN voucher are quite similar.

This offers mixed evidence as to whether a more fungible incentive generates more behavior change, and suggests that value is not the only voucher attribute that matters to consumers.

With respect to the effect of *incentive level* on number of debit card transactions, the treatment effect for 300 MXN cash is approximately three times larger than for 60 MXN cash, and the treatment effect for the 300 MXN voucher is approximately two times larger than for the 60 MXN voucher. Thus, increasing the value of each incentive type fivefold, from 60 MXN to 300 MXN, resulted in effects sizes only two to threefold larger. This is consistent with diminishing sensitivity to reward value (Kahneman and Tversky, 1979).

The effect of incentives on amount of debit card spending. Generally speaking, the effect of incentives on the amount of debit card transactions is weaker than for the number of debit card transactions. As compared to control, the total amount of debit card spending was 3.0% higher in the 300 MXN cash condition ($\beta = 0.030$, $SE = .007$, $p < .010$), 1.4% in the 300 MXN voucher condition ($\beta = 0.014$, $SE = .007$, $p < .100$), 0.5% in the 60 MXN cash condition ($\beta = 0.005$, $SE = .007$, $p > .100$), and 1.6% in the 60 MXN voucher condition ($\beta = 0.016$, $SE = .007$, $p < .050$). The mean amount of debit card spending in the control group is 1,029.14 MXN; thus, the most successful intervention—300 MXN cash—increased the average amount spent by approximately 30.87 MXN per customer. If scaled up across the entire sample of approximately 1.5 million customers, this would result in approximately 45 million MXN more debit card spending in one month.

With respect to the effect of *incentive type* on the amount of debit card spending, the treatment effect for 300 MXN cash is approximately twice as large as for the 300 MXN voucher, but the treatment effect for the 60 MXN voucher is roughly three times larger than for 60 MXN cash. In other words, cash creates more behavior change than the voucher at 300 MXN but the voucher creates more behavior change than cash at 60 MXN. This again offers mixed evidence as to whether a more fungible incentive generates more behavior change, and further suggests that value is not the only voucher attribute that matters to consumers.

With respect to the effect of *incentive level* on the amount of debit card spending, the treatment effect for 300 MXN cash is approximately six times larger than for 60 MXN cash, and the treatment effect for the 300 MXN voucher and the 60 MXN voucher are roughly the same. Thus, the effects for cash rewards *increasing* sensitivity to reward value between 60 and 300 MXN, and the effects for vouchers suggest essentially no sensitivity. Viewed in light of the results for number of transactions, this suggests that the effect of incentive level on behavior depends partly on the behavior itself: in this case, results on the extensive margin (i.e., whether to transact) are uniformly consistent with prospect theory but results on the intensive margin (i.e., how much to spend) are not.

Nudges. As compared to control, none of the nudges produced statistically significant increases in number of transactions. Amount of debit card spending is 3.2% higher in the

security nudge condition than in the control ($\beta = 0.032$, $SE = .019$, $p < .100$), but the convenience and monitoring conditions did not produce statistically significant treatment effects.

Nudges + Cash. As compared to control, the Security nudge + 300 MXN cash intervention increased the number of debit card transactions by 4.8% ($\beta = 0.048$, $SE = .017$, $p < .010$). This is the largest average effect size in the study, but the confidence interval includes the point estimate for 300 MXN cash, which suggests that adding the Security nudge to the 300 MXN cash incentive did not create a meaningfully larger effect than offering a 300 MXN cash incentive alone. However, we offer this interpretation cautiously because the smaller sample size in the Nudge + Cash conditions contributes to noticeably wider confidence intervals. The Convenience nudge + 300 MXN cash did not significantly increase the number of debit card transactions vs control ($\beta = 0.014$, $SE = .017$, $p > .100$), and the effect size is noticeably smaller than in the 300 MXN cash condition (1.4% vs 4.0%), suggesting that some nudges may actually detract from the positive effect of cash incentives. The Monitoring nudge + 300 MXN cash increased transactions by 3.9% ($\beta = 0.039$, $SE = .017$, $p < .050$) vs control. This is virtually the same effect size as for 300 MXN cash, suggesting that adding the nudge to the cash incentive had no meaningful effect. In sum, these findings indicate that adding a nudge to the 300 MXN cash incentive did not reliably change the results as compared to offering only the 300 MXN cash incentive.

The results for the nudge + 60 MXN cash conditions are more promising. Recall that the average effect of the 60 MXN cash incentive on number of debit card transaction is 1.2%. In contrast, the average effect of Security + 60 MXN cash is 3.0% ($\beta = 0.030$, $SE = .016$, $p < .100$), the average effect of Convenience + 60 MXN cash is 3.9% ($\beta = 0.039$, $SE = .017$, $p < .050$), and the average effect of Monitoring + 60 MXN is 3.6% ($\beta = 0.036$, $SE = .017$, $p < .050$) vs control. Thus, the average effect sizes in the nudge + 60 MXN cash conditions are 2.5 to 3.25 times larger than in the 60 MXN cash condition. The wide confidence intervals prevent firm conclusions about statistical significance, but one implication of these findings is that adding a nudge to a low-value incentive like 60 MXN may have greater marginal benefit than adding a nudge to a high-value incentive like 300 MXN. The results and conclusions for total amount of debit card spending are substantively the same.

Duration. Comparing the one month and two month conditions (i.e., whether someone was randomly assigned to participate only in November or in November and December) is a placebo test because the main experiment results in Figure 3 are for November only, and a consumer’s behavior in November cannot be affected by whether they are randomly assigned to be treated again at the start of December. Thus, it is assuring that the confidence interval for each of these conditions includes the point estimate of the other.

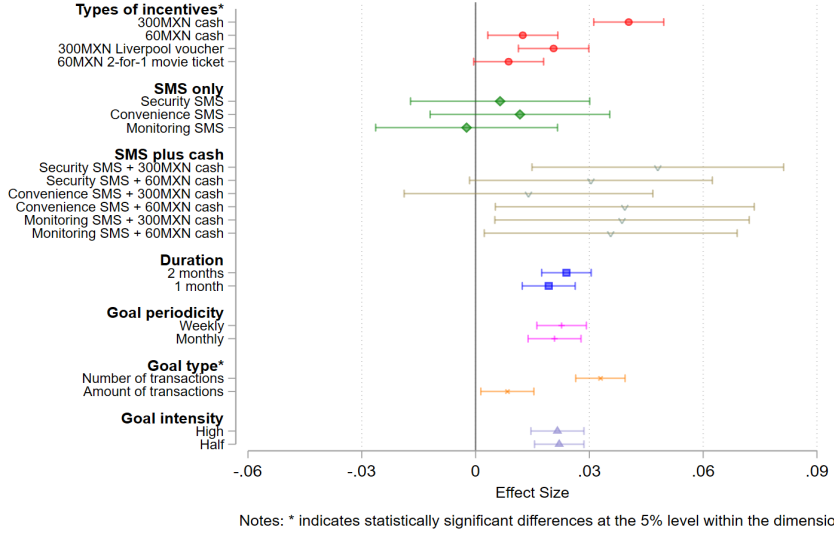
Goal period. As compared to control, the number of debit card transactions was 2.3% higher in the week condition ($\beta = 0.023$, $SE = .003$, $p < .010$) and 2.1% higher in the month

condition ($\beta = 0.021$, $SE = .004$, $p < .010$). The similarity of these effect sizes suggests that number of transactions was not meaningfully affected by whether the spending goal was framed on a weekly or monthly basis. Amount of debit card spending was 2.1% higher than control in the week condition ($\beta = 0.021$, $SE = .005$, $p < .010$) and 1.1% higher in the month condition ($\beta = 0.011$, $SE = .006$, $p < .100$). The difference in these effect sizes suggests that the total amount of debit card spending may have been more strongly affected by framing the personal challenge as weekly rather than monthly.

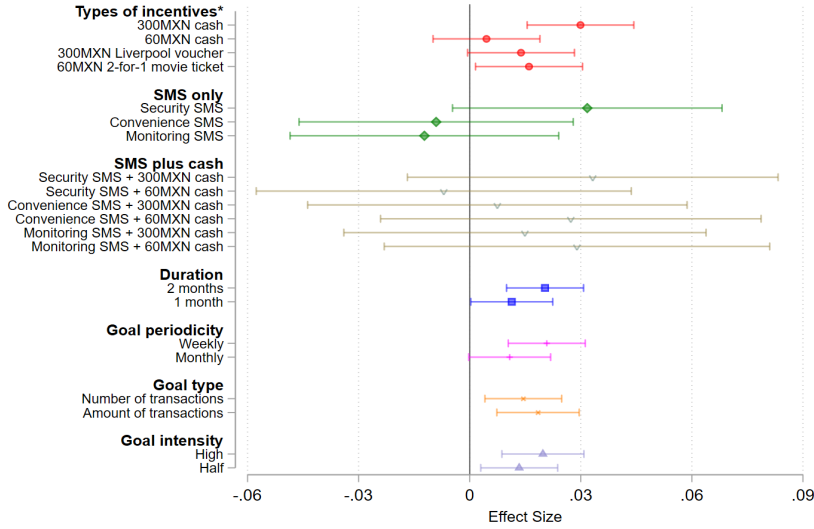
Goal type. As compared to control, the number of debit card transactions is 3.3% higher in the number of transactions goal condition ($\beta = 0.033$, $SE = .003$, $p < .010$) and 0.8% higher in the amount of debit card spending goal condition ($\beta = 0.008$, $SE = .004$, $p < .050$). The large relative difference between these effect sizes suggests that goal-behavior alignment is a winning strategy when the target behavior is number of transactions. Interestingly, the same is not true for total amount of debit card spending. As compared to control, total amount of debit card spending is 1.4% higher in the number of transactions goal condition ($\beta = 0.014$, $SE = .005$, $p < .010$) and 1.8% higher in the amount goal condition ($\beta = 0.018$, $SE = .006$, $p < .010$). The relative similarity between these effect sizes suggests that goal-behavior alignment may not be consequential when the target behavior is total amount of debit card spending.

Goal intensity. The average effect on transactions and amount spent in the goal intensity conditions is significantly higher than in the control ($ps < .050$) but not substantially different between the high and half conditions. Thus, the goal levels we tested did not meaningfully impact our focal outcomes.

Figure 3: Main Experiment Results: Average Treatment Effects



(a) Number of debit card transactions



(b) Amount of debit card spending

Figure 3 shows the average treatment effects of the primary experimental treatments on the number of debit card transactions (panel 3a) and the total amount of debit card spending (panel 3b). The items in bold denote the variables and the items listed below them are the levels (or categories) associated with each variable. The point estimates in both coefficient plots are estimated using Poisson regression (equation 1) and the whiskers depict 95% confidence interval. The x-axis represents the fractional change in outcome variable relative to the control group. The mean amount and number of debit card transactions in the control group are 1,029.14 MXN and 1.82 transactions, respectively.

6.2 Heterogeneous Treatment Effects

We next study treatment effect heterogeneity using causal forests (Athey et al., 2019). We train fourteen models corresponding to the seven core treatments (300 MXN cashback, 60

MXN cashback, Liverpool voucher, 2-for-1 movie tickets, SMS security, SMS convenience, and SMS monitoring) and two outcome variables: the amount and number of debit card transactions.

For this part of the analysis we restrict the sample to individuals in the payroll group. We do so for three reasons. The first reason is technical. Because the non-payroll group was over-sampled, the probability of assignment to each treatment differs between payroll and non-payroll customers. However, payroll status is also a strong predictor of the outcomes we study and, as a result, pooling the two groups would let the forest split on artificial differences in the precision of its local estimates rather than on differences in treatment effects (Athey and Wager, 2019); Appendix A explains this in more detail. The second and third reasons are substantive. Individuals in the payroll group receive their payroll at our partner bank, and we therefore have a fuller picture of their financial situation, including more accurate measures of income and tenure with employer. Furthermore, the payroll group is the largest segment of the bank’s client base, and the one in which the bank reported seeing the most potential for a deeper relationship. In addition, we use only the “minimum control” subset of the control group. We do so to prevent treatment propensities significantly below 10%, which would otherwise negatively affect the performance (precision and sensitivity to specification) of the treatment effect estimates used for model training (Crump et al., 2009). All performance results are calculated on a held-out evaluation sample that was not used at any stage of model training.

Table 3 reports, for each treatment and outcome, the Rank-Weighted Average Treatment Effect (RATE) of Yadlowsky et al. (2021), computed on a held-out evaluation sample. The RATE measures how much larger the treatment effect is among the customers the forest ranks highest than among all customers, averaged over every possible cutoff. A positive and statistically significant RATE means that the ranking identifies customers who respond more than average, so that the heterogeneity the forest predicts could be used for targeting. A RATE that is not distinguishable from zero means that the ranking does no better than ordering customers at random. Treatment effects may still differ across customers in that case; but the forest is not able to predict those differences. The ranking is significant for all seven treatments on the number of transactions, and for the 300 MXN cashback, the 60 MXN cashback, and SMS monitoring on the amount (Table 3). Appendix A.2 describes computation and inference.

Figures 4 and 5 sort consumers in the held-out evaluation sample into quartiles of the predicted treatment effect and report, within each quartile, the difference in mean outcomes between treated and control customers. For the amount of debit card transactions, the top quartile has the largest point estimate and the bottom quartile the smallest in all seven treatments, and the two middle quartiles are close to zero throughout (Figure 4). For the 300 MXN cashback, the effect in the top quartile is approximately 100 MXN and statistically different from zero, and the estimates for the other three quartiles are indistinguishable from

Table 3: Rank-Weighted Average Treatment Effect (RATE)

Treatment	RATE (Transactions)	RATE (Amount)
300 MXN cashback	0.06** (0.02)	53.11** (22.07)
60 MXN cashback	0.05** (0.02)	44.42** (20.82)
Liverpool voucher	0.06** (0.03)	9.37 (22.17)
2-for-1 movie tickets	0.04** (0.02)	40.72 (24.84)
SMS Security	0.11* (0.06)	79.87 (49.52)
SMS Convenience	0.15** (0.07)	55.34 (40.94)
SMS Monitoring	0.11* (0.06)	112.64** (51.17)
Std. errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.		

This table reports Rank-Weighted Average Treatment Effect (RATE) estimates targeting the Area Under the Targeting Operator Characteristic curve (AUROC) (Yadlowsky et al., 2021), computed on the held-out evaluation sample. The columns correspond to the two outcomes, the number and the amount of debit card transactions. Each cell reports the RATE estimate with its standard error in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

zero (panel a). The average effect of this treatment reported in Table 3 therefore comes from one quarter of customers. For the SMS monitoring message, the bottom quartile reduces spending by 328 MXN, with a confidence interval from -515 to -142 MXN, while the top quartile has an estimate of 83 MXN that is not statistically distinguishable from zero (panel g). For the remaining five treatments the top-quartile estimates range from 35 to 147 MXN and the bottom-quartile estimates from -91 to -20 MXN, but none is significant at the 5% level.

The pattern for the number of transactions is similar but less pronounced (Figure 5). The top quartile has the largest point estimate in six of the seven treatments, and the effect there is 0.18 transactions for the 300 MXN cashback, 0.12 for the Liverpool voucher, and 0.26 for the SMS convenience message, each significant at the 5% level. The bottom quartile of the SMS monitoring message reduces the number of transactions by 0.37, and no other quartile in any treatment has an effect distinguishable from zero.

Figure 4: Average Treatment Effect by Quartiles of Predicted Treatment Effect on Transaction Amount

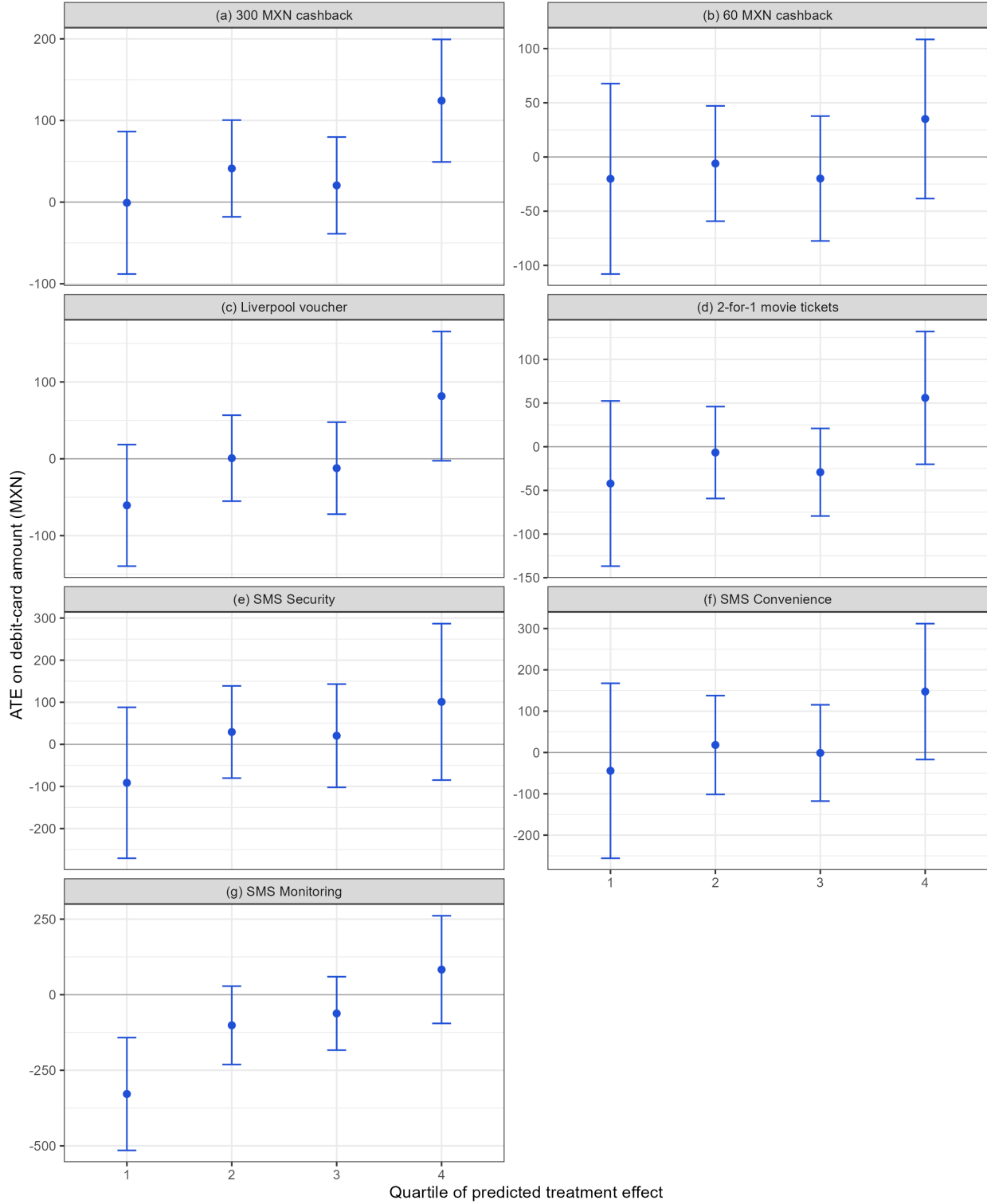


Figure 4 shows the average treatment effect on transaction amount by quartile of predicted treatment effect, computed on the held-out evaluation sample. The x-axis divides the sample into quartiles of the predicted treatment effect, where Q1 (Q4) is the quartile with the lowest (highest) predicted effect. The y-axis plots the difference-in-averages estimate of the average treatment effect within each quartile, along with its 95% confidence interval.

Figure 5: Average Treatment Effect by Quartiles of Predicted Treatment Effect on Number of Transactions

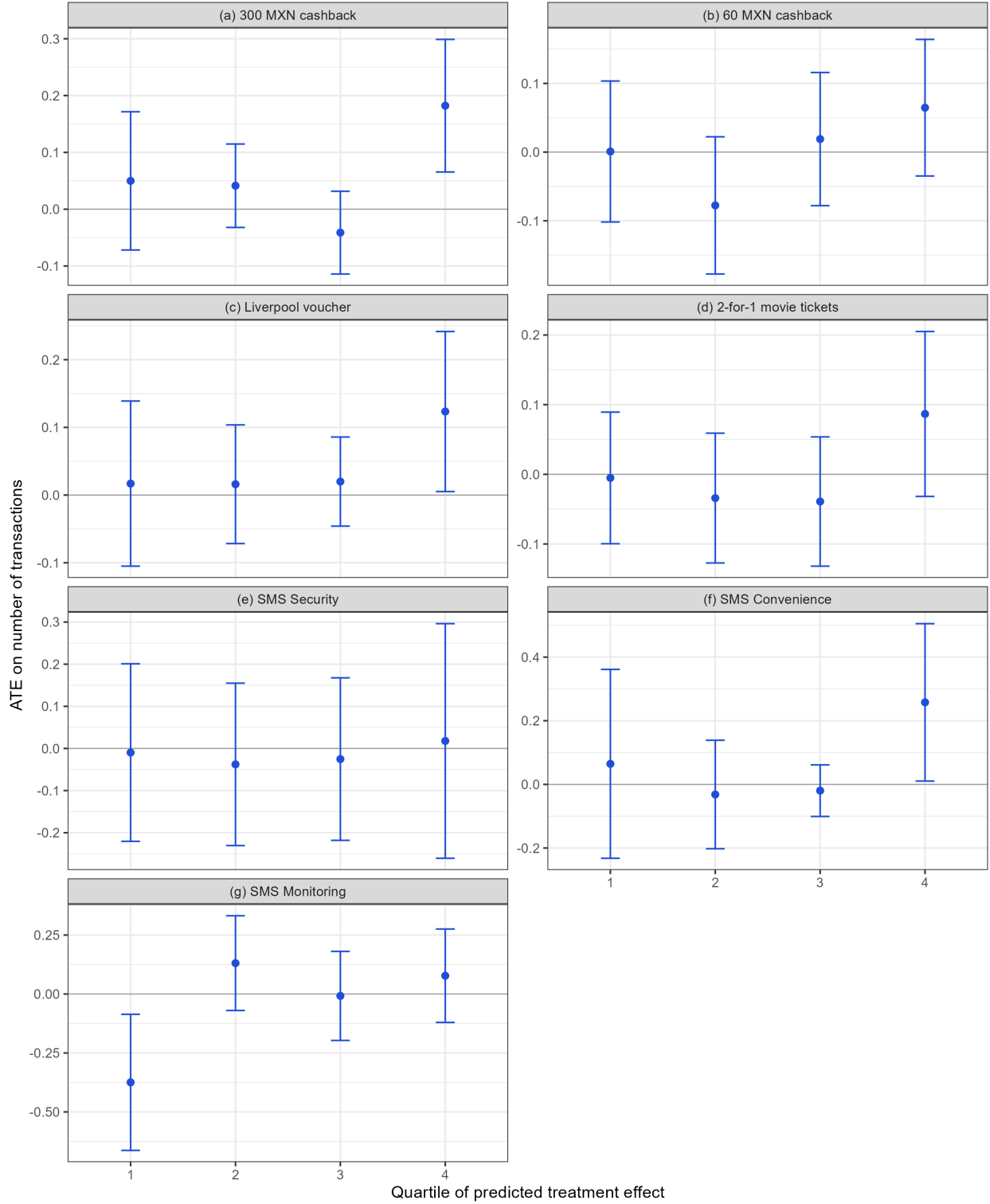


Figure 5 shows the average treatment effect on the number of transactions by quartile of predicted treatment effect, computed on the held-out evaluation sample. The x-axis divides the sample into quartiles of the predicted treatment effect, where Q1 (Q4) is the quartile with the lowest (highest) predicted effect. The y-axis plots the difference-in-averages estimate of the average treatment effect within each quartile, along with its 95% confidence interval.

Appendix A.3 reports average pre-treatment characteristics by quartile of predicted treatment effect for each treatment. Customers in the top quartile for the 300 MXN cashback are the oldest, have the longest tenure with the bank, and have higher income, debit card use, and credit card limits than customers in the middle quartiles, although these measures are highest in the bottom quartile. For the SMS monitoring message, customers in the bottom quartile, whose response is negative, have the highest income and credit card limits and the lowest arrears of any quartile. These comparisons are descriptive. The heterogeneous treatment effects are driven by high-dimensional combinations of covariates rather than by any single characteristic in isolation, and in practice a bank would use the predicted treatment effect scores directly for targeting rather than relying on individual covariates, in a manner similar to how credit scores are used: customers with a predicted response above a threshold determined by the cost of the treatment would receive it.

6.3 Spillovers from Increase in Debit Card Use

An increase in debit card spending can come from new spending or a shift of existing spending away from cash and electronic transfers. In our data, for each customers we observe ATM withdrawals and electronic transfers to third parties in addition to debit card use. We define total spending as the sum of these three amounts. Table 4 reports the effect of each treatment on the three outcomes in the full sample. None of the treatments has a statistically significant effect on total spending (Panel A).

Table 5 repeats the exercise for customers in the top quartile of predicted treatment effects, where the increases in debit card use are concentrated. As in the previous section, the sample is the held-out evaluation sample, and the base category is the minimum-control group within that sample. Among customers ranked by the predicted effect on the amount of debit card transactions, the 300 MXN cashback raises debit card spending by 91 MXN and the Liverpool voucher by 82 MXN. The effects of these two treatments on ATM withdrawals, electronic transfers, and total spending are not statistically significant (Panel A). Among customers ranked by the predicted effect on the number of transactions, the 300 MXN cashback raises the number of debit card transactions by 0.17, the amount by 82 MXN, ATM withdrawals by 265 MXN, and total spending by 385 MXN (Panel B). The standard errors on total spending are between 130 and 410 MXN, several times the size of the effects on debit card spending, so for most treatments we cannot tell whether the additional debit card spending replaces cash and electronic transfers or adds to overall spending.

Table 4: Average Treatment Effects on Spending Outcomes

	ATM Withdrawals	Electronic Transfers	Total Spending
	(1)	(2)	(3)
Panel A: With and without partner-bank credit card			
300MXN cashback	12.410 (25.417)	5.078 (23.952)	50.597 (37.244)
60MXN cashback	-33.437 (25.315)	-18.911 (23.745)	-46.142 (36.991)
Liverpool voucher	-18.290 (25.382)	16.694 (24.005)	9.739 (37.041)
2-for-1 movie ticket	-66.882*** (25.276)	27.602 (24.102)	-16.101 (37.170)
SMS Security	-90.309 (64.597)	57.709 (64.682)	-3.210 (98.805)
SMS Convenience	-69.731 (64.752)	55.840 (64.400)	-40.897 (97.876)
SMS Monitoring	-16.695 (64.995)	28.311 (63.748)	9.076 (96.727)
Observations	1,475,877	1,475,877	1,475,877
Panel B: Without partner-bank credit card only			
300MXN cashback	13.377 (27.976)	-15.872 (23.803)	26.995 (38.936)
60MXN cashback	-26.165 (27.961)	-4.685 (23.877)	-28.207 (38.884)
Liverpool voucher	-31.631 (27.991)	13.603 (24.068)	-3.566 (38.971)
2-for-1 movie ticket	-76.476*** (27.842)	24.694 (24.125)	-21.629 (39.034)
SMS Security	-97.318 (71.056)	45.234 (64.271)	-25.793 (102.092)
SMS Convenience	-110.111 (71.162)	120.818* (65.190)	-26.727 (102.948)
SMS Monitoring	-16.632 (71.851)	-12.247 (63.001)	-27.905 (100.813)
Observations	1,172,923	1,172,923	1,172,923

This table shows the average treatment effects of the core interventions on ATM withdrawals, electronic transfers, and total spending. Total spending is calculated as the sum of amounts from ATM withdrawals, electronic transfers to third parties, and debit card transactions. All columns are estimated using OLS (controlling for strata fixed effect) and the control group is the base category. The outcome variables are in MXN and winsorized at 1th and 99th percentiles. Panel A includes all customers, while Panel B restricts the sample to customers without partner-bank credit card only. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Treatment Effects on Spending Outcomes for Customers in the Top Quartile of Predicted Treatment Effects on Debit Card Use

	Number of Debit Transactions (1)	Amount of Debit Transactions (2)	Amount of ATM Withdrawals (3)	Amount of Electronic Transfers (4)	Total Spending (5)
Panel A: Top quartile by predicted treatment effect on debit card use (amount, MXN)					
300 MXN cashback	0.061 (0.048)	91.352*** (30.165)	23.985 (131.742)	-55.574 (58.629)	47.785 (156.547)
60 MXN cashback	0.058 (0.046)	37.543 (29.616)	156.209 (142.215)	-67.123 (58.598)	111.757 (167.628)
Liverpool voucher	0.087* (0.045)	82.095** (32.308)	-50.301 (149.976)	8.030 (75.400)	44.560 (181.398)
2-for-1 movie tickets	-0.001 (0.042)	21.926 (29.863)	39.643 (137.514)	-72.910 (61.032)	-17.356 (163.829)
SMS Security	0.061 (0.112)	80.937 (76.016)	100.686 (338.812)	14.149 (182.837)	187.043 (407.892)
SMS Convenience	0.151 (0.106)	122.762* (68.502)	-309.798 (278.824)	-190.857 (129.923)	-497.372 (337.057)
SMS Monitoring	0.123 (0.112)	59.024 (69.805)	159.208 (299.106)	-193.425 (133.158)	11.182 (368.790)
Panel B: Top quartile by predicted treatment effect on debit card use (number of transactions)					
300 MXN cashback	0.173*** (0.057)	81.525** (34.488)	265.185* (135.600)	52.689 (64.384)	384.901** (162.861)
60 MXN cashback	0.086* (0.047)	76.167*** (28.249)	85.508 (123.563)	-12.621 (53.348)	159.731 (144.975)
Liverpool voucher	0.146** (0.057)	4.448 (29.982)	128.465 (128.887)	-10.996 (58.843)	112.180 (154.233)
2-for-1 movie tickets	0.083 (0.057)	15.363 (26.218)	-117.751 (110.144)	1.816 (47.939)	-112.280 (128.416)
SMS Security	0.058 (0.139)	-82.485 (66.669)	335.629 (285.124)	-64.184 (149.901)	152.528 (344.515)
SMS Convenience	0.291** (0.125)	113.730 (75.877)	39.928 (317.166)	-65.015 (152.044)	256.534 (385.088)
SMS Monitoring	0.106 (0.102)	-10.374 (46.681)	-207.679 (246.562)	54.386 (104.527)	-157.613 (298.553)

This table reports treatment effects of the core interventions on five outcomes for customers in the top quartile of predicted treatment effects on debit-card use, estimated on the held-out evaluation sample. Column (1) is the number of debit-card transactions; columns (2)–(5) are in Mexican pesos (MXN): the amount of debit-card transactions, ATM withdrawals, electronic transfers to third parties, and total spending (the sum of the previous three); these monetary outcomes are winsorized at the 1st and 99th percentiles. In Panel A the top quartile is defined by the predicted effect on the *amount* of debit-card transactions; in Panel B, by the predicted effect on the *number* of transactions. Each column is estimated by OLS with strata fixed effects; the treatment group is the arm indicated by the row and the base category is the minimum-control group. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

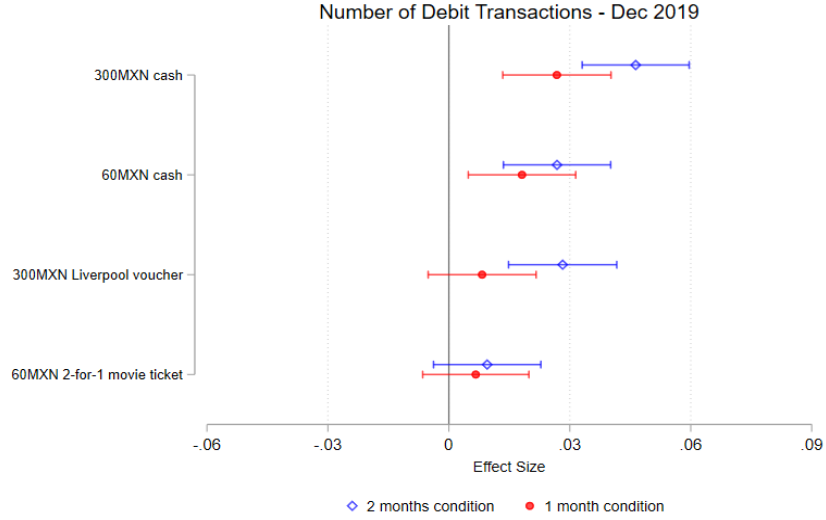
6.4 Exploring Treatment Effect Persistence

Figure 6 plots the average effect of each incentive condition *in December* for customers in the one month condition (i.e., who were only treated in November) and the two month condition (i.e., who treated in November and December). Customers in the one month cash incentive conditions continued to have more debit card transactions than control participants

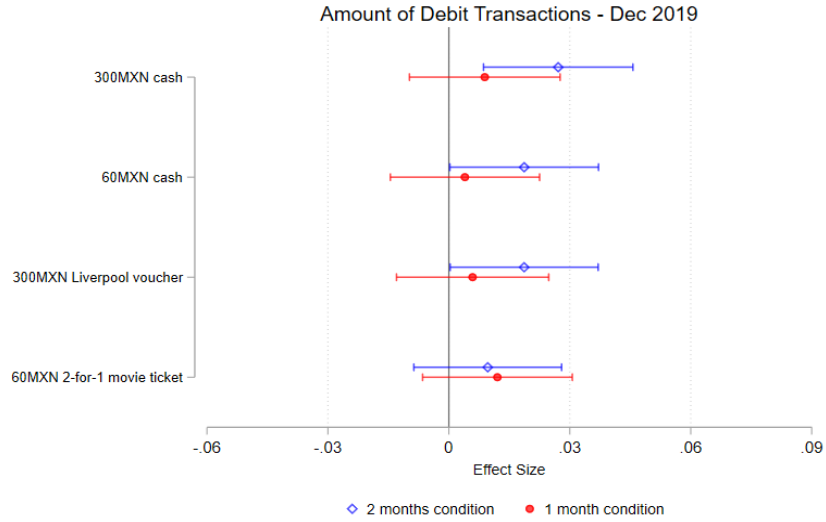
in December, suggesting that the effect of our cash incentives on transactions persisted from one month to the next. However, the same is not true for total amount of debit card spending: customers in the one month cash incentive conditions did not spend significantly more money with their debit card in December than control participants. Additionally, there is no significant evidence of persistence in the voucher conditions. These results suggest that treatment effect persistence can depend on the nature of the treatment (e.g., cash vs voucher) and the focal behavior (e.g., whether to transact vs how much to spend).

It is also informative to examine the December results for customers who were treated twice, because it could be the case that additional treatment produces larger, smaller, or equivalent effects in month two than in month one. Among customers in the two month condition, number of debit card transactions is somewhat larger in December than November for 300 MXN cash, 60 MXN cash, and 300 MXN Liverpool voucher. Thus, additional treatment created some marginal behavior change. However, total amount of debit card spending is very similar in December and November. These findings suggest that the marginal benefit of continued treatment also depends on the focal behavior.

Figure 6: December 2019 - 1 month vs 2 months Duration



(a) Number of debit card transactions



(b) Amount of debit card transactions

6.5 Cost-Effectiveness of the Interventions

The treatments we study differ substantially in their costs to the bank. Thus, an intervention that generates larger increases in debit card use, if expensive, may not be the most cost-effective, while an inexpensive treatment may be attractive even if its effect on debit card use is modest. Moreover, the substantial treatment-effect heterogeneity documented above creates an additional opportunity for the bank to improve cost-effectiveness by directing

treatments toward the customers who are most responsive to them. We therefore ask which treatments generate the largest increase in debit card use per MXN spent, and how much their cost-effectiveness can be improved through targeting.

A complete calculation of the profitability of increased debit card use is beyond the scope of our analysis. As discussed in Section 2, greater debit card use can benefit the bank through several channels, including interchange revenue, lower cash-management costs, higher account balances, and cross-selling opportunities. We instead make the simplifying assumption that an additional MXN of debit card use has the same value to the bank regardless of which treatment generates it. Under this assumption, differences in cost-effectiveness across treatments depend only on how much additional debit card use they generate and how much they cost.

According to the partner bank, cash incentives are the most expensive because the bank incurs taxes on rewards: each 1 MXN paid to a customer costs the bank 1.25 MXN. In-kind incentives, by contrast, can be purchased in bulk at a discount. The bank pays approximately 240 MXN for a Liverpool voucher with a face value of 300 MXN and approximately 10 MXN for a 2-for-1 movie ticket with a nominal value of 60 MXN. Because incentives are paid only to customers who complete their debit card challenge, these costs must be scaled by the rate at which customers complete. The bank reports an aggregate completion rate of 2.74%, but we do not observe completion at the customer level. We therefore consider alternative assumptions about how completion vary across customers.

We ask how a bank with a fixed marketing budget should spend it to maximize debit card use. We take as our benchmark the budget of the experiment itself. The bank treated 599,633 customers, against 926,544 assigned to the control group, at an average cost of 9.07 MXN per treated customer, for a total of 5.44 million MXN.

We consider three policies that differ in what the bank knows when it allocates its budget. In the first, the bank knows only the average effect of a single treatment, as it might from A/B testing conducted before scaling a campaign, but does not know which customers are more responsive to it. The bank offers the treatment to as many customers as the budget allows, selecting customers at random if the budget does not cover the entire user base, and offers nothing if the treatment is not expected to increase debit card use on average. We report this policy separately for each of the seven treatments. In the second, the bank also knows how the effect of the treatment its considering varies across customers, and spends the budget on those with the highest predicted responses to it. In the third, it knows the predicted response of every customer to every treatment, and chooses both whom to treat and which treatment each of them receives. The first policy is our untargeted benchmark. Comparing it with the second measures what the bank gains from knowing whom to treat, and comparing the second with the third measures what it gains from knowing what to send. Because the second and third policies use predicted treatment effects, we evaluate all three

on a sample that was not used in any aspect of model training. The predicted responses come from the seven causal forests of Section 6.2, one per treatment. We estimate them separately rather than as a single multi-arm forest to prioritize flexibility in training, since a joint forest would grow partitions common to all arms, thereby fixing a single set features and tuning parameters across them.

We report results under two assumptions about costs. The first holds the cost of each treatment constant across customers. The second allows the cost of the incentive treatments to increase with the predicted treatment effect, reflecting that customers whose debit card use is predicted to increase by more are also more likely to reach the spending target and collect the reward. The three SMS treatments carry no reward, so their cost remains 4.55 MXN for every customer regardless of the predicted response. We model the probability of completing the challenge as a logistic function of the predicted treatment effect, with an intercept calibrated to match the aggregate completion rate of 2.74%. We calibrate the slope so that customers in the top decile of predicted response complete at five times the aggregate rate, while customers in the bottom decile complete with negligible probability. This introduces variation in expected costs both across incentive treatments, which differ in their average effects, and across customers receiving the same incentive, who differ in their predicted responses. Appendix D gives the parametrization and reports results for a wider range of gradients.

Because the treatments differ in cost, a budget of 5.44 million MXN would reach a different number of customers under each. The three SMS messages cost 4.55 MXN per customer, so the budget covers 1,195,097 of them, 78.3% of the client base. The 300 MXN cash incentive costs 14.83 MXN and covers 366,668, or 24.0%.

A treatment is worth offering to a randomly selected group of customers only if it raises debit card use, and for most of the treatments we cannot establish that it does. Table 6 reports each treatment’s effect, its cost-effectiveness, and the additional debit card use generated when it is offered to as many customers as the budget allows. On the amount of debit card transactions, only the 300 MXN cash incentive has an effect significantly greater than zero, and offering it to 24.0% of customers generates 12.0 million MXN of additional debit card spending (Panel A). The Convenience SMS message has a similar estimated effect, 31.79 MXN per customer against 32.75, and because it costs less than a third as much its cost-effectiveness is three times higher, 6.99 MXN of additional use per MXN spent against 2.21. Its effect is too imprecisely estimated to act on, however ($t = 0.86$), as are those of four other treatments, and the Monitoring SMS message is estimated to reduce debit card use ($t = -2.58$). Two treatments clear the bar on the number of transactions (Panel B): the Convenience SMS message generates 103,689 additional transactions and the 300 MXN cash incentive 23,887, so on that outcome the cheaper message both reaches more customers and generates more transactions. The right-hand block of Table 6 prices each incentive at the rate its own customers complete rather than at the average rate across incentives. The

300 MXN cash incentive is the one incentive that raises debit card use, so it is also the one whose reward is claimed: its expected cost rises from 14.83 to 27.72 MXN per customer, the budget reaches 12.9 per cent of the client base rather than 24.0, and the additional spending it generates falls from 12.0 to 6.4 million MXN. The three SMS messages carry no reward and are unchanged.

Cost-effectiveness improves substantially when the bank chooses which customers to treat (Table 7). Spending the same 5.44 million MXN on the 300 MXN cash incentive, directed to the customers with the highest predicted responses rather than to a random group, generates 69.1 million MXN of additional debit card spending, compared with 12.0 million. The two policies reach the same 24.0% of customers and spend the same amount, and differ only in which customers they reach. The targeted estimate is also more precise ($t = 5.56$ compared with $t = 2.15$). The 60 MXN cash incentive, whose average effect is indistinguishable from zero, generates 37.7 million MXN when targeted ($t = 2.54$).

Allowing the bank to choose the treatment as well as the customer generates little further improvement. At the same budget the personalized policy generates 72.3 million MXN, compared with 69.1 million from targeting the 300 MXN cash incentive alone. On the number of transactions it does slightly worse than the best single treatment, generating 157,605 additional transactions against 166,872 from targeting the Convenience SMS message. The largest gain from targeting therefore comes from selecting customers rather than from personalizing which treatments they receive.

Allowing cost to rise with the predicted effect reduces the estimated gains from targeting on the amount of transactions. Targeting the 300 MXN cash incentive generates 24.3 million MXN rather than 69.1 million, and the personalized policy 11.3 million rather than 72.3 million. Targeting selects the customers with the largest predicted responses, but these are also the customers most likely to complete the challenge and claim the reward. The results for the number of transactions change very little. Targeting the Convenience SMS message generates 166,872 additional transactions under both cost models, and the personalized policy generates 157,605 and 162,092. This message carries no reward, so its cost does not vary with the response it produces.

Appendix Table E.1 reports every policy with different budgets, under each cost gradient. Reducing the budget by half reduces the additional debit card use generated by considerably less than half. With 2.72 million MXN the personalized policy reaches 24.5% of customers and generates 55.4 million MXN, 77% of what twice the budget generates.

Table 6: Treatment effects, costs, and coverage under a fixed budget

Treatment	Treatment effect	Cost does not vary				Cost rises with predicted effect			
		Cost per customer	Cost-eff.	Treated (%)	Total effect	Cost per customer	Cost-eff.	Treated (%)	Total effect
<i>Panel A. Amount of debit card transactions (MXN)</i>									
300 MXN cash incentive	32.75** (15.24)	14.83	2.21	24.0	12,007,148	27.72	1.18	12.9	6,424,335
Convenience SMS	31.79 (36.94)	—	—	—	—	—	—	—	—
Security SMS	0.97 (35.40)	—	—	—	—	—	—	—	—
2-for-1 movie ticket	-0.94 (14.68)	—	—	—	—	—	—	—	—
Liverpool voucher	-2.50 (14.49)	—	—	—	—	—	—	—	—
60 MXN cash incentive	-7.58 (14.39)	—	—	—	—	—	—	—	—
Monitoring SMS	-81.77*** (31.66)	—	—	—	—	—	—	—	—
<i>Panel B. Number of debit card transactions</i>									
Convenience SMS	0.087* (0.047)	4.55	0.0191	78.3	103,689	4.55	0.0191	78.3	103,689
300 MXN cash incentive	0.065*** (0.019)	14.83	0.0044	24.0	23,887	17.04	0.0038	20.9	20,794
Liverpool voucher	0.027 (0.019)	—	—	—	—	—	—	—	—
60 MXN cash incentive	0.016 (0.018)	—	—	—	—	—	—	—	—
2-for-1 movie ticket	-0.002 (0.018)	—	—	—	—	—	—	—	—
Monitoring SMS	-0.003 (0.044)	—	—	—	—	—	—	—	—
Security SMS	-0.027 (0.044)	—	—	—	—	—	—	—	—

Notes: This table shows treatment effects, costs, cost effectiveness, coverage and aggregate effects of each treatment under a uniform policy. A treatment is offered to a randomly chosen group of customers, as large as a budget of 5.4 million MXN allows. Treatment effects are OLS estimates with randomization-stratum fixed effects and the baseline value of the outcome as a control; heteroskedasticity-robust standard errors in parentheses. For consistency of comparison with targeted policies that use machine learning predictions, all results are calculated on a random sample of customers set aside for evaluation of the causal forests. A treatment is offered to customers only where its effect is significantly greater than zero. — marks treatments that do not have a significantly positive treatment effect and are thus not offered to any client. Cost-effectiveness is the treatment effect divided by the cost per customer, and total effect is the treatment effect multiplied by the fraction of customers covered by the policy, scaled to the 1,526,177 clients in the client base. In the left-hand block the probability of completing a challenge is constant. In the right-hand block it rises with predicted effects, following the parametrization described in Appendix D. *** p<0.01, ** p<0.05, * p<0.10.

Table 7: Coverage, aggregate effects and cost effectiveness of different targeting policies

Treatment	Cost does not vary			Cost rises with predicted effect		
	Treated (%)	Total effect	Cost-effectiveness	Treated (%)	Total effect	Cost-effectiveness
<i>Panel A. Amount of debit card transactions (MXN)</i>						
300 MXN cash incentive	24.0	69,050,563 (5.56)***	12.70	15.4	24,271,923 (3.13)***	4.46
60 MXN cash incentive	51.7	37,741,609 (2.54)**	7.23	36.5	42,997,151 (3.22)***	7.91
Monitoring SMS	15.7	15,637,718 (0.60)	14.35	15.7	15,637,718 (0.60)	14.35
Convenience SMS	—	—	—	—	—	—
Security SMS	—	—	—	—	—	—
Liverpool voucher	—	—	—	—	—	—
2-for-1 movie ticket	—	—	—	—	—	—
Personalized assignment	39.3	72,261,106 (2.64)***	13.29	61.5	11,272,002 (0.28)	2.07
<i>Panel B. Number of debit card transactions</i>						
Convenience SMS	77.5	166,872 (3.21)***	0.0310	77.5	166,872 (3.21)***	0.0310
Liverpool voucher	32.0	65,070 (3.93)***	0.0120	25.5	31,131 (2.27)**	0.0057
300 MXN cash incentive	24.0	59,858 (3.63)***	0.0110	19.1	36,750 (2.94)***	0.0068
60 MXN cash incentive	53.9	45,874 (2.30)**	0.0084	35.0	25,564 (1.55)	0.0047
2-for-1 movie ticket	35.0	17,136 (0.95)	0.0067	35.0	17,136 (0.95)	0.0068
Monitoring SMS	—	—	—	—	—	—
Security SMS	—	—	—	—	—	—
Personalized assignment	74.1	157,605 (3.19)***	0.0290	75.6	162,092 (3.20)***	0.0298

Notes: This table shows percentage of user base treatment, total effects and cost effectiveness under two different targeting policies. Under the first targeting policy a treatment is offered to the customers with the highest predicted responses to it, and treats as many customers as the budget allows. The results for each treatment under this policy are shown on the row with the corresponding treatment name. The second targeting policy allows the partner-bank to choose which treatment to offer to each customer, based on predicted treatment effects. The results of this policy are shown on the row label personalized assignment. Predicted responses come from the causal forests of Section 6.2. All results are calculated on a random sample of customers set aside for evaluation of the causal forests. Total effect is scaled to the 1,526,177 clients in the client base; t -statistics in parentheses. In the left-hand block the probability of completing a challenge is constant. In the right-hand block it rises with predicted effects following the parametrization described in Appendix D. The budget is 5.4 million MXN. — marks treatments that do not have a significant RATE and do not have a significantly positive average treatment effect and are thus not offered to any client. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

7 Conclusion

We conduct a large-scale field experiment with 1.5 million bank customers to investigate how consumer spending can be shifted from cash to debit cards. Overall, the average treatment effects of our interventions are modest. However, we find that these modest average effects mask significant heterogeneity in treatment effects.

Improved understanding of what interventions alter consumer behavior can advance our understanding of consumer psychology, and solve problems of practical importance to firms and consumers. Our setting allows for a comprehensive set of comparisons between levels of cash and in-kind incentives, as well as with informational SMS campaigns. Our findings provide practitioners and policymakers with a comprehensive comparative benchmark for alternative strategies to influence consumers. Our findings show that inexpensive informational campaigns do not substitute for financial incentives on average, but that targeting can make both considerably more cost-effective. We also highlight the importance of targeting interventions for greater effectiveness.

References

- Agarwal, S., Ghosh, P., Li, J., and Ruan, T. (2024). Digital payments and consumption: Evidence from the 2016 demonetization in india. *The Review of Financial Studies*, page hhae005.
- Anderson, M. 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):1481–1495.
- Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *Annals of Statistics*, 47(2):1148–1178.
- Athey, S. and Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational studies*, 5(2):37–51.
- Bachas, P., Gertler, P., Higgins, S., and Seira, E. (2021). How debit cards enable the poor to save more. *The Journal of finance*, 76(4):1913–1957.
- Bagozzi, R. P. and Dholakia, U. (1999). Goal setting and goal striving in consumer behavior. *Journal of marketing*, 63(4_suppl1):19–32.
- Borzekowski, R., Kiser, E. K., and Ahmed, S. (2008). Consumers’ use of debit cards: Patterns, preferences, and price response. *Journal of Money, Credit and Banking*, 40(1):149–172.
- Brown, M., Hentschel, N., Mettler, H., and Stix, H. (2022). The convenience of electronic payments and consumer cash demand. *Journal of Monetary Economics*, 130:86–102.
- Bruno, P., Jeenah, U., Gandhi, A., and Gancho, I. (2024). Global payments in 2024: Simpler interfaces, complex reality. *McKinsey & Company*.
- Chaudhuri, R., Gathinji, C., Tayar, G., and Williams, E. (2022). Sustaining digital payments growth: Winning models in emerging markets. *mckinsey*. <https://www.mckinsey.com/industries/financial-services/our-insights/sustaining-digital-payments-growth-winning-models-in-emerging-markets>.
- Chen, Y., Lee, J.-Y., Sridhar, S., Mittal, V., McCallister, K., and Singal, A. G. (2020). Improving cancer outreach effectiveness through targeting and economic assessments: Insights from a randomized field experiment. *Journal of Marketing*, 84(3):1–27.
- Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I. (2018). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. Technical report, National Bureau of Economic Research.
- Clarke, D. (2021). RWOLF2: Stata module to calculate Romano-Wolf stepdown p-values for multiple hypothesis testing. Statistical Software Components, Boston College Department of Economics.

- Crump, R. K., Hotz, V. J., Imbens, G. W., and Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1):187–199.
- Duflo, E., Glennerster, R., and Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4:3895–3962.
- Greenberg, A. E., Hershfield, H. E., Shu, S. B., and Spiller, S. A. (2023). What motivates social security claiming age intentions? testing behaviorally informed interventions alongside individual differences. *Journal of Marketing Research*, 60(6):1052–1070.
- Heath, C. and Soll, J. B. (1996). Mental budgeting and consumer decisions. *Journal of consumer research*, 23(1):40–52.
- Higgins, S. (2024). Financial technology adoption: Network externalities of cashless payments in mexico. *American Economic Review*, 114(11):3469–3512.
- Hirschman, E. C. (1979). Differences in consumer purchase behavior by credit card payment system. *Journal of consumer Research*, 6(1):58–66.
- Hirshleifer, S., McKenzie, D., Almeida, R., and Ridao-Cano, C. (2016). The impact of vocational training for the unemployed: Experimental evidence from turkey. *The Economic Journal*, 126(597):2115–2146.
- Kahneman, D. and Tversky, A. (1979). 1prospect theory: An analysis of decision under risk1. *Econometrica*, 47.
- Kosse, A. (2010). The safety of cash and debit cards: a study on the perception and behaviour of dutch consumers.
- Lukas, M. F. and Howard, R. C. (2023). The influence of budgets on consumer spending. *Journal of Consumer Research*, 49(5):697–720.
- McConnell, B. and Vera-Hernández, M. (2015). Going beyond simple sample size calculations: a practitioner’s guide.
- McKenzie, D. (2017). Should we require balance t-tests of baseline observables in randomized experiments? *World Bank blog*, <https://blogs.worldbank.org/en/impac evaluations/should-we-require-balance-t-tests-baseline-observables-randomized-experiments>.
- McKenzie, D. (2022). More on using a waitlist as part of your experimental design. World Bank Development Impact. Published July 5, 2022; accessed August 31, 2026.
- McKenzie, D. (2025). Designing and analysing powerful experiments: Practical tips for applied researchers. *Fiscal Studies*, 46(3):305–322.
- Mrkva, K., Posner, N. A., Reeck, C., and Johnson, E. J. (2021). Do nudges reduce disparities? choice architecture compensates for low consumer knowledge. *Journal of Marketing*, 85(4):67–84.

- Perez, M. (2020). Mexico seeks to reduce consumers’ longstanding reliance on cash. *Federal Reserve Bank of Dallas*.
- Prelec, D. and Loewenstein, G. (1998). The red and the black: Mental accounting of savings and debt. *Marketing science*, 17(1):4–28.
- Reinholtz, N., Bartels, D. M., and Parker, J. R. (2015). On the mental accounting of restricted-use funds: How gift cards change what people purchase. *Journal of Consumer Research*, 42(4):596–614.
- Schuh, S. and Stavins, J. (2010). Why are (some) consumers (finally) writing fewer checks? the role of payment characteristics. *Journal of Banking & Finance*, 34(8):1745–1758.
- Shaffer, V. A. and Arkes, H. R. (2009). Preference reversals in evaluations of cash versus non-cash incentives. *Journal of Economic Psychology*, 30(6):859–872.
- Shah, A. M., Osborne, M., Lefkowitz Kalter, J., Fertig, A., Fishbane, A., and Soman, D. (2023). Identifying heterogeneity using recursive partitioning: evidence from sms nudges encouraging voluntary retirement savings in mexico. *PNAS nexus*, 2(5):pgad058.
- Sharif, M. A. and Woolley, K. (2020). The effect of categorization on goal progress perceptions and motivation. *Journal of Consumer Research*, 47(4):608–630.
- Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.
- Soman, D. (2015). *The last mile: Creating social and economic value from behavioral insights*. University of Toronto Press.
- Statista (2023). Digital payments: Market insights report. *Market Insights by Statista*.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, 4(3):199–214.
- Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making*, 12(3):183–206.
- Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, New Haven, CT.
- Visa (2022). Optimizing your debit card business. Technical report, Visa Caribbean. Accessed: September 26, 2024.
- Wood, W. and Neal, D. T. (2009). The habitual consumer. *Journal of Consumer Psychology*, 19(4):579–592.
- World Bank (2012). Developing a comprehensive national retail payments strategy. Financial infrastructure series: Payment systems policy and research, World Bank, Washington, DC.
- World Bank (2022). The global index database 2021: Use of accounts. World Bank. Accessed August 19, 2026.

- Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., and Wager, S. (2021). Evaluating treatment prioritization rules via rank-weighted average treatment effects.” arxiv. *arXiv preprint arXiv:2111.07966*.
- Zandi, M., Singh, V., and Irving, J. (2013). The impact of electronic payments on economic growth. *Moody’s Analytics: Economic and Consumer Credit Analytics*, 217(2):1–16.

Appendix

A Causal Forests

A.1 Detailed Description

We use causal forests to study treatment effect heterogeneity (Athey and Wager, 2019). Causal forests combine experimental variation with machine learning to predict (causal) treatment effects at the individual level. Specifically, using a potentially large set of individual level pre-treatment variables, the algorithm recursively partitions the covariate space to identify sub-populations with larger (and smaller) treatment effects.

One potential risk of recursively splitting a sample and comparing treatment effects across subpopulations is overfitting. Overfitting would occur if our individual-level predictions for treatment effects were influenced by noise, which we would spuriously attribute to the treatment. To avoid overfitting, the algorithm follows a multiple split-sample procedure, by which half the training sample is used to identify candidate partitions of the covariate space, and the other half is used to calculate treatment effects. This way, the outcomes of any given individual are never used in the process of calculating its individual level prediction. The algorithm produces a model to predict treatment effects at the individual level; that is, a set of rules that assign to each observation an individual level predicted treatment effect, based on its covariates. The performance of this model is then evaluated in a validation sample that was not used in training.²

We evaluate the performance of the model calculating the Area Under the Targeting Operator Characteristic curve (AUROC) (Yadlowsky et al., 2021). The Targeting Operator Characteristic curve or TOC curve shows the benefits (in terms of treatment effects) of sorting individuals based on the treatment effects predicted by the causal forests. Specifically, individuals are sorted and split into groups where each group is the top q -th fraction of individuals with the largest predicted treatment effect. The TOC shows, for q running between 0 and 1, the difference in treatment effect of that group and the overall average treatment effect, considering all observations in the sample. A TOC that runs over positive values and reaches zero for $q=1$ will be indicative of a causal forest that successfully identifies treatment effect heterogeneity, since sorting based on its predictions would lead to actual treatment effects following the same ordering; that is, high (low) predicted treatment effects would follow high (low) actual treatment effects.

As with all machine learning methods, the process of variable selection and hyperparameter tuning is important to further prevent overfitting, and ensure appropriate performance out of sample. We tune all tunnable parameters with the default method for parameter tuning embedded in the package `grf`, and use a wrapper method for variable se-

²For an efficient use of the sample, we implement cross-validation over 10 folds.

lection. For variable selection, our starting point is the 120 pre-treatment variables provided by the partner-bank. In our setting, model performance measured by AUROC is significantly better with a small number of covariates than with the full set of 120 variables. We thus define the wrapping procedure to sequentially identify the combination of up to 5 variables that lead to the best performing model in terms of AUROC. To prevent overfitting and to make sure that testing information is not influencing modelling decisions (beyond the “few covariate” intuition described above), wrapping is implemented with nested cross validation, by which RATE comparisons are performed in one sample that identifies a set of important covariates, and the final model is trained with those covariates in a separate sample.

A.2 Test for Heterogeneous Treatment Effects

Figure A.1: Rank-Weighted Average Treatment Effect (RATE) plots (Outcome: Amount of Transactions)

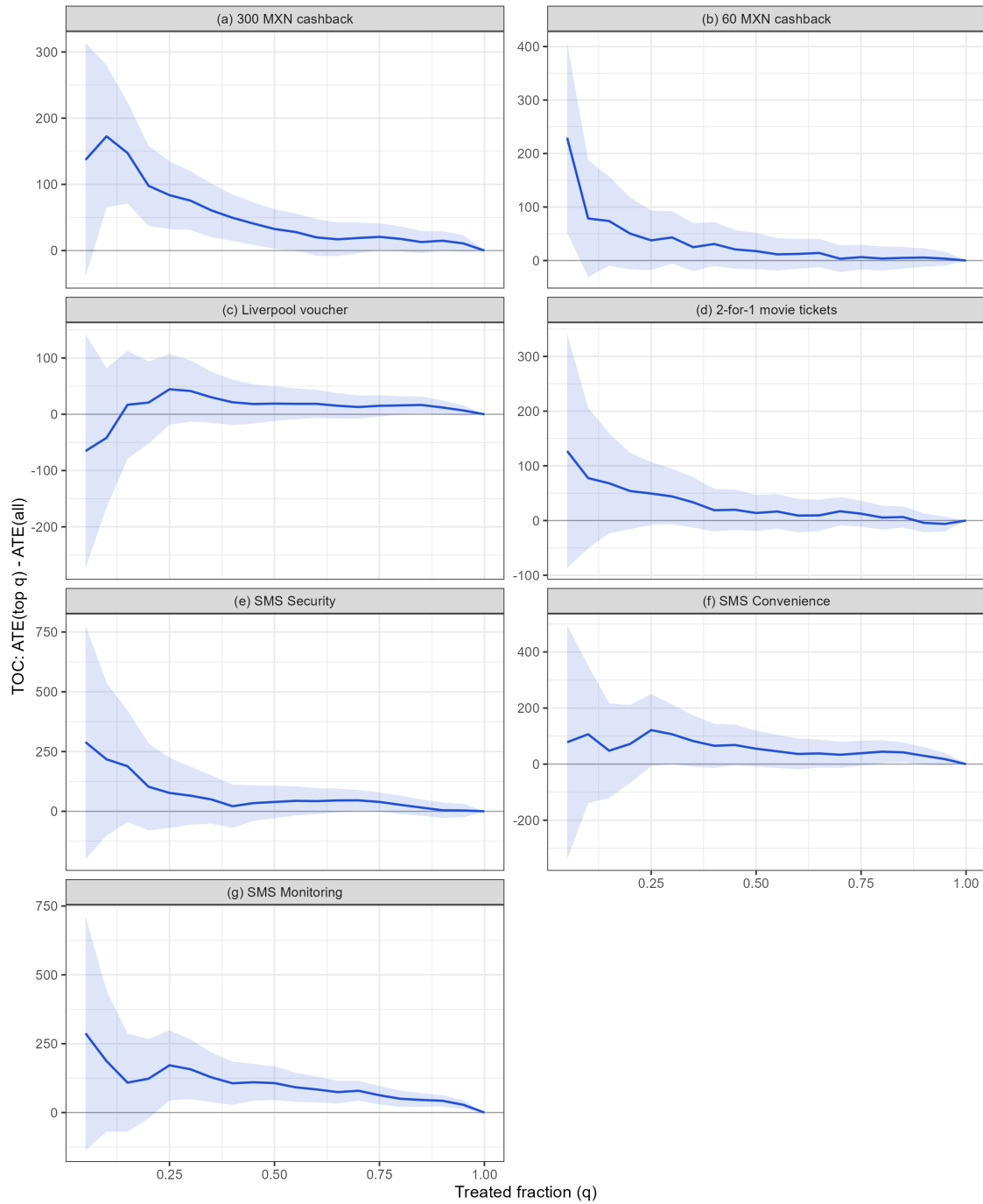
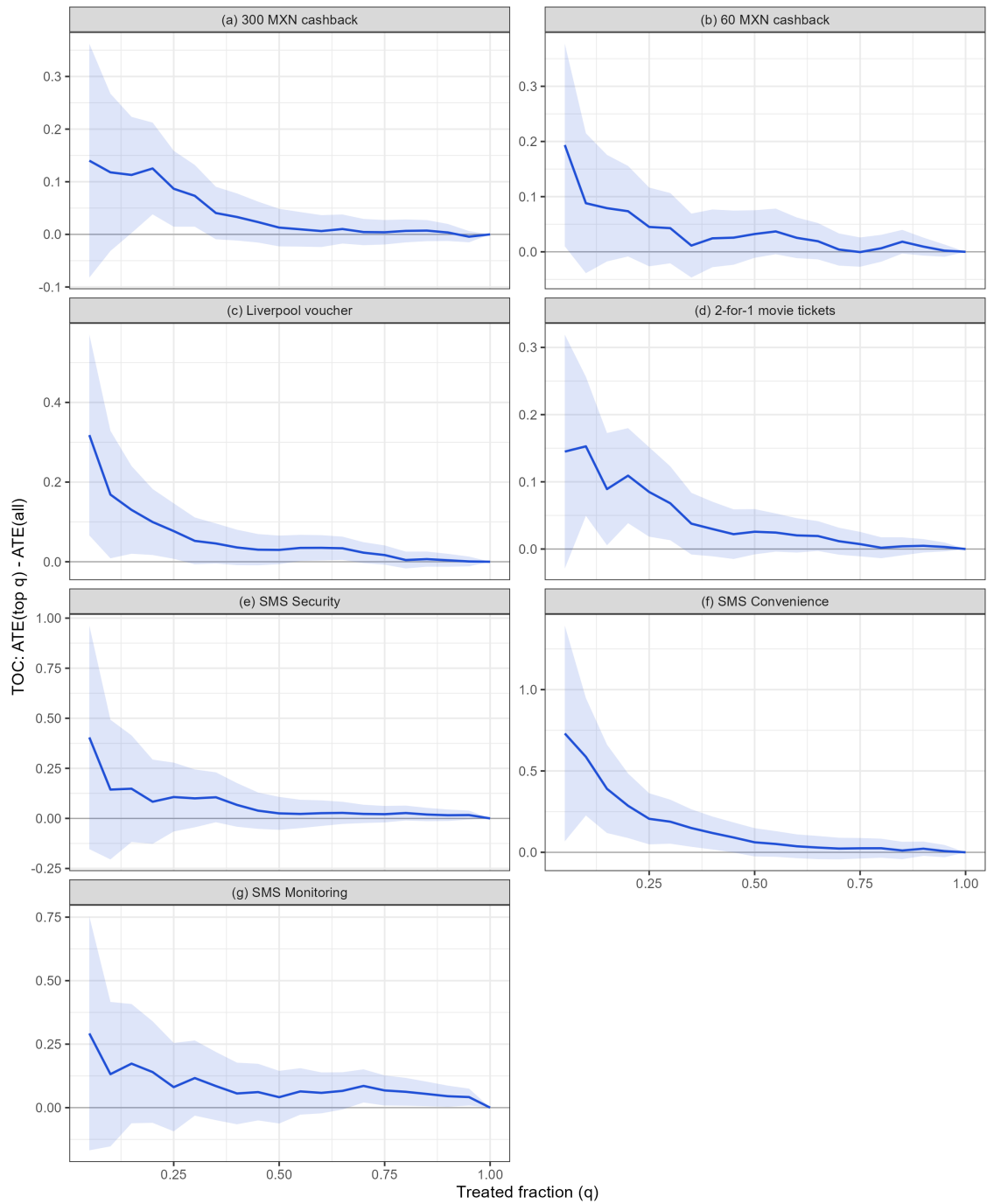


Figure A.2: Rank-Weighted Average Treatment Effect (RATE) plots (Outcome: Number of Transactions)



A.3 Average Covariate Values Within Each Quartile

Table A.1: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect: Liverpool Voucher

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	46.15	43.36	43.10	45.69	<0.001	<0.001
Monthly income (MXN)	13,859.73	10,341.94	10,647.71	17,694.73	<0.001	<0.001
Tenure with bank (months)	87.98	73.43	78.05	93.62	<0.001	<0.001
Total amount of debit card trans (MXN)	1,313.34	928.67	1,066.60	1,623.83	0.020	<0.001
Debit card spending on Necessity MCC (MXN)	224.05	119.89	133.51	228.29	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	107.80	32.69	31.53	94.14	<0.001	<0.001
Partner-bank credit card limit (MXN)	7,862.73	4,033.95	4,179.35	9,642.98	<0.001	<0.001
Monthly credit card closing balance (MXN)	2,626.56	1,472.88	1,430.18	2,548.30	<0.001	<0.001
Credit card interest charged (MXN)	44.98	26.59	26.49	36.42	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	1.41%	2.45%	1.62%	1.27%	0.311	<0.001

Table A.1 reports average pre-treatment characteristics by quartile of the predicted treatment effect on transaction amount, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.2: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect: 2-for-1 Movie Tickets

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	45.50	43.94	44.55	44.62	<0.001	<0.001
Monthly income (MXN)	15,027.56	11,940.26	12,719.00	12,658.00	<0.001	<0.001
Tenure with bank (months)	87.71	78.15	87.48	79.32	<0.001	<0.001
Total amount of debit card trans (MXN)	1,449.06	1,135.40	1,181.96	1,194.95	0.003	0.003
Debit card spending on Necessity MCC (MXN)	200.53	166.80	161.83	171.97	0.037	0.173
Debit card spending on Luxury MCC (MXN)	78.98	57.35	66.12	61.47	<0.001	<0.001
Partner-bank credit card limit (MXN)	7,779.45	5,592.65	6,377.70	5,939.15	<0.001	<0.001
Monthly credit card closing balance (MXN)	2,431.48	1,702.49	1,951.35	1,907.56	<0.001	<0.001
Credit card interest charged (MXN)	39.91	27.27	29.81	33.30	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	1.44%	1.67%	1.64%	1.61%	0.225	0.334

Table A.2 reports average pre-treatment characteristics by quartile of the predicted treatment effect on transaction amount, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.3: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect: SMS Security

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	43.13	43.53	44.55	46.96	<0.001	<0.001
Monthly income (MXN)	13,334.72	10,143.37	11,377.03	17,177.27	<0.001	<0.001
Tenure with bank (months)	78.00	74.17	81.03	98.66	<0.001	<0.001
Total amount of debit card trans (MXN)	1,481.87	815.41	1,008.95	1,607.07	0.323	<0.001
Debit card spending on Necessity MCC (MXN)	235.31	87.36	121.20	250.77	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	91.83	28.40	38.20	98.90	<0.001	<0.001
Partner-bank credit card limit (MXN)	7,728.49	3,503.28	4,153.76	9,883.55	<0.001	<0.001
Monthly credit card closing balance (MXN)	2,600.02	1,188.49	1,182.30	3,016.22	0.006	<0.001
Credit card interest charged (MXN)	42.41	21.17	20.62	49.38	0.009	<0.001
With arrears on other partner-bank credit products (=1)	2.24%	2.50%	0.86%	0.82%	<0.001	<0.001

Table A.3 reports average pre-treatment characteristics by quartile of the predicted treatment effect on transaction amount, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.4: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect: SMS Convenience

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	46.09	43.71	43.96	44.84	<0.001	<0.001
Monthly income (MXN)	16,767.28	10,683.46	10,534.85	14,281.18	<0.001	<0.001
Tenure with bank (months)	96.78	73.87	76.01	85.40	<0.001	<0.001
Total amount of debit card trans (MXN)	1,681.43	1,012.84	975.74	1,243.29	<0.001	<0.001
Debit card spending on Necessity MCC (MXN)	314.72	101.37	92.19	186.85	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	151.77	15.25	8.99	86.95	<0.001	<0.001
Partner-bank credit card limit (MXN)	10,462.87	3,387.73	2,325.90	9,294.24	<0.001	<0.001
Monthly credit card closing balance (MXN)	3,327.56	1,012.17	708.11	3,044.77	<0.001	<0.001
Credit card interest charged (MXN)	52.03	16.91	12.46	52.92	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	1.51%	1.69%	1.78%	1.89%	0.036	0.177

Table A.4 reports average pre-treatment characteristics by quartile of the predicted treatment effect on transaction amount, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.5: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: 300 MXN Cashback

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	38.25	44.85	50.01	45.55	<0.001	<0.001
Monthly income (MXN)	11,414.48	12,600.08	13,842.79	14,331.35	<0.001	<0.001
Tenure with bank (months)	40.83	69.02	117.66	105.75	<0.001	<0.001
Total amount of debit card trans (MXN)	1,434.82	918.89	931.94	1,679.10	<0.001	<0.001
Debit card spending on Necessity MCC (MXN)	318.43	56.89	40.94	286.20	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	122.21	17.13	17.29	103.67	<0.001	<0.001
Partner-bank credit card limit (MXN)	7,178.71	4,407.81	4,404.70	9,678.38	0.600	<0.001
Monthly credit card closing balance (MXN)	2,393.57	1,474.20	1,347.76	2,844.59	0.128	<0.001
Credit card interest charged (MXN)	34.80	27.09	23.64	48.25	0.543	<0.001
With arrears on other Partner-bank credit products (=1)	0.78%	2.61%	2.07%	1.03%	0.019	<0.001

Table A.5 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.6: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: 60 MXN Cashback

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	42.46	45.15	50.22	40.60	<0.001	<0.001
Monthly income (MXN)	14,505.90	12,560.96	12,906.38	12,026.65	<0.001	<0.001
Tenure with partner-bank (months)	82.45	82.29	91.21	76.76	<0.001	<0.001
Total amount of debit card trans (MXN)	1,285.70	1,229.21	1,136.22	1,252.84	0.376	<0.001
Debit card spending on Necessity MCC (MXN)	186.00	175.45	156.80	180.29	0.508	<0.001
Debit card spending on Luxury MCC (MXN)	71.11	61.29	61.38	64.29	0.907	<0.001
Partner-bank credit card limit (MXN)	7,071.66	6,392.96	5,852.72	5,777.41	<0.001	<0.001
Monthly credit card closing balance (MXN)	2,157.27	1,986.72	1,725.95	1,880.01	<0.001	<0.001
Credit card interest charged (MXN)	35.37	33.21	29.20	31.67	0.001	<0.001
With arrears on other partner-bank credit products (=1)	1.87%	1.83%	1.37%	1.49%	0.009	<0.001

Table A.6 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.7: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: Liverpool Voucher

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	40.36	43.60	52.23	42.12	<0.001	<0.001
Monthly income (MXN)	11,985.08	12,946.20	14,254.32	13,358.73	<0.001	<0.001
Tenure with bank (months)	55.18	93.38	107.75	76.77	<0.001	<0.001
Total amount of debit card trans (MXN)	1,501.41	1,161.58	887.14	1,382.30	<0.001	<0.001
Debit card spending on Necessity MCC (MXN)	232.75	150.68	64.91	257.40	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	80.79	61.78	28.73	94.85	0.285	<0.001
Partner-bank credit card limit (MXN)	5,382.73	7,633.09	5,383.50	7,319.77	<0.001	<0.001
Monthly credit card closing balance (MXN)	1,737.80	2,348.51	1,610.34	2,381.29	<0.001	<0.001
Credit card interest charged (MXN)	31.32	40.25	25.98	36.92	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	0.65%	1.91%	2.39%	1.81%	<0.001	<0.001

Table A.7 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.8: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: 2-for-1 Movie Tickets

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	50.80	43.78	43.62	40.40	<0.001	<0.001
Monthly income (MXN)	20,637.94	11,486.30	10,242.92	9,979.99	<0.001	<0.001
Tenure with bank (months)	124.23	73.56	81.18	54.00	<0.001	<0.001
Total amount of debit card trans (MXN)	2,249.18	1,116.76	856.10	738.82	<0.001	<0.001
Debit card spending on Necessity MCC (MXN)	326.27	143.45	106.21	124.81	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	143.08	48.98	32.77	39.25	<0.001	<0.001
Partner-bank credit card limit (MXN)	15,445.87	4,405.19	3,445.08	2,409.68	<0.001	<0.001
Monthly credit card closing balance (MXN)	3,945.72	1,635.86	1,422.11	990.87	<0.001	<0.001
Credit card interest charged (MXN)	55.74	29.62	25.87	18.93	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	0.65%	1.59%	1.94%	2.19%	<0.001	<0.001

Table A.8 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.9: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: SMS Security

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	47.25	43.30	46.80	40.80	<0.001	<0.001
Monthly income (MXN)	12,856.59	11,985.32	13,663.61	13,526.14	0.038	<0.001
Tenure with bank (months)	103.05	66.87	97.53	64.37	<0.001	<0.001
Total amount of debit card trans (MXN)	1,025.73	1,095.20	1,220.13	1,572.48	<0.001	<0.001
Debit card spending on Necessity MCC (MXN)	111.12	109.71	118.94	354.97	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	0.09	0.22	0.42	256.70	<0.001	<0.001
Partner-bank credit card limit (MXN)	3,625.72	3,502.35	5,544.80	12,599.26	<0.001	<0.001
Monthly credit card closing balance (MXN)	1,118.19	1,194.64	1,839.04	3,836.17	<0.001	<0.001
Credit card interest charged (MXN)	21.53	23.27	33.30	55.51	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	1.55%	1.69%	1.98%	1.21%	0.037	<0.001

Table A.9 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.10: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: SMS Convenience

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	40.24	41.80	48.75	47.81	<0.001	<0.001
Monthly income (MXN)	12,288.16	11,366.89	14,046.94	14,565.68	<0.001	<0.001
Tenure with bank (months)	67.42	66.09	88.28	110.28	<0.001	<0.001
Total amount of debit card trans (MXN)	1,628.48	1,352.04	335.78	1,596.82	<0.001	<0.001
Debit card spending on Necessity MCC (MXN)	248.33	138.90	89.90	218.01	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	79.03	42.46	51.27	90.20	<0.001	<0.001
Partner-bank credit card limit (MXN)	4,944.39	4,679.57	7,407.94	8,439.85	<0.001	<0.001
Monthly credit card closing balance (MXN)	1,805.00	1,564.88	2,186.26	2,536.74	<0.001	<0.001
Credit card interest charged (MXN)	32.73	25.89	33.93	41.77	0.011	<0.001
With arrears on other partner-bank credit products (=1)	0.76%	1.77%	2.62%	1.70%	<0.001	<0.001

Table A.10 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

Table A.11: Pre-treatment Characteristics by Quartile of Predicted Treatment Effect on Number of Transactions: SMS Monitoring

Pre-treatment variable	Q1	Q2	Q3	Q4	Q1 vs Q4 (p-value)	All quartiles (p-value)
Age (years)	43.41	46.24	46.17	42.54	<0.001	<0.001
Monthly income (MXN)	14,927.93	13,877.39	13,424.17	9,827.57	<0.001	<0.001
Tenure with bank (months)	73.77	96.14	93.55	68.65	0.017	<0.001
Total amount of debit card trans (MXN)	1,410.14	1,271.14	1,216.94	949.60	0.060	0.026
Debit card spending on Necessity MCC (MXN)	305.98	144.92	133.67	105.07	<0.001	<0.001
Debit card spending on Luxury MCC (MXN)	160.56	29.22	28.61	35.01	<0.001	<0.001
Partner-bank credit card limit (MXN)	9,621.17	6,070.44	6,038.27	3,252.50	<0.001	<0.001
Monthly credit card closing balance (MXN)	3,047.14	1,846.06	1,772.72	1,054.83	<0.001	<0.001
Credit card interest charged (MXN)	47.82	32.41	31.95	19.63	<0.001	<0.001
With arrears on other partner-bank credit products (=1)	1.56%	2.12%	1.70%	1.55%	1.000	0.003

Table A.11 reports average pre-treatment characteristics by quartile of the predicted treatment effect on the number of debit card transactions, computed on the held-out evaluation sample. Q1 (Q4) denotes the quartile with the lowest (highest) predicted treatment effect. All monetary variables are in Mexican pesos (MXN), and variables marked (=1) are indicator variables. Debit card spending on Necessity (Luxury) MCC is the sum of spending in the food, personal, services, and transportation (entertainment and luxury) merchant category codes. “With arrears on other partner-bank credit products” equals one if the customer is in arrears on payroll credit, a mortgage, or an auto loan. The “Q1 vs Q4” column reports the p -value for a test of equality between the first and fourth quartiles (Wilcoxon rank-sum test for continuous variables; Pearson’s χ^2 test for indicator variables); the “All quartiles” column reports the p -value for a test of equality across all four quartiles (Kruskal–Wallis test for continuous variables; Pearson’s χ^2 test for indicator variables).

B Tables

Table B.1: Balance table

Variable	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		F-test for balance across all groups	
	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	N	Mean (SE)	F-stat	P-value
Age (years)	926544	45.890 (0.018)	124996	45.870 (0.047)	125004	45.886 (0.047)	124614	45.903 (0.048)	125008	45.900 (0.047)	16572	45.875 (0.130)	16567	45.874 (0.130)	16572	45.880 (0.130)	16772	45.870 (0.129)	16757	45.908 (0.130)	16771	45.916 (0.129)	1526177	0.482 0.903
Female (=1)	926544	0.420 (0.001)	124996	0.420 (0.001)	125004	0.420 (0.001)	124614	0.420 (0.001)	125008	0.420 (0.001)	16572	0.420 (0.004)	16567	0.420 (0.004)	16572	0.419 (0.004)	16772	0.421 (0.004)	16757	0.420 (0.004)	16771	0.421 (0.004)	1526177	1.847** 0.048
At least college-educated (=1)	926544	0.098 (0.000)	124996	0.099 (0.001)	125004	0.098 (0.001)	124614	0.098 (0.001)	125008	0.097 (0.001)	16572	0.099 (0.002)	16567	0.096 (0.002)	16572	0.100 (0.002)	16772	0.101 (0.002)	16757	0.096 (0.002)	16771	0.101 (0.002)	1526177	0.518 0.879
Monthly income (MXN)	926544	14,790.981 (21.737)	124996	14,837.954 (44.675)	125004	14,847.398 (44.744)	124614	14,813.306 (44.379)	125008	14,846.042 (44.342)	16572	14,798.571 (122.185)	16567	14,878.896 (122.296)	16572	14,692.134 (119.074)	16772	14,743.038 (123.023)	16757	14,763.243 (119.323)	16771	14,793.031 (120.632)	1526177	0.728 0.698
Amount of debit card transactions (MXN)	926544	504.741 (2.196)	124996	511.378 (3.737)	125004	502.141 (3.669)	124614	504.834 (3.697)	125008	507.975 (3.737)	16572	501.288 (10.055)	16567	496.857 (9.927)	16572	504.864 (10.075)	16772	495.687 (9.917)	16757	493.961 (9.897)	16771	505.851 (10.203)	1526177	1.086 0.368
Number of debit card transactions	926544	1.558 (0.004)	124996	1.571 (0.009)	125004	1.556 (0.009)	124614	1.556 (0.009)	125008	1.546 (0.009)	16572	1.581 (0.024)	16567	1.537 (0.024)	16572	1.538 (0.024)	16772	1.564 (0.024)	16757	1.543 (0.024)	16771	1.557 (0.024)	1526177	1.056 0.393
Amount of withdrawals (MXN)	926544	43,916.252 (233.660)	124996	44,988.508 (374.070)	125004	44,371.455 (371.939)	124614	43,791.408 (366.665)	125008	44,395.531 (369.977)	16572	43,611.739 (967.474)	16567	45,144.405 (1,017.574)	16572	44,687.211 (1,009.149)	16772	42,802.436 (978.289)	16757	44,523.689 (1,031.477)	16771	43,541.469 (994.472)	1526177	1.207 0.280
Number of withdrawal transactions	926544	6.883 (0.009)	124996	6.930 (0.020)	125004	6.880 (0.020)	124614	6.885 (0.019)	125008	6.854 (0.019)	16572	6.840 (0.053)	16567	6.805 (0.053)	16572	6.846 (0.053)	16772	6.900 (0.054)	16757	6.877 (0.054)	16771	6.868 (0.054)	1526177	1.288 0.230
With partner-bank credit card (=1)	926544	0.206 (0.000)	124996	0.206 (0.000)	125004	0.206 (0.001)	124614	0.205 (0.001)	125008	0.206 (0.001)	16572	0.204 (0.003)	16567	0.204 (0.003)	16572	0.205 (0.003)	16772	0.208 (0.003)	16757	0.208 (0.003)	16771	0.207 (0.003)	1526177	0.090 1.000
Partner-bank credit card limit (MXN)	926544	9,862.999 (47.014)	124996	9,969.080 (93.474)	125004	9,849.201 (92.604)	124614	9,887.953 (93.002)	125008	9,898.609 (92.972)	16572	9,338.767 (244.771)	16567	9,838.873 (255.078)	16572	10,190.331 (268.658)	16772	10,010.188 (257.338)	16757	10,110.992 (257.154)	16771	10,051.296 (256.946)	1526177	1.232 0.264
Partner-bank credit card tenure (months)	926544	7.663 (0.033)	124996	7.778 (0.070)	125004	7.684 (0.070)	124614	7.660 (0.070)	125008	7.674 (0.070)	16572	7.287 (0.182)	16567	7.591 (0.190)	16572	7.627 (0.189)	16772	7.875 (0.195)	16757	7.756 (0.191)	16771	7.749 (0.192)	1526177	0.918 0.515
Partner-bank credit card closing bal (MXN)	926544	2,836.063 (14.835)	124996	2,872.014 (20.337)	125004	2,853.867 (20.224)	124614	2,836.883 (20.159)	125008	2,827.745 (20.175)	16572	2,762.876 (82.117)	16567	2,803.485 (82.758)	16572	2,969.314 (86.323)	16772	2,914.474 (83.887)	16757	2,893.780 (82.735)	16771	2,903.912 (83.616)	1526177	0.301 0.951
Partner-bank credit card int (MXN)	926544	45.786 (0.283)	124996	44.397 (0.600)	125004	46.217 (0.621)	124614	45.582 (0.614)	125008	44.595 (0.607)	16572	43.535 (1.649)	16567	44.526 (1.643)	16572	44.260 (1.641)	16772	45.182 (1.715)	16757	47.664 (1.715)	16771	46.006 (1.673)	1526177	1.300 0.223
Number of Bank credit card transactions	926544	0.706 (0.004)	124996	0.725 (0.009)	125004	0.712 (0.009)	124614	0.708 (0.009)	125008	0.697 (0.009)	16572	0.657 (0.023)	16567	0.680 (0.024)	16572	0.707 (0.025)	16772	0.723 (0.026)	16757	0.734 (0.025)	16771	0.713 (0.025)	1526177	1.298 0.225
Tenure with Bank (months)	926544	91.622 (0.085)	124996	91.499 (0.218)	125004	91.500 (0.219)	124614	91.443 (0.218)	125008	91.691 (0.219)	16572	90.939 (0.599)	16567	91.815 (0.604)	16572	91.281 (0.598)	16772	91.547 (0.598)	16757	91.164 (0.592)	16771	91.553 (0.597)	1526177	1.051 0.397
F-test of joint significance (F-stat)																								
F-test, number of observations																								

Variable	(2)-(1)		(3)-(1)		(4)-(1)		(5)-(1)		(6)-(1)		(7)-(1)		(8)-(1)		(9)-(1)		(10)-(1)		(11)-(1)	
	N	Normalized difference	N	Normalized difference	N	Normalized difference	N	Normalized difference	N	Pairwise t-test	N	Normalized difference	N	Normalized difference	N	Normalized difference	N	Normalized difference	N	Normalized difference
Age (years)	1051540	-0.001	1051548	-0.000	1051158	0.001	1051552	0.001	943116	-0.001	943111	-0.001	943116	-0.000	943316	-0.001	943301	0.001	943315	0.002
Female (=1)	1051540	0.000	1051548	0.000	1051158	-0.000	1051552	0.000	943116	-0.001	943111	0.000***	943116	-0.002	943316	0.001	943301	0.000	943315	0.002
At least college-educated (=1)	1051540	0.002	1051548	0.001	1051158	-0.000	1051552	-0.002	943116	0.002	943111	-0.007	943116	0.006	943316	0.009	943301	-0.007	943315	0.008
Monthly income (MXN)	1051540	0.003	1051548	0.003	1051158	0.001	1051552	0.003	943116	0.000	943111	0.005	943116	-0.006	943316	-0.003	943301	-0.002	943315	0.000
Amount of debit card transactions (MXN)	1051540	0.004*	1051548	-0.002	1051158	0.000	1051552	0.002	943116	-0.002	943111	-0.005	943116	0.000	943316	-0.006	943301	-0.007	943315	0.001
Number of debit card transactions	1051540	0.004	1051548	-0.000	1051158	-0.002	1051552	-0.004	943116	0.007	943111	-0.006	943116	-0.006	943316	0.002	943301	-0.004	943315	-0.000
Amount of withdrawals (MXN)	1051540	0.005*	1051548	0.003	1051158	-0.001	1051552	0.003	943116	-0.002	943111	0.008	943116	0.005	943316	-0.007	943301	0.004	943315	-0.002
Number of withdrawal transactions	1051540	0.007**	1051548	-0.000	1051158	0.000	1051552	-0.004	943116	-0.006	943111	-0.003	943116	-0.005	943316	0.002	943301	-0.001	943315	-0.002
With Partner-bank credit card (=1)	1051540	-0.000	1051548	-0.000	1051158	-0.002	1051552	-0.000	943116	-0.006	943111	-0.006	943116	-0.004	943316	0.005	943301	0.004	943315	0.003
Partner-bank credit card limit (MXN)	1051540	0.003	1051548	-0.000	1051158	0.001	1051552	0.001	943116	-0.015**	943111	-0.001	943116	0.009*	943316	0.004	943301	0.007	943315	0.005
Partner-bank credit card tenure (months)	1051540	0.005**	1051548	0.001	1051158	-0.000	1051552	0.000	943116	-0.015**	943111	-0.003	943116	-0.001	943316	0.008	943301	0.004	943315	0.003
Partner-bank credit card closing bal (MXN)	1051540	0.001	1051548	-0.000	1051158	-0.002	1051552	-0.003	943116	-0.008	943111	-0.005	943116	0.005	943316	0.005	943301	0.003	943315	0.004
Partner-bank credit card int (MXN)	1051540	-0.006**	1051548	0.002	1051158	-0.001	1051552	-0.005*	943116	-0.010	943111	-0.006	943116	-0.007	943316	-0.003	943301	0.008	943315	0.001
Number of Partner-bank credit card transactions	1051540	0.000**	1051548	0.002	1051158	0.001	1051552	-0.003	943116	-0.015*	943111	-0.008	943116	0.000	943316	0.005	943301	0.009	943315	0.002
Tenure with Bank (months)	1051540	-0.002	1051548	-0.000	1051158	-0.002	1051552	0.001	943116	-0.009*	943111	0.003	943116	-0.005	943316	-0.001	943301	-0.006*	943315	-0.001
F-test of joint significance (F-stat)	2.178***		0.375		0.518		0.927		1.189		1.505*		0.892		0.567		0.734		0.314	
F-test, number of observations	1051540		1051548		1051158		1051552		943116		943111		943116		943316		943301		943315	

Fixed effect used in pairwise and f-test regressions: [strata]. Sampling probability weights are applied as non payroll clients were oversampled in the experiment. Robust standard errors in parentheses. Significance: ***=.01, **=.05, *=.1.

Table B.2: Average Treatment Effect of Types of Incentives

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
300MXN cash	0.024 (0.006) [0.000]*** {0.001}*** {{0.001}}***	0.024 (0.006) [0.000]*** {0.001}*** {{0.001}}***	0.038 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.038 (0.004) [0.000]*** {0.001}*** {{0.001}}***
60 MXN cash	0.008 (0.006) [0.174] {0.046}** {{0.043}}**	0.008 (0.006) [0.175] {0.046}** {{0.043}}**	0.017 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.017 (0.004) [0.000]*** {0.001}*** {{0.001}}***
300MXN Liverpool voucher	0.013 (0.006) [0.032]** {0.022}** {{0.003}}***	0.013 (0.006) [0.032]** {0.022}** {{0.003}}***	0.019 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.019 (0.004) [0.000]*** {0.001}*** {{0.001}}***
60MXN 2-for-1 movie ticket	0.013 (0.006) [0.027]** {0.022}** {{0.003}}***	0.013 (0.006) [0.027]** {0.022}** {{0.003}}***	0.008 (0.004) [0.052]* {0.031}** {{0.008}}***	0.008 (0.004) [0.052]* {0.031}** {{0.009}}***
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	2,824,937	2,824,937	2,824,937	2,824,937
p(300MXN cash=60MXN cash)	0.043	0.043	0.000	0.000
p(300MXN cash=Liverpool voucher)	0.166	0.166	0.001	0.001
p(300MXN cash=Movie ticket)	0.175	0.175	0.000	0.000
p(60MXN cash=Liverpool voucher)	0.530	0.530	0.714	0.715
p(60MXN cash=Movie ticket)	0.504	0.504	0.100	0.099
p(Liverpool voucher=Movie ticket)	0.970	0.971	0.045	0.045

Table B.2 shows the average treatment effect by type of incentive. All columns are estimated using Poisson regression (equation 1) and the control group is the base category. The sample pools November and December 2019 and all columns include month fixed effects; the main text reports November only. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }}. *** p<0.01, ** p<0.05, * p<0.1

Table B.3: Average Treatment Effect of SMS reminders only

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
Security SMS	0.035 (0.015) [0.020]** {0.140} {{0.001}}***	0.035 (0.015) [0.020]** {0.140} {{0.001}}***	0.015 (0.011) [0.174] {0.770} {{0.146}}	0.015 (0.011) [0.174] {0.767} {{0.147}}
Convenience SMS	-0.009 (0.015) [0.574] {1.000} {{0.681}}	-0.008 (0.015) [0.575] {1.000} {{0.689}}	0.010 (0.011) [0.376] {1.000} {{0.467}}	0.010 (0.011) [0.374] {1.000} {{0.463}}
Monitoring SMS	0.006 (0.015) [0.689] {1.000} {{0.681}}	0.006 (0.015) [0.689] {1.000} {{0.689}}	0.007 (0.011) [0.542] {1.000} {{0.681}}	0.007 (0.011) [0.546] {1.000} {{0.689}}
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	1,933,332	1,933,332	1,933,332	1,933,332
p(SMS Security=SMS Convenience)	0.039	0.039	0.731	0.733
p(SMS Security=SMS Monitoring)	0.169	0.169	0.591	0.587
p(SMS Convenience=SMS Monitoring)	0.491	0.491	0.845	0.839

Table B.3 shows the average treatment effect of SMS reminders. All columns are estimated using Poisson regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }}. *** p<0.01, ** p<0.05, * p<0.1

Table B.4: Average Treatment Effect of SMS reminders plus unpooled cash

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
Security SMS + 300MXN cash	0.028 (0.021) [0.174] {0.256} {{0.235}}	0.028 (0.021) [0.174] {0.256} {{0.235}}	0.050 (0.015) [0.001]*** {0.011}** {{0.001}}***	0.050 (0.015) [0.001]*** {0.011}** {{0.001}}***
Convenience SMS + 300MXN cash	0.011 (0.022) [0.616] {0.587} {{0.804}}	0.011 (0.022) [0.617] {0.588} {{0.806}}	0.006 (0.015) [0.707] {0.628} {{0.830}}	0.006 (0.015) [0.708] {0.630} {{0.833}}
Monitoring SMS + 300MXN cash	0.017 (0.020) [0.389] {0.516} {{0.618}}	0.017 (0.020) [0.389] {0.516} {{0.621}}	0.042 (0.015) [0.006]*** {0.036}** {{0.001}}***	0.042 (0.015) [0.006]*** {0.036}** {{0.001}}***
Security SMS + 60MXN cash	0.001 (0.021) [0.968] {0.806} {{0.958}}	0.001 (0.021) [0.968] {0.804} {{0.961}}	0.035 (0.015) [0.017]** {0.059}* {{0.003}}***	0.035 (0.015) [0.017]** {0.059}* {{0.003}}***
Convenience SMS + 60MXN cash	0.013 (0.021) [0.535] {0.555} {{0.793}}	0.013 (0.021) [0.535] {0.555} {{0.793}}	0.030 (0.016) [0.052]* {0.092}* {{0.028}}**	0.030 (0.016) [0.052]* {0.092}* {{0.026}}**
Monitoring SMS + 60MXN cash	0.024 (0.022) [0.260] {0.352} {{0.411}}	0.024 (0.022) [0.260] {0.352} {{0.411}}	0.035 (0.016) [0.027]** {0.071}* {{0.012}}**	0.035 (0.016) [0.027]** {0.071}* {{0.012}}**
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	1,934,665	1,934,665	1,934,665	1,934,665
p(Security SMS + 300MXN cash=Convenience SMS + 300MXN cash)	0.556	0.555	0.037	0.037
p(Security SMS + 300MXN cash=Monitoring SMS + 300MXN cash)	0.704	0.704	0.688	0.686
p(Security SMS + 300MXN cash=Security SMS + 60MXN cash)	0.353	0.352	0.483	0.481
p(Security SMS + 300MXN cash=Convenience SMS + 60MXN cash)	0.604	0.604	0.361	0.361
p(Security SMS + 300MXN cash=Monitoring SMS + 60MXN cash)	0.889	0.889	0.481	0.482
p(Convenience SMS + 300MXN cash=Monitoring SMS + 300MXN cash)	0.822	0.821	0.093	0.093
p(Convenience SMS + 300MXN cash=Security SMS + 60MXN cash)	0.740	0.741	0.159	0.159
p(Convenience SMS + 300MXN cash=Convenience SMS + 60MXN cash)	0.940	0.939	0.257	0.255
p(Convenience SMS + 300MXN cash=Monitoring SMS + 60MXN cash)	0.658	0.657	0.182	0.180
p(Monitoring SMS + 300MXN cash=Monitoring SMS + 60MXN cash)	0.816	0.816	0.757	0.759
p(Monitoring SMS + 300MXN cash=Convenience SMS + 60MXN cash)	0.881	0.881	0.605	0.606
p(Monitoring SMS + 300MXN cash=Security SMS + 60MXN cash)	0.570	0.569	0.769	0.769
p(Security SMS + 60MXN cash=Convenience SMS + 60MXN cash)	0.681	0.681	0.814	0.816
p(Security SMS + 60MXN cash=Monitoring SMS + 60MXN cash)	0.437	0.436	0.979	0.982
p(Convenience SMS + 60MXN cash=Monitoring SMS + 60MXN cash)	0.710	0.710	0.840	0.839

Table B.4 shows the average treatment effect of SMS reminders plus unpooled cash. All columns are estimated using Poisson regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }}. *** p<0.01, ** p<0.05, * p<0.1

Table B.5: Average Treatment Effect of SMS reminders plus cash

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
Security SMS + cash	0.015 (0.015) [0.322] {0.254} {{0.264}}	0.015 (0.015) [0.322] {0.254} {{0.264}}	0.043 (0.011) [0.000]*** {0.001}*** {{0.001}}***	0.043 (0.011) [0.000]*** {0.001}*** {{0.001}}***
Convenience SMS + cash	0.012 (0.015) [0.431] {0.276} {{0.264}}	0.012 (0.015) [0.432] {0.276} {{0.264}}	0.018 (0.011) [0.101] {0.156} {{0.044}}**	0.018 (0.011) [0.101] {0.156} {{0.044}}**
Monitoring SMS + cash	0.021 (0.015) [0.162] {0.193} {{0.094}}*	0.021 (0.015) [0.161] {0.193} {{0.093}}*	0.038 (0.011) [0.001]*** {0.002}*** {{0.001}}***	0.038 (0.011) [0.001]*** {0.002}*** {{0.001}}***
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	1,934,665	1,934,665	1,934,665	1,934,665
p(SMS Security plus cash=SMS Convenience plus cash)	0.889	0.889	0.099	0.099
p(SMS Security plus cash=SMS Monitoring plus cash)	0.776	0.776	0.757	0.758
p(SMS Convenience plus cash=SMS Monitoring plus cash)	0.673	0.672	0.188	0.187

Table B.5 shows the average treatment effect of SMS reminders plus cash. All columns are estimated using Poisson regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }}. *** p<0.01, ** p<0.05, * p<0.1

Table B.6: Average Treatment Effect of Duration

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
2 months	0.019 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.019 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.026 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.026 (0.003) [0.000]*** {0.001}*** {{0.001}}***
1 month	0.009 (0.005) [0.042]** {0.011}** {{0.001}}***	0.009 (0.005) [0.042]** {0.011}** {{0.001}}***	0.017 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.017 (0.003) [0.000]*** {0.001}*** {{0.001}}***
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	2,924,617	2,924,617	2,924,617	2,924,617
p(2 months=1 month)	0.065	0.065	0.014	0.014

Table B.6 shows the average treatment effect of duration. All columns are estimated using Poisson regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }. *** p<0.01, ** p<0.05, * p<0.1

Table B.7: Average Treatment Effect of Goal Periodicity

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
Monthly	0.012 (0.005) [0.012]** {0.004}*** {{0.001}}***	0.012 (0.005) [0.012]** {0.004}*** {{0.001}}***	0.022 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.022 (0.003) [0.000]*** {0.001}*** {{0.001}}***
Weekly	0.017 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.017 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.022 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.022 (0.003) [0.000]*** {0.001}*** {{0.001}}***
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	2,924,617	2,924,617	2,924,617	2,924,617
p(Monthly=Weekly)	0.273	0.272	0.888	0.892

Table B.7 shows the average treatment effect of goal periodicity. All columns are estimated using Poission regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }}. *** p<0.01, ** p<0.05, * p<0.1

Table B.8: Average Treatment Effect of Goal Type

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
Amount of transactions	0.013 (0.005) [0.005]*** {0.003}*** {{0.001}}***	0.013 (0.005) [0.006]*** {0.003}*** {{0.001}}***	0.010 (0.003) [0.002]*** {0.002}*** {{0.001}}***	0.010 (0.003) [0.002]*** {0.002}*** {{0.001}}***
Number of transactions	0.016 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.016 (0.004) [0.000]*** {0.001}*** {{0.001}}***	0.032 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.032 (0.003) [0.000]*** {0.001}*** {{0.001}}***
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	2,924,617	2,924,617	2,924,617	2,924,617
p(Amount=Transactions)	0.514	0.514	0.000	0.000

Table B.8 shows the average treatment effect of goal type. All columns are estimated using Poission regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }}. *** p<0.01, ** p<0.05, * p<0.1

Table B.9: Average Treatment Effect of Goal Intensity

	Outcome Variable: Amount of Debit Card Trans.		Outcome Variable: Number of Debit Card Trans.	
	(1)	(2)	(3)	(4)
Half	0.014 (0.004) [0.001]*** {0.001}*** {{0.001}}***	0.014 (0.004) [0.001]*** {0.001}*** {{0.001}}***	0.023 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.023 (0.003) [0.000]*** {0.001}*** {{0.001}}***
High	0.015 (0.005) [0.001]*** {0.001}*** {{0.001}}***	0.015 (0.005) [0.001]*** {0.001}*** {{0.001}}***	0.021 (0.003) [0.000]*** {0.001}*** {{0.001}}***	0.021 (0.003) [0.000]*** {0.001}*** {{0.001}}***
Mean of control group	1,137.28	1,137.28	1.90	1.90
Strata	✓	✓	✓	✓
Month	✓	✓	✓	✓
Baseline outcome	✓	✓	✓	✓
Controls		✓		✓
Observations	2,924,617	2,924,617	2,924,617	2,924,617
p(High=Half)	0.904	0.904	0.715	0.714

Table B.9 shows the average treatment effect of goal intensity. All columns are estimated using Poisson regression (equation 1) and the control group is the base category. Both outcome variables are winsorized at 1th and 99th percentiles. Columns 2 and 4 are the preferred specifications and they control for the imbalanced baseline covariates. In the bottom panel, we report the p-value from testing the equality of two coefficients, where the null hypothesis is that the coefficients are equal. Sampling probability weights are applied to make inferences about the treatment effects on the bank's debit card customer population as non-payroll clients were oversampled in the experiment. Robust standard errors in round parentheses (). Unadjusted p-values in brackets []. Sharpened q-values in curly brackets { }. Romano-Wolf stepdown adjusted p-values in double curly brackets {{ }. *** p<0.01, ** p<0.05, * p<0.1

Table B.10: Heterogeneous Treatment Effects, by Experimental Strata

Group definition	Outcome Variable: Amount of Debit Card Trans.						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Sex	Has Bank Credit Card?	Managed by Payroll group	Median of tenure with bank	Level of transactionality	Quartiles of age	Quartiles of income
Omitted Category (Group ₁)	Male	No	Yes	M1	Low	Q1	Q1
300MXN cashback*Group ₂	-0.023 (0.016)	0.002 (0.018)	-0.013 (0.016)	0.026* (0.016)	-0.069 (0.058)	0.002 (0.021)	0.073*** (0.026)
300MXN cashback*Group ₃					-0.120** (0.056)	0.049** (0.022)	0.055** (0.025)
300MXN cashback*Group ₄						0.016 (0.023)	0.047** (0.022)
60MXN cashback*Group ₂	0.003 (0.016)	0.003 (0.018)	0.011 (0.016)	-0.001 (0.016)	0.022 (0.059)	0.012 (0.021)	0.005 (0.026)
60MXN cashback*Group ₃					0.006 (0.057)	-0.000 (0.022)	-0.011 (0.025)
60MXN cashback*Group ₄						0.030 (0.023)	0.000 (0.023)
Liverpool voucher*Group ₂	0.023 (0.016)	-0.021 (0.018)	-0.007 (0.016)	-0.009 (0.016)	-0.160*** (0.058)	0.013 (0.021)	0.009 (0.027)
Liverpool voucher*Group ₃					-0.158*** (0.056)	-0.019 (0.021)	-0.032 (0.026)
Liverpool voucher*Group ₄						-0.005 (0.023)	-0.015 (0.023)
2-for-1 movie ticket*Group ₂	0.002 (0.016)	-0.021 (0.018)	0.006 (0.016)	0.012 (0.016)	-0.021 (0.062)	-0.013 (0.021)	0.009 (0.027)
2-for-1 movie ticket*Group ₃					-0.004 (0.060)	0.015 (0.022)	0.005 (0.026)
2-for-1 movie ticket*Group ₄						0.007 (0.023)	0.005 (0.023)
SMS Security*Group ₂	0.016 (0.039)	-0.000 (0.044)	-0.009 (0.038)	0.021 (0.039)	0.050 (0.149)	-0.047 (0.053)	0.145** (0.065)
SMS Security*Group ₃					-0.009 (0.144)	-0.045 (0.055)	0.159** (0.065)
SMS Security*Group ₄						-0.038 (0.057)	0.175*** (0.058)
SMS Monitoring*Group ₂	-0.007 (0.039)	-0.005 (0.043)	0.048 (0.038)	0.091** (0.039)	0.065 (0.165)	0.067 (0.053)	0.033 (0.072)
SMS Monitoring*Group ₃					0.070 (0.161)	0.054 (0.054)	-0.021 (0.067)
SMS Monitoring*Group ₄						0.017 (0.056)	0.021 (0.059)
SMS Convenience*Group ₂	-0.015 (0.040)	0.064 (0.045)	-0.020 (0.039)	0.075* (0.040)	-0.106 (0.155)	0.035 (0.055)	-0.038 (0.067)
SMS Convenience*Group ₃					-0.021 (0.150)	-0.069 (0.055)	-0.011 (0.066)
SMS Convenience*Group ₄						0.056 (0.059)	-0.043 (0.059)
SMS Security plus cash*Group ₂	-0.007 (0.038)	-0.084* (0.044)	-0.023 (0.037)	-0.033 (0.038)	-0.059 (0.149)	0.031 (0.051)	0.037 (0.066)
SMS Security plus cash*Group ₃					-0.079 (0.144)	0.025 (0.053)	-0.004 (0.063)
SMS Security plus cash*Group ₄						0.112** (0.056)	-0.063 (0.056)
SMS Monitoring plus cash*Group ₂	0.110*** (0.038)	0.005 (0.043)	0.022 (0.038)	-0.017 (0.038)	-0.183 (0.146)	0.007 (0.052)	-0.082 (0.066)
SMS Monitoring plus cash*Group ₃					-0.239* (0.142)	0.042 (0.053)	-0.019 (0.066)
SMS Monitoring plus cash*Group ₄						0.062 (0.057)	0.044 (0.057)
SMS Convenience plus cash*Group ₂	0.023 (0.040)	-0.039 (0.046)	-0.071* (0.038)	-0.035 (0.039)	-0.257* (0.141)	0.050 (0.054)	0.052 (0.069)
SMS Convenience plus cash*Group ₃					-0.256* (0.136)	-0.065 (0.055)	0.080 (0.065)
SMS Convenience plus cash*Group ₄						-0.049 (0.058)	0.036 (0.058)
Observations	1,526,177	1,526,177	1,526,177	1,526,177	1,526,177	1,526,177	1,526,177

This table presents heterogeneous treatment effects on the amount of debit card transactions using the baseline covariates that define the experimental strata. The treatment effects are estimated using Poisson regression with the following specification: $y_i = \alpha_s + Treatment_i + Group_{i,j} + Treatment * Group_{i,j} + controls + \epsilon_i$ where α_s represents strata fixed effects, $Group_{i,j}$ is a dummy variable indicating whether customer i belongs to group j , and $controls$ are aspects of the "challenge" structure (goal type, goal intensity, duration, and goal condition). For all specifications Group 1 is the base category defined as: Sex (male), Has partner-bank credit card? (no), Managed by payroll group? (yes), Median tenure with bank (M1), Level of transactionality (low level), Quartiles of age (Q1), and Quartiles of income (Q1). Non-omitted categories for the Level of Transactionality variable: medium (Group 2) and high (Group 3). Robust standard errors in parentheses. Sampling weights were applied due to the oversampling of customers from the non-payroll group..

*** p<0.01, ** p<0.05, * p<0.1

C Figures

Figure C.1: Full list of treatment arms

	Treatment	Type of incentive	Amount of incentive	Goal intensity	Duration of participation	Goal type	Goal periodicity
Cash incentives	1	cash	60	high	1 month	amount	weekly
	2	cash	60	high	1 month	transactions	weekly
	3	cash	60	high	1 month	amount	monthly
	4	cash	60	high	1 month	transactions	monthly
	5	cash	60	high	2 months	amount	weekly
	6	cash	60	high	2 months	transactions	weekly
	7	cash	60	high	2 months	amount	monthly
	8	cash	60	high	2 months	transactions	monthly
	9	cash	60	half	1 month	amount	weekly
	10	cash	60	half	1 month	transactions	weekly
	11	cash	60	half	1 month	amount	monthly
	12	cash	60	half	1 month	transactions	monthly
	13	cash	60	half	2 months	amount	weekly
	14	cash	60	half	2 months	transactions	weekly
	15	cash	60	half	2 months	amount	monthly
	16	cash	60	half	2 months	transactions	monthly
	17	cash	300	high	1 month	amount	weekly
	18	cash	300	high	1 month	transactions	weekly
	19	cash	300	high	1 month	amount	monthly
	20	cash	300	high	1 month	transactions	monthly
	21	cash	300	high	2 months	amount	weekly
	22	cash	300	high	2 months	transactions	weekly
	23	cash	300	high	2 months	amount	monthly
	24	cash	300	high	2 months	transactions	monthly
	25	cash	300	half	1 month	amount	weekly
	26	cash	300	half	1 month	transactions	weekly
	27	cash	300	half	1 month	amount	monthly
	28	cash	300	half	1 month	transactions	monthly
	29	cash	300	half	2 months	amount	weekly
	30	cash	300	half	2 months	transactions	weekly
	31	cash	300	half	2 months	amount	monthly
	32	cash	300	half	2 months	transactions	monthly
In-kind incentives	33	in-kind	60	high	1 month	amount	weekly
	34	in-kind	60	high	1 month	transactions	weekly
	35	in-kind	60	high	1 month	amount	monthly
	36	in-kind	60	high	1 month	transactions	monthly
	37	in-kind	60	high	2 months	amount	weekly
	38	in-kind	60	high	2 months	transactions	weekly
	39	in-kind	60	high	2 months	amount	monthly
	40	in-kind	60	high	2 months	transactions	monthly
	41	in-kind	60	half	1 month	amount	weekly
	42	in-kind	60	half	1 month	transactions	weekly
	43	in-kind	60	half	1 month	amount	monthly
	44	in-kind	60	half	1 month	transactions	monthly
	45	in-kind	60	half	2 months	amount	weekly
	46	in-kind	60	half	2 months	transactions	weekly
	47	in-kind	60	half	2 months	amount	monthly
	48	in-kind	60	half	2 months	transactions	monthly
	49	in-kind	300	high	1 month	amount	weekly
	50	in-kind	300	high	1 month	transactions	weekly
	51	in-kind	300	high	1 month	amount	monthly
	52	in-kind	300	high	1 month	transactions	monthly
	53	in-kind	300	high	2 months	amount	weekly
	54	in-kind	300	high	2 months	transactions	weekly
	55	in-kind	300	high	2 months	amount	monthly
	56	in-kind	300	high	2 months	transactions	monthly
	57	in-kind	300	half	1 month	amount	weekly
	58	in-kind	300	half	1 month	transactions	weekly
	59	in-kind	300	half	1 month	amount	monthly
	60	in-kind	300	half	1 month	transactions	monthly
	61	in-kind	300	half	2 months	amount	weekly
	62	in-kind	300	half	2 months	transactions	weekly
	63	in-kind	300	half	2 months	amount	monthly
	64	in-kind	300	half	2 months	transactions	monthly
SMS	65 - SMS security	NA	NA	NA	NA	NA	NA
	66 - SMS Convenience	NA	NA	NA	NA	NA	NA
	67 - SMS Monitoring	NA	NA	NA	NA	NA	NA
SMS plus cash	68 - SMS Security	cash	60	half	2 months	transactions	weekly
	69 - SMS Security	cash	300	half	2 months	transactions	weekly
	70 - SMS Convenience	cash	60	half	2 months	transactions	weekly
	71 - SMS Convenience	cash	300	half	2 months	transactions	weekly
	72 - SMS Monitoring	cash	60	half	2 months	transactions	weekly
	73 - SMS Monitoring	cash	300	half	2 months	transactions	weekly

Notes: NA means “not applicable”. Treatments 1-64 were created by cross-randomizing (full factorial: $2^6 = 64$) the following aspects of the challenge: (i) types of incentives (cash or in-kind); (ii) amounts of incentives (300 MXN or 60 MXN); (iii) duration of participation (one month or two months); (iv) goal type (amount or number of debit card transactions); (v) goal intensity (half or high target); and (vi) goal periodicity (weekly or monthly target). Under SMS only, there were three treatment arms corresponding to the themes of the text messages. Finally, SMS plus cash had six treatment arms resulting from combining each SMS message with either 300 MXN or 60 MXN cash.

D Costs and completion rate assumptions

The costs in Section 6.5 decompose into a communication cost paid for every treated customer and a reward paid only on completion. Seven weekly SMS messages at 0.65 MXN give $f = 4.55$ MXN. The reward costs the bank r_k when it is claimed: 10 MXN for the movie ticket and 240 MXN for the Liverpool voucher, the bulk purchase prices of vouchers with face values of 60 and 300 MXN, and 75 and 375 MXN for the 60 and 300 MXN cash incentives, which are the nominal amounts multiplied by 1.25 because the bank pays 1.25 MXN in tax for every MXN of cashback. Writing $\bar{p} = 2.74\%$ for the aggregate completion rate, $f + r_k\bar{p}$ reproduces each of the seven costs used in the main text to the cent. The three SMS treatments have $r_k = 0$.

The fixed-cost calculations charge every treated customer the same $f + r_k\bar{p}$. To let cost rise with the predicted effect we replace the aggregate rate with a customer-specific completion probability, so that the cost of treatment k for customer i is

$$c_k(X_i) = c_k + r_k(p_k(X_i) - \bar{p}), \quad p_k(X_i) = \frac{1}{1 + \exp\{-a - b z_k(X_i)\}},$$

where c_k is the published cost of treatment k and $z_k(X_i)$ is the customer’s calibrated predicted response to it, standardized over the four incentive treatments pooled. The intercept a is common to the four treatments and solves $\mathbb{E}[p] = \bar{p}$ over them jointly. Completion therefore differs across treatments according to their average effects, rather than being held at $\bar{p}$ within each. Anchoring on c_k rather than on $f + r_k\bar{p}$ absorbs the rounding in the published figures, so that $b = 0$ reproduces the fixed-cost columns exactly. Completion varies within a treatment only where its forest passed the rank-weighted average treatment effect test; where it did not, every customer is assigned that treatment’s arm-level rate.

We set $b = 1.2$ in the main text. At that value the top decile of predicted response completes with probability 0.159 on the amount of debit card transactions and 0.141 on the number of transactions, against an aggregate rate of 2.74 per cent, and the bottom decile with probability 0.005 and 0.002. Table E.1 reports $b \in \{0, 0.4, 1.2, 2, 2.9\}$ at half the bank’s budget, at its budget, and at twice its budget.

E Cost-effectiveness across budgets and cost gradients

Cost-effectiveness falls as the budget rises. Under constant cost the personalized policy returns 20.39 MXN of debit card use per MXN spent at half the bank’s budget, 13.29 at its budget and 4.32 at twice its budget, treating 24.5, 39.3 and 100.0 per cent of the client base. The same policy returns 0.0615, 0.0290 and 0.0177 additional transactions per MXN at the three budgets. A larger budget extends a policy to customers with smaller predicted

responses, and some policies stop before it is spent: the 60 MXN cash incentive treats 51.7 per cent of the client base at the bank's budget and the same share at twice it.

Cost-effectiveness is not monotone in b . For the 300 MXN cash incentive at the bank's budget it falls from 12.70 MXN of debit card use per MXN spent at $b = 0$ to 2.82 at $b = 2$, then rises to 5.39 at $b = 2.9$, while the share of the client base treated goes from 24.0 to 26.3 to 37.9 per cent. Cost falls as well as rises with the predicted effect: a customer predicted to respond weakly is one the bank is unlikely to have to pay, so a steep gradient brings that customer's cost close to the 4.55 MXN of communication alone, and a selective policy can then be satisfied by customers who are cheap rather than responsive. Because we observe only the aggregate completion rate, and no completion outcome at the customer level, b cannot be estimated.

Table E.1: Targeted policies across the full range of cost gradients and budgets

<i>Panel A. MXN of debit card use per MXN spent</i>				
Budget		2,718,847	5,437,693	10,875,387
300 MXN cash incentive	no gradient	18.71 [12.0]	12.70 [24.0]	9.12 [48.1]
	$b = 0.4$	11.08 [5.2]	8.68 [11.5]	6.75 [27.0]
	$b = 1.2$	5.60 [7.7]	4.46 [15.4]	3.33 [30.0]
	$b = 2$	2.89 [13.8]	2.82 [26.3]	3.74 [45.3]
	$b = 2.9$	3.67 [21.2]	5.39 [37.9]	4.13 [56.8]
60 MXN cash incentive	no gradient	13.56 [27.0]	7.23 [51.7]	7.23 [51.7]
	$b = 0.4$	11.95 [22.0]	7.08 [48.6]	6.56 [51.7]
	$b = 1.2$	5.58 [17.2]	7.91 [36.5]	5.58 [51.7]
	$b = 2$	7.83 [22.0]	7.52 [33.5]	5.36 [51.7]
	$b = 2.9$	10.98 [25.6]	7.17 [36.8]	5.34 [51.7]
Monitoring SMS	any gradient	14.35 [15.7]	14.35 [15.7]	14.35 [15.7]
2-for-1 movie ticket	no gradient	—	—	—
	$b = 0.4$	—	—	—
	$b = 1.2$	—	—	—
	$b = 2$	—	—	—
	$b = 2.9$	—	—	—
Liverpool voucher	no gradient	—	—	—
	$b = 0.4$	—	—	—
	$b = 1.2$	—	—	—
	$b = 2$	—	—	—
	$b = 2.9$	—	—	—
Security SMS	any gradient	—	—	—
Convenience SMS	any gradient	—	—	—
Personalized assignment	no gradient	20.39 [24.5]	13.29 [39.3]	4.32 [100.0]
	$b = 0.4$	14.98 [26.7]	12.51 [65.8]	3.90 [100.0]
	$b = 1.2$	12.78 [24.3]	2.07 [61.5]	2.19 [100.0]
	$b = 2$	10.66 [27.8]	5.86 [66.9]	0.88 [100.0]
	$b = 2.9$	11.50 [31.0]	4.30 [67.4]	3.87 [100.0]
<i>Panel B. Debit card transactions per MXN spent</i>				
Budget		2,718,847	5,437,693	10,875,387
Convenience SMS	any gradient	0.0617 [39.1]	0.0310 [77.5]	0.0310 [77.5]
Liverpool voucher	no gradient	0.0220 [16.0]	0.0120 [32.0]	0.0079 [60.8]
	$b = 0.4$	0.0209 [9.4]	0.0124 [23.9]	0.0074 [55.6]
	$b = 1.2$	0.0062 [11.8]	0.0057 [25.5]	0.0058 [50.4]
	$b = 2$	0.0055 [19.1]	0.0037 [40.2]	0.0069 [55.5]
	$b = 2.9$	0.0056 [27.3]	0.0064 [50.7]	0.0059 [56.7]
300 MXN cash incentive	no gradient	0.0131 [12.0]	0.0110 [24.0]	0.0069 [48.0]
	$b = 0.4$	0.0096 [6.6]	0.0091 [15.6]	0.0065 [36.7]
	$b = 1.2$	0.0103 [9.0]	0.0068 [19.1]	0.0043 [40.9]
	$b = 2$	0.0045 [15.5]	0.0060 [32.3]	0.0035 [66.1]
	$b = 2.9$	0.0102 [23.9]	0.0061 [49.6]	0.0042 [74.7]
60 MXN cash incentive	no gradient	0.0128 [26.6]	0.0084 [53.9]	0.0064 [62.0]
	$b = 0.4$	0.0114 [21.4]	0.0073 [47.5]	0.0058 [62.0]
	$b = 1.2$	0.0075 [17.6]	0.0047 [35.0]	0.0049 [62.0]
	$b = 2$	0.0082 [22.1]	0.0046 [40.8]	0.0044 [62.0]
	$b = 2.9$	0.0058 [25.4]	0.0067 [47.3]	0.0043 [62.0]
2-for-1 movie ticket	no gradient	0.0067 [35.0]	0.0067 [35.0]	0.0067 [35.0]
	$b = 0.4$	0.0066 [35.0]	0.0066 [35.0]	0.0066 [35.0]
	$b = 1.2$	0.0068 [35.0]	0.0068 [35.0]	0.0068 [35.0]
	$b = 2$	0.0069 [35.0]	0.0069 [35.0]	0.0069 [35.0]
	$b = 2.9$	0.0070 [35.0]	0.0070 [35.0]	0.0070 [35.0]
Security SMS	any gradient	—	—	—
Monitoring SMS	any gradient	—	—	—
Personalized assignment	no gradient	0.0615 [39.0]	0.0290 [74.1]	0.0177 [99.0]
	$b = 0.4$	0.0617 [39.2]	0.0267 [74.1]	0.0150 [97.7]
	$b = 1.2$	0.0617 [39.2]	0.0298 [75.6]	0.0149 [97.4]
	$b = 2$	0.0617 [39.2]	0.0285 [76.1]	0.0134 [97.7]
	$b = 2.9$	0.0617 [39.2]	0.0260 [76.4]	0.0145 [97.9]

Notes: Each cell reports cost-effectiveness, the total additional debit card use a policy generates divided by what it spends, with the share of the client base it treats in brackets. Columns are three budgets: half what the experiment spent, what it spent, and twice what it spent. Rows are the five cost gradients, from constant cost to one under which the top decile of predicted response is several hundred times more likely to complete a challenge than the bottom decile; $b = 1.2$ is the gradient used in the main text and $b = 0$ reproduces the constant-cost figures of Table 7 exactly. The three SMS messages carry no conditional reward, so their costs do not depend on the gradient and they occupy a single row. Entries are — where a policy treats no one, either because its budget buys too few customers to estimate an effect or because the treatment does not clear the rule described in Table 7. Cost-effectiveness is not monotone in the gradient at the smallest budgets: cost falls as well as rises with the predicted effect, so a steep gradient makes weakly responding customers cheap enough to re-enter a selective policy.