

# Do Behavioral Frictions Prevent Firms from Adopting Profitable Opportunities?\*

Paul Gertler      Sean Higgins      Ulrike Malmendier      Waldo Ojeda

January 14, 2025

## Abstract

Firms frequently fail to adopt profitable business opportunities even when they do not face informational or liquidity constraints. We explore three behavioral frictions that explain inertia among individuals—present bias, limited memory, and distrust—in a managerial setting. In partnership with a FinTech payments company in Mexico, we randomly offer 33,978 firms the opportunity to pay a lower merchant fee. We vary whether the offer has a deadline, reminder, pre-announced reminder, and the size of the fee reduction. Reminders increase take-up by 15%, suggesting a role of memory. Announced reminders increase take-up by an additional 7%. Survey data reveal the likely mechanism: When the FinTech company follows through with the pre-announced reminder, firms’ trust in the offer increases. The deadline does not affect larger firms, implying limited or no present bias, but does increase take-up by 8% for smaller firms. Overall, behavioral frictions contribute significantly to explaining profit-reducing firm behavior.

KEYWORDS: Managerial inertia, FinTech, trust, memory, present bias, firms, Mexico

---

\* Gertler: UC Berkeley, Haas School of Business, and NBER; gertler@berkeley.edu. Higgins: Northwestern University, Kellogg School of Management; sean.higgins@kellogg.northwestern.edu. Malmendier: UC Berkeley, Department of Economics and Haas School of Business, and NBER; ulrike@berkeley.edu. Ojeda: Columbia University, Department of Economics; wo2198@columbia.edu. We are very grateful to our discussants Manuel Adelino, Tomomichi Amano, Jie Bai, Milo Bianchi, Miriam Bruhn, Michael Ewens, Camille Hebert, Rawley Heimer, Kai Li, Luca Lin, Song Ma, Michaela Pagel, Ryan Pratt, and Melanie Wallskog. We thank seminar participants at Arizona, Baruch, Ben Gurion, Cape Town, Chicago, Duke, Harvard, IMF, KAIST and Korea, Maryland, Melbourne, Monash, Northwestern, Peking, Sydney, Tulane, and Washington, as well as participants at several conferences including NBER Corporate Finance, NBER Organizational Economics, and SITE Psychology and Economics for helpful comments. We thank Noah Forougi, Mohammad Atif Haidry, Miguel Angel Jimenez, César Landín, Nicolas Min, Alexandra Wall, Honghao Wang, and Tiange Ye for research assistance. We are grateful for funding from the CEGA-Visa Financial Inclusion Lab and LIFT at UC Berkeley. The authors declare that they have no financial or material interests in the results of this research. UC Berkeley IRB approvals 2018-02-10796 and 2020-03-13091; AEA RCT Registry: AEARCTR-0006540.

# 1 Introduction

A standard assumption in classical economic models of firm behavior is that they maximize profits. In practice, however, firms often fail to adopt profitable business opportunities, including financial technologies, management practices, cost-saving machinery, and optimal pricing. The failure to adopt readily-available technologies and practices has been observed across many industries, including manufacturing, banking, retail, and healthcare.<sup>1</sup> It has caused firms of all sizes to forgo substantial profits.<sup>2</sup>

Several economic frictions contribute to firms' failure to adopt profitable opportunities, including lack of information and managerial capital, fixed costs, uncertainty, liquidity constraints, labor constraints, and principal-agent problems.<sup>3</sup> Even when these frictions are removed, though, firms frequently fail or are slow to adopt. For example, Bloom et al. (2013) find that "even if the owners became convinced of the need to adopt a practice, they would often take several months to do so," and DellaVigna and Gentzkow (2019) and Mishra et al. (2022) document "managerial inertia" as well as "stickiness in organizational structures and practices."

Why do firms exhibit such inertia and fail to take advantage of new opportunities even though they are forgoing profits? In partnership with a financial technology (FinTech) company, we design a randomized controlled trial (RCT) that allows us to assess the role of three frictions that have been shown to explain inertia and lack of behavioral change in *non-managerial* settings: present bias, limited memory, and lack of trust. Present bias affects health-related choices such as gym attendance and smoking cessation (DellaVigna and Malmendier, 2006; Giné et al., 2010) as well as financial choices such as saving, borrowing, and loan repayment (Laibson, 1997; DellaVigna and Malmendier, 2004; Ashraf et al., 2006; Kuchler and Pagel, 2021). Limited memory also hampers health-related behavior such as gym attendance and vaccine take-up (Dai et al., 2021; Calzolari and Nardotto, 2017) as well as financial choices such as saving and loan repayment (Karlan et al., 2016a,b). Finally, distrust interferes with financial decisions such as saving, borrowing, and refinancing (Karlan et al., 2009; Johnson et al., 2019; Bachas et al., 2021). We ask whether these

---

<sup>1</sup> See Atkin et al. (2017) on the failure to adopt cost-savings technologies in manufacturing; Bloom et al. (2013), Bruhn et al. (2018), and Giorcelli (2019) on management practices in manufacturing, commerce, and services firms; Mishra et al. (2022) and Higgins (2024) on financial technology adoption in banks and retail firms; Celhay et al. (2019) on the adoption of care practices by healthcare firms; and DellaVigna and Gentzkow (2019) and Strulov-Shlain (2023) on non-optimal pricing by large retail firms.

<sup>2</sup> Microenterprises in Banerjee et al. (2024b) forgo a 60% increase in profits; small and medium enterprises in Bruhn et al. (2018) forgo a 28% increase in productivity; medium and large firms in Bloom et al. (2013) forgo a 17% increase in productivity; large retail chains in DellaVigna and Gentzkow (2019) forgo 12% of profits; and Lyft forgoes \$160 million of annual profits in List et al. (2023).

<sup>3</sup> See Bloom et al. (2013), Bruhn et al. (2018), and Giorcelli (2019) on the role of information and managerial capital; Abel and Eberly (1994) on fixed costs with uncertainty; Banerjee et al. (2024a) on fixed costs with liquidity constraints; Jagannathan et al. (2016) and Hardy and McCasland (2023) on labor constraints; and Atkin et al. (2017) and Rigol and Roth (2024) on principal-agent problems.

frictions also explain profit-reducing *managerial* behavior and, as such, help resolve the puzzle of firms not adopting profitable, readily-available business opportunities.

There are several reasons why these frictions may work differently in managerial settings. First, managers, including entrepreneurs and CEOs, differ from the general population in their ability, motivation, and personality (Schoar, 2010; Kaplan and Sorensen, 2021). Managers tend to have higher levels of education, cognitive ability, optimism, and risk tolerance (Murphy et al., 1991; de Mel et al., 2010; Gennaioli et al., 2013; Graham et al., 2013; Levine and Rubinstein, 2017). Second, firms can implement management and information systems that mitigate the effects of individual biases. Partners and employees might also help reduce the effects of a manager’s biases. Finally, stakes might be higher and the adverse effects of inertia more immediate in businesses than in personal lives.

We test whether, nonetheless, these individual-level frictions help explain managerial inertia. In our RCT, a FinTech payments provider in Mexico offered 33,978 firms that were already active users of their payments technology the opportunity to be charged a lower merchant fee for each payment they receive from customers. By adopting the new contract these firms would reduce their costs. For the median firm, which is relatively small with three employees, the expected cost savings from the reduced fee equal 3% of profits.<sup>4</sup>

In order to examine the effect of these three behavioral frictions, we designed the RCT to randomly vary (i) the amount of the lower fee (2.75%, 3%, or a control group that retained their current fee of 3.5–3.75%), (ii) a reminder, (iii) whether the FinTech company told the firms in advance that they would receive a reminder (“announced reminder”), and (iv) a deadline to accept the offer. An initial message was sent to all firms receiving the offer on the first day of the experiment. The deadline was on the eighth day, and we sent reminders on the morning of the seventh day. Firms that received an announced reminder were told in the initial message on which day they would receive such a reminder.

To show how the design allows us to test for the three proposed mechanisms—present bias, limited memory, and lack of trust in other firms—we build on the model from Ericson (2017), which studies how present bias and limited memory affect task completion. We augment the model to include the notion of trust, and apply it to managerial decision-making, where the manager’s objective is to maximize the net present value of firm profits.<sup>5</sup> The model illustrates that present bias can lower adoption rates of profitable opportunities because the costs to adopt are borne immediately and the benefits are in the future. Naïve present-biased agents procrastinate thinking they will adopt tomorrow, but when tomorrow arrives their present bias causes them to procrastinate again.

---

<sup>4</sup> In the subsample of firms we surveyed, the largest firm has 150 employees. The distribution of number of employees of the firms in our RCT is similar to that of 99.7% of firms in Mexico.

<sup>5</sup> The manager’s objective function reflects that the recipient of the cost-saving offer in our RCT was the firm owner in 88.7% of cases. In most small firms, the firm owner is also the manager.

Deadlines can help overcome present bias because, at the deadline, the manager cannot delay any longer. In our model, the more present-biased a manager is, the larger is the treatment effect of a deadline.

Limited memory can also decrease adoption as managers forget about the profitable opportunity. Reminders help overcome this failure. Advance notice of a future reminder, in turn, increases managers' expectation that they will remember the offer later and thus decreases initial take-up among managers who are aware that they have limited memory.

Variation in trust generates different effects of the announced reminder and deadline. If managers lack trust in the validity of the offer, a pre-announced reminder can increase take-up: when the FinTech firm promises to send a reminder and follows through with its promise, trust in the offer and thus the perceived value of the offer increase. Deadlines could instead reduce trust as the manager might assume that the purpose of the deadline is to get them to fall for a deceptive offer or scam. The model shows that the deadline can reduce cumulative take-up by any day up until the deadline only if it reduces trust (or changes the actual or perceived probability of remembering), but not otherwise.

Our RCT allows us to test these theoretical predictions.

We estimate a strong role of limited memory, overestimation of future memory, and distrust, while present bias does not explain managerial inertia, at least not in larger firms.

First, we show that managers are forgetful: unannounced reminders cause a large and significant increase in adoption by 3.9 pp (15.2% relative to 25.5% take-up without reminder). The difference in take-up occurs almost entirely on the day we sent the reminder.

The announced-reminder group displays no difference in behavior relative to the unannounced reminder group on the first day (when we sent the initial email). This null result—combined with reminders having a large effect and not all managers adopting immediately—suggests that managers are not only forgetful but also overconfident about memory. On the day we sent the reminder, however, take-up increases by 2 pp (7.8%) more if it was pre-announced, and the difference in take-up persists over several months after the deadline.<sup>6</sup> This result cannot be explained if announced reminders only impact the probability (or perceived probability) of remembering. Instead, the announced reminder must increase the perceived value of accepting the offer, for example by increasing trust in the offer.

To better understand the announced-reminder effect, we conduct a survey of a subsample of managers in our RCT, eliciting their valuation of the offer and trust. We find that managers who received an announced reminder are 16.1 pp more likely to state that the reminder changed their perception of the offer's value (39.2% compared to 23.1% in the unannounced-reminder group).

---

<sup>6</sup> Firms in the no-deadline arm could adopt after the deadline. The effect of the announced reminder persists for the entire period we have data, both in the full sample and restricting to those without a deadline.



Moreover, the effect of the announced relative to the unannounced reminder is concentrated among managers who generally distrust advertised offers. We find similar evidence in the administrative data when we use the length of the manager’s prior business relationship with the FinTech firm as a proxy for their trust in the FinTech firm. These findings suggest that the announced reminder increases managers’ trust and, as a result, their perception of the offer’s value. Auxiliary analyses allow us to rule out alternative explanations such as different behavior induced by the announced reminder (e.g., checking the offer’s profitability in preparation for the reminder). The result on trust could have broad implications for firms’ adoption of profitable opportunities, as these often require firm-to-firm interactions where distrust may be an important barrier.

We find two interesting effects of the deadline. First, the deadline *decreases* take-up on the day the offers are sent, which is not possible in standard models of present bias. This is possible in our model only if the deadline affects trust, memory, or perceived memory. We show suggestive evidence that the channel appears to be (reduced) trust. Second, the negative effect of the deadline on day-one take-up does not persist: The treatment effect on cumulative take-up by the deadline is insignificantly positive. However, while this effect is close to zero and not statistically significant in the subsample of larger firms, the deadline does increase take-up by 2.1 pp (7.8%) among smaller firms. One possible explanation for the differential effect is that decision-makers in smaller firms exhibit more present bias, e.g., because smaller firms are run by more present-biased individuals due to selection, or because firm growth is a function of (less) present bias. Another possibility is that smaller firms lack organizational structures to mitigate managers’ present bias.<sup>7</sup>

We conclude that non-standard (behavioral) frictions are significant determinants of managerial decision-making, above and beyond the standard economic frictions analyzed in prior literature. While there is substantial evidence that these frictions induce inertia in non-managerial actions, we provide evidence on how these barriers affect managerial decisions and prevent firms from maximizing profits. As a result, firms need to apply “organizational repairs” (Camerer and Malmendier, 2007) to prevent limited memory, overconfidence about memory, and distrust from affecting managers’ decision-making. To a more limited extent the same holds for present bias, albeit mostly among smaller firms.

**Related Literature.** In *non-managerial* settings, the effects of present bias and its economic costs have been extensively studied, starting with the seminal paper by Laibson (1997). Focusing on farmers, Duflo et al. (2011) find that present bias inhibits the adoption of newer, more efficient fertilizer, but time-limited subsidies increase adoption, especially among impatient farmers.

---

<sup>7</sup> The latter mechanism would be consistent with larger firms having more formalized information and management systems (Daft, 2015) and using better management practices (Bloom et al., 2014; McKenzie and Woodruff, 2017), including being more likely to use the correct concepts from corporate finance when evaluating new investments (Graham and Harvey, 2001).

However, in many settings deadlines do not help individuals overcome present bias. For example, individuals do not switch health plans despite large benefits of doing so and a deadline imposed by the open enrollment period (Handel, 2013; Ericson, 2014). Evidence on the effectiveness of commitment devices in helping sophisticated present-biased individuals is also mixed (Bryan et al., 2010; Carrera et al., 2022).

Individuals' limited memory has been documented in a number of non-managerial domains, as cited above. We show that limited memory also affects managerial decisions and prevents some firms from adopting a profitable opportunity. Overconfidence also affects decision-making in a number of domains, including managerial decisions about firm entry, corporate investment, acquisitions, and filing for bankruptcy (Camerer and Lovallo, 1999; Malmendier and Tate, 2005, 2008; Bernstein et al., 2024). Overconfidence about memory is less-studied, even in non-managerial settings, but has been shown to reduce take-up of delayed payments and mail-in rebates (Ericson, 2011; Tasoff and Letzler, 2014).

A novel contribution of our paper is that we identify trust in other firms as an important friction in managerial decision-making. Prior literature has documented that a lack of trust can have significant effects on consumer decision-making. For example, distrust in banks leads individuals to avoid using banks (Guiso et al., 2004), and interventions that increase trust can lead to increased savings (Bachas et al., 2021). Distrust also reduces stock-market participation, savings, borrowing, mortgage refinancing, risk pooling, and the take-up of insurance products (Guiso et al., 2008; Osili and Paulson, 2008; D'Acunto et al., 2019; Karlan et al., 2009; Johnson et al., 2019; Feigenberg et al., 2013; Cole et al., 2013).

Evidence on the role of trust in *interfirm* relationships is limited. McMillan and Woodruff (1999) find that suppliers in Vietnam are more likely to offer trade credit to buyer firms that they trust. Banerjee and Duflo (2000) document the importance of interfirm trust and reputation in the Indian software industry. Cai and Szeidl (2018) find that a lack of trust is a barrier to creating business partnerships in China, and randomizing regular meetings between firms increases trust. In Alfaro-Ureña et al. (2022), local supplier firms in Costa Rica cite gaining the trust of multinational corporations as an import precursor to exporting. In the US, a reduction of interfirm trust following the Enron scandal led to an increase in the number of contingencies included in contracts (D'Acunto et al., 2024). Within multinational firms, higher trust increases decentralization and raises aggregate productivity (Bloom et al., 2012). The findings of our paper suggest a mechanism to increase trust: We find that when a firm informs other firms that they will take an action and then follows through on that action, this increases trust and adoption of a profitable opportunity.

Turning to the literature on behavioral biases within firms, DellaVigna and Gentzkow (2019) identify managerial inertia as a key friction. They define managerial inertia as “agency frictions and behavioral factors that prevent firms from implementing optimal policies even though the

benefits of doing so exceed the economic costs.” However, there is limited evidence on which behavioral factors are driving this inertia within firms. Kremer et al. (2013) argue that loss aversion prevents small retail firms from stocking sufficient inventory. Beaman et al. (2014) find that limited attention prevents them from keeping sufficient small change. Selective attention appears to lead both seaweed farmers and multi-billion dollar companies to fail to attend to important features of the data (Hanna et al., 2014; List et al., 2023). In all of these cases, these behavioral factors led to lower profits.

There is also substantial evidence cited above on *other barriers* that firms face; see Verhoogen (2023) for an extensive survey on firm technology and product upgrading in developing countries, as well as the barriers that prevent firms from adopting these opportunities. Our contribution is to test whether—in addition to these barriers documented by other studies—present bias, limited memory (including distorted beliefs about memory), and a lack of trust in other firms prevent firms from adopting profitable opportunities.

## 2 Model

We present a simple theoretical framework to illustrate the potential roles of present bias, limited memory, and distrust in the take-up of business opportunities, and generate predictions for our empirical implementation. We build on the model in Ericson (2017), which allows for present bias, limited memory, and naïveté (overconfidence) about present bias and memory. We augment the model by introducing a role for trust: agents might discount the value of an opportunity if they do not fully trust their business partner to follow through.

### 2.1 Model Assumptions

A manager decides whether to take up a business opportunity that increases future profits but has an immediate cost of adoption. We consider the choice of adoption over  $T$  periods, from  $t = 1$  (when the offer is received) to  $t = T$ , including the possibility  $T \rightarrow \infty$ . The manager at time  $t$  maximizes the net present value of firm profits,  $U_t = \pi_t + \beta \left( \sum_{k=1}^{\infty} \delta^k \pi_{t+k} \right)$ , where  $\delta$  is the standard exponential discount factor and  $\beta \leq 1$  captures the possibility of present bias. The manager has beliefs  $\hat{\beta} \in [\beta, 1]$ , and is (partially) naïve if  $\hat{\beta} > \beta$ .

Managers may have imperfect memory. The parameter  $\rho_t$  measures the probability of remembering the offer in period  $t$  conditional on having remembered it in period  $t - 1$ . On the day the offer is received ( $t = 1$ ), we assume  $\rho_1 = 1$ . Managers have beliefs  $\hat{\rho}_t \in [0, 1]$ , and are overconfident about their memory if  $\hat{\rho}_t > \rho_t$ .

Adopting the offer in period  $\tau$  has an immediate cost  $c_\tau$  drawn from distribution  $F(c)$  that is continuous, differentiable, and has positive density over a range  $[\underline{c}, \bar{c}]$ . It generates a flow of future

benefits  $\{y_t\}_{t=\tau+1}^{\infty}$  starting in  $\tau + 1$ , which is the increase in profits each period after adopting. Thus for a manager adopting at  $\tau$ , the present value of future benefits as of  $\tau + 1$  is  $\sum_{k=0}^{\infty} \delta^k y_{\tau+1+k} \equiv y$ , and it is  $\beta \delta y$  as of  $\tau$ .

We incorporate trust into the model with a trust parameter  $\alpha_t$  such that the expected benefit from adopting the offer is  $\alpha_t y$ . Thus,  $\alpha_t$  can be thought of as the probability that the offer is not a scam, or the probability that the FinTech company is not trying to take advantage of the firm in some way. The subscript  $t$  allows trust to change over time.

Hence, the manager decides whether to adopt based on the value function

$$V_t = \begin{cases} \beta \delta \alpha_t y - c_t & \text{if adopting at } t, \\ \hat{\rho}_{t+1} \beta \delta \mathbb{E}_t[\hat{V}_{t+1}^t] & \text{otherwise,} \end{cases} \quad (1)$$

where  $\mathbb{E}_t[\hat{V}_{t+1}^t]$  is the expected value (over cost draws) of the perceived continuation value as of time  $t$  of not adopting at  $t$  (but potentially adopting at  $t + 1$  or later), with

$$\hat{V}_{t+1}^t = \begin{cases} \delta \alpha_{t+1} y - c_{t+1} & \text{if } \hat{\beta} \delta \alpha_{t+1} y - c_{t+1} \geq \hat{\rho}_{t+2} \hat{\beta} \delta \mathbb{E}_{t+1}[\hat{V}_{t+2}^{t+1}], \\ \hat{\rho}_{t+2} \delta \mathbb{E}_{t+1}[\hat{V}_{t+2}^t] & \text{otherwise.} \end{cases} \quad (2)$$

Note that, as of time  $t$ , a naïve present-biased manager *perceives* that the decision about adopting at  $t + 1$  will depend on  $\hat{\beta}$  rather than  $\beta$ . We denote this perception with a hat over  $V$ . The superscript and subscript on  $\hat{V}_{t+1}^t$  denote what the manager believes as of time  $t$  to receive for taking the action they perceive they will take at time  $t + 1$ . Since the manager evaluates  $\hat{V}_{t+1}^t$  at time  $t$  and all future benefits are already discounted by  $\beta$  in equation (1),  $\beta$  and  $\hat{\beta}$  do not enter the values themselves, but rather  $\hat{\beta}$  affects whether the manager thinks they will adopt at  $t + 1$ . If there is a deadline, the continuation value at the deadline is zero, as is the expected value of the perceived continuation value:  $\mathbb{E}_T[\hat{V}_{T+1}^T] = 0$  for each  $t \leq T$ .

## 2.2 Equilibrium Behavior

We impose the equilibrium notion of a perception-perfect equilibrium (O'Donoghue and Rabin, 1999), which requires that the action taken in each period maximizes preferences as of that period, given dynamically consistent beliefs about future strategies. We show in Appendix A that the model implies an optimal cutoff strategy where the manager adopts in period  $t$  if the offer is active and the cost draw  $c_t$  is below a threshold  $c_t^*$ . The offer is active if the manager has neither adopted nor forgotten about it by period  $t$ , and the deadline has not passed ( $t \leq T$ ). The optimal cost threshold is defined recursively:

$$c_t^* = \beta \delta (\alpha_t y - \hat{\rho}_{t+1} \mathbb{E}_t[\hat{V}_{t+1}^t]), \quad (3)$$

$$\mathbb{E}_t[\hat{V}_{t+1}^t] = \int_{\underline{c}}^{\hat{c}_{t+1}^*} (\delta \alpha_{t+1} y - c) dF(c) + (1 - F(\hat{c}_{t+1}^*)) \delta \hat{\rho}_{t+2} \mathbb{E}_{t+1}[\hat{V}_{t+2}^t], \quad (4)$$

where  $\hat{c}_{t+1}^*$  is what the manager at time  $t$  believes their optimal cost threshold will be at time  $t + 1$ . The definition of  $\hat{c}_{t+1}^*$  is thus identical to that of  $c_{t+1}^*$  but replacing  $\beta$  with  $\hat{\beta}$ . Appendix A proves the existence and uniqueness of this equilibrium.

### 2.3 Model Predictions

The model generates several testable predictions, which we will take to the data. In all predictions, “take-up” refers to cumulative take-up, and in some predictions we specify (cumulative) take-up by a specific date (e.g., take-up at  $t = 1$ , or take-up by the deadline). We use “final take-up” for predictions about overall cumulative take-up that do not necessarily hold for cumulative take-up in earlier periods. All proofs are in Appendix A.<sup>8</sup>

First, as a check of the basic economics underlying our experiment, we relate the value of the offer to the client firms’ take-up:

**Prediction 1** (Offer Value and Take-Up). *A higher value of the offer increases take-up.*

This prediction holds regardless of behavioral frictions being at work. The remaining predictions allow us to test for the role of these frictions, starting from the role of memory.

**Prediction 2** (Reminder and Memory). *If a reminder (unannounced or announced) increases the probability of remembering, it will also increase take-up.*

While Prediction 2 holds for both announced and unannounced reminders, they differ in their effect on take-up. We focus on the scenario where firms are forgetful and at least some firms do not adopt on the first day, e.g., because they wait for a better cost draw.

**Prediction 3** (Announced Reminder and Beliefs about Memory). *The announced reminder (a) reduces take-up at  $t = 1$ , compared to the unannounced reminder, if managers do not believe they have perfect memory, and (b) has no differential effect on take-up at  $t = 1$  if managers believe they have perfect memory.*

The reason for the predicted first-day effects is that the announced reminder increases the manager’s belief about their future ability to remember and adopt. When managers expect a reminder, they do not have to worry about forgetting. Thus the announced reminder lowers take-up on day 1 if and only if managers think they have imperfect memory.

---

<sup>8</sup> Note that most predictions—as well as the proof that there exists a unique equilibrium—are conditional on  $\hat{\beta}$  being not too small. Our simulations in Appendix B confirm that this restriction typically rules out only unreasonably small values of  $\hat{\beta}$ , i.e., beliefs that one’s future self will be severely present-biased.

**Prediction 4** (Announced Reminder and Trust). *The announced reminder (a) does not affect final take-up, compared to the unannounced reminder, if managers inherently trust the offer; and (b) increases final take-up if some managers distrust the offer and their trust increases after receiving the announced reminder.*

Prediction 4 describes the trust-based (rather than memory-based) implications of an announced reminder. On the day the announced reminder is received, the manager's belief in the trustworthiness of the FinTech firm and its offer may increase, and thus the expected value of adopting also increases.

We next turn to predictions about the effect of a deadline.<sup>9</sup>

**Prediction 5** (Deadline and Offer Value). *A higher value of the offer implies a lower treatment effect of a deadline on take-up by the deadline.*

Prediction 5 implies that if we increase the value of the offer (as we do experimentally by randomizing the fee reduction we offer), the effect of the deadline should decrease.

**Prediction 6** (Deadline and Take-Up). *The treatment effect of a deadline on take-up by any date up until the deadline is always positive if the deadline does not affect trust, memory, or perceived memory.*

Prediction 6 implies that if the deadline has a *negative* treatment effect on cumulative take-up on any day prior to the deadline, the deadline must affect trust, memory, or beliefs about memory.

**Prediction 7** (Deadline and Present Bias). *The more present-biased a manager is, the larger is the treatment effect of a deadline on take-up by the deadline.*

The contrapositive of Prediction 7 implies limited or no present bias if a deadline does not have an effect on take-up by the deadline.

### **3 Institutional Context and Experimental Setting**

We partnered with a FinTech payments company in Mexico to study whether these behavioral frictions—present bias, limited memory, and a lack of trust in other firms, as well as possibly distorted beliefs—help explain firms' failure to adopt a profitable opportunity. The FinTech company provides its customers with point-of-sale (POS) hardware and an app to accept debit and credit card payments, similar to Square in the US. For each electronic card payment that a client firm processes, the FinTech company charges a merchant fee that is a percentage of the payment amount.

---

<sup>9</sup> Predictions 5 and 7 impose additional parameter restrictions, formalized in Appendix A. The simulations in Appendix B show that these predictions nevertheless tend to hold for reasonable parameter values.

The FinTech partner charges 199 pesos (\$9 USD) to purchase the POS terminal, no monthly fee, and a 3.5–3.75% merchant fee.

FinTech and cashless payments have expanded substantially in Mexico since the 2010s. A government program that rolled out debit cards to poor households in 2009–2012 led to an increase in small retail firms’ adoption of POS terminals (Higgins, 2024). Mexico passed a FinTech Law in 2018 and by 2020 had 441 FinTech startups and over 50 million users of FinTech payment products.<sup>10</sup> FinTech companies offering POS terminals entered the market in 2013, and our FinTech partner estimates that in 2019, of the 3.3 million total POS terminals in use in Mexico, 1.3 million were issued by FinTech payments companies, while another 1.3 million were issued by banks and 0.6 million were issued by other issuers. Our partner had a 10% market share among FinTech-issued POS terminals, and a 4% market share among all POS terminals. Among small firms, FinTech-issued POS terminals are substantially more popular than bank-issued POS terminals: according to survey data from 2022 collected by Higgins (2024), 61.1% of small retail firms with a POS terminal had a FinTech-issued POS terminal, while 21.7% had a bank-issued POS terminal and the remainder had a POS from other issuers. In focus groups we conducted with small firms using our FinTech partner’s product prior to the RCT, most managers stated that prior to adopting the FinTech company’s technology they did not accept card payments.

The FinTech company’s motivation for partnering with us for this experiment was two-fold. First, they were interested in testing a lower fee to increase customer retention (i.e., to lose fewer customers to competitor FinTech companies or banks). They also wanted to see what messages would increase adoption of this lower fee (and hence potentially further increase retention). Second, they did not know the elasticity of customer firms’ card sales with respect to the fee, and thus whether they were charging the optimal fee.

## **4 Experimental Design**

### **4.1 Intervention**

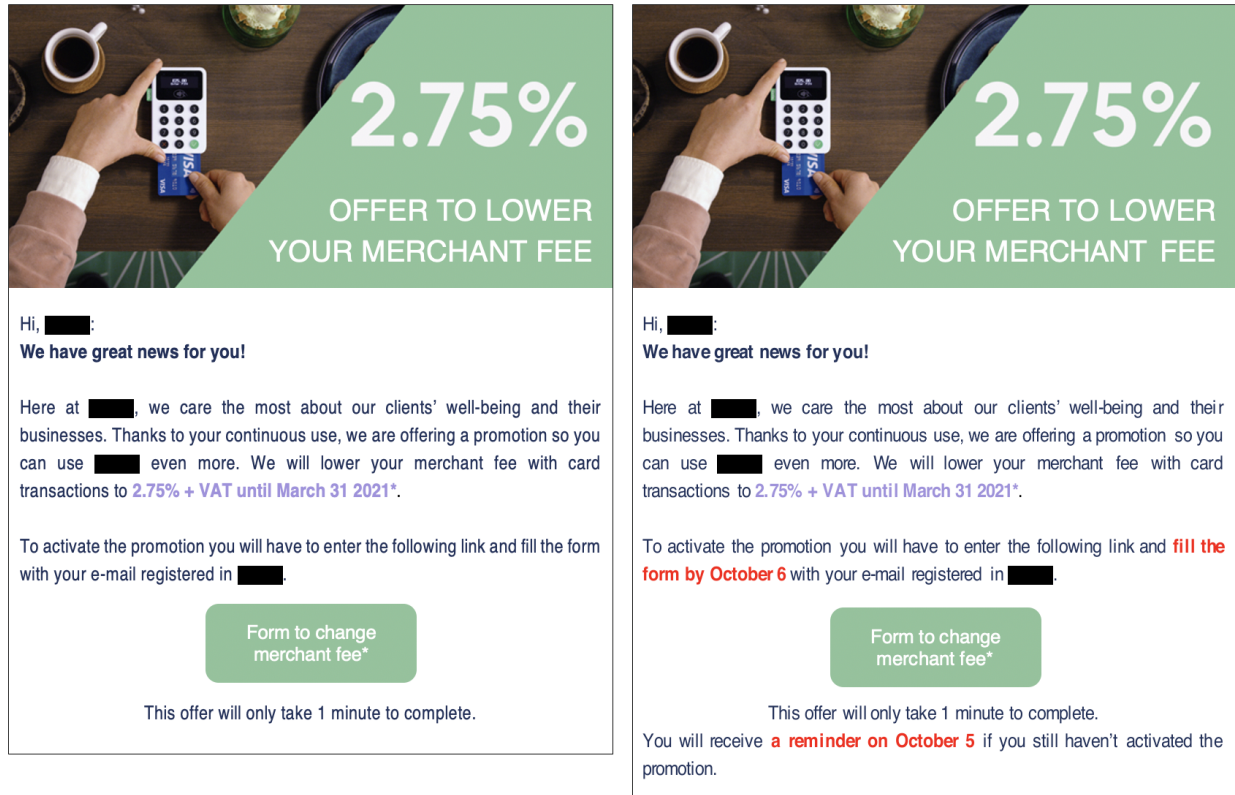
Our FinTech partner randomly offered the opportunity to obtain a lower merchant fee to firms that were already users of their technology. Managers were informed through both email and SMS text messages in order to maximize awareness of the offer. Figure 1 shows two sample emails. Each email had a button that linked to a short online form that managers had to fill out to activate the fee reduction, which was generally activated within one day. The form required managers to fill in basic information they had previously shared with the FinTech firm: name, email, and national

---

<sup>10</sup> See <https://www.finnovista.com/en/radar/fintech-radar-mexico-23-eng/> and <https://www.debate.com.mx/economia/Pagos-digitales-en-Mexico-estadisticas-y-tendencias-del-mercado-20230429-0095.html>.

identification number (which is frequently used in Mexico for many types of transactions). The email informed the user that, based on the FinTech company’s best estimate, the form would only take one minute to complete.

Figure 1: Sample Emails with Lower Rate Offers



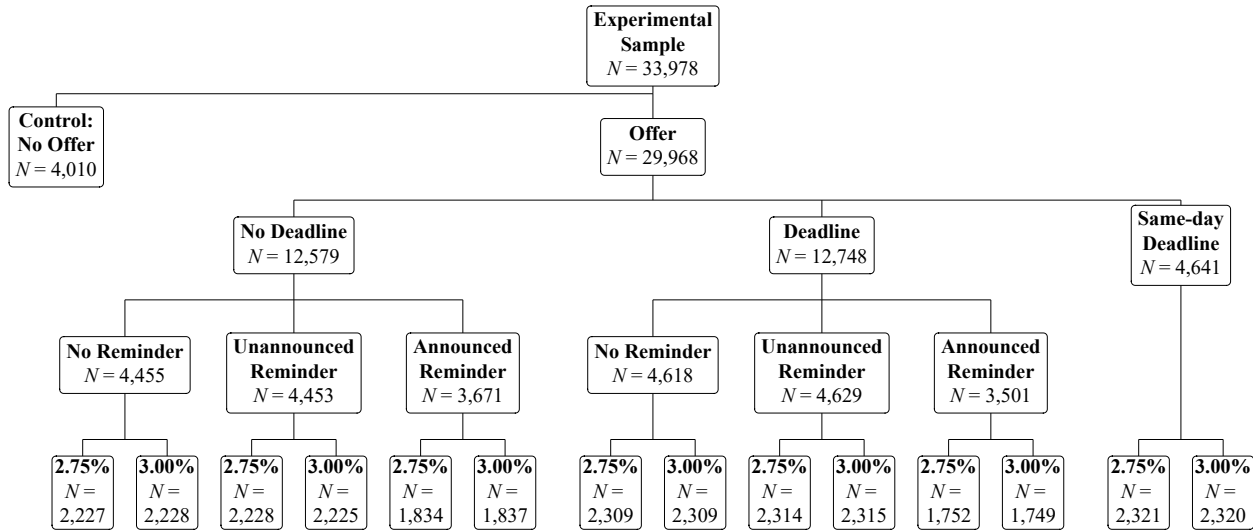
This figure shows screenshots of the emails sent to different treatment arms, on the left to the arm with no deadline and either no or an unannounced reminder, and on the right to the deadline, announced-reminder arm. Emails for other treatment arms were designed accordingly: For example, the no-deadline, announced-reminder arm would exclude the “fill the form by October 6” sentence but include “reminder on October 5” from the right panel. The text is translated from the original Spanish into English. Asterisks at the end of the purple text refer to the fine print at the bottom of the email, which reads: “By filling out the form you authorize [redacted] to change the fee on your [redacted] account to a 2.75% + VAT fee per successful card payment transaction until March 31, 2021. Starting April 1, 2021, the fee will revert back to the fee you had before activating this promotion.”

Figure 2 illustrates the various treatment arms in the experimental design. Among the 33,978 firms in the RCT, 4,010 firms were randomly assigned to the control group that was eligible to receive the offer based on our sample selection criteria, but did not receive it. The control group size was based on institutional constraints of the FinTech partner. The remaining firms were assigned to one of the seven other groups combining deadlines and reminders: (i) no deadline, no reminder (4,455 firms); (ii) no deadline, announced reminder (3,671); (iii) no deadline, unannounced reminder (4,453); (iv) eight-day deadline, no reminder (4,618); (v) eight-day deadline, announced reminder (3,501); (vi) eight-day deadline, unannounced reminder (4,629); and (vii)



same-day deadline, no reminder (4,641).<sup>11</sup> The sample sizes were determined based on power calculations using the results from our May 2019 randomized pilot and simulations of the model presented in Section 2.

Figure 2: Experimental Design



This figure shows how many firms were randomly assigned to each treatment arm.

Within each of these treatment groups, we also experimentally varied the value of the offer, 3% versus 2.75%. Prior to the experiment, firms were charged either 3.75% or 3.5%, which was a function of when they started using our FinTech partner’s technology. Thus, the fee reduction ranges from 50 basis points—for those reduced from 3.5% to 3%—to 100 basis points—for those reduced from 3.75% to 2.75%. (Part of the size of this reduction is random based on their randomized new fee offer, and part is not random based on whether they had a 3.75% or 3.5% fee before the experiment.) The lower fee lasted for six months (until March 31, 2021), after which the firm’s rate returned to their pre-intervention rate. All of this information was included in the emails they received. The reason that the fee reduction was temporary was that our FinTech partner worried that firms’ use of the technology might be inelastic with respect to the lower fee, in which case the FinTech company could lose a substantial amount of money by lowering the fee permanently.

We stratified our randomization by business type and quartiles of average monthly sales before our intervention, where business type is one of six categories: small retailers, professionals, beauty, clothing, restaurants, and other.

## 4.2 Timeline

The timeline of the experiment is as follows. The control group did not receive any offer or messages related to the experiment. The treatment groups received an initial email and SMS message

<sup>11</sup> The same-day deadline naturally precluded a reminder, which was sent on day 7 in the other arms.

with the offer on September 29, 2020 at 10 am Central Standard Time (CST), which is the time zone that covers most of Mexico. We did not randomize the method of contact (email vs. SMS) as we wanted to maximize awareness of the offer.

The no-deadline, no-reminder group received no subsequent messages; this group could adopt any time (Figure 1, left panel). The no-deadline, unannounced-reminder group received an initial email identical to that of the no-deadline, no-reminder group, but also an unannounced reminder email on October 5 at 10 am CST. The no-deadline, announced-reminder group received the initial email from September 29 with an additional sentence stating that they would receive a reminder on October 5; this group then received a reminder email identical to that of the no-deadline, unannounced-reminder group on October 5 if they had not already adopted. The deadline, no-reminder group received an initial email on September 29 with an additional sentence informing them of the deadline on October 6 (at midnight). The deadline, unannounced reminder group received the same initial email on September 29 as the deadline, no-reminder group informing them of the October 6 deadline, but also a reminder on October 5. The deadline, announced reminder group received an initial email on September 29 that informed them of the deadline on October 6 and that they would receive a reminder on October 5 (Figure 1, right panel). Finally, we had a treatment arm with a same-day deadline and no reminder, which was informed in the initial email that the deadline was “today, September 29” at midnight. The purpose of the same-day deadline arm was to isolate variation in costs from the probability of forgetting.

The experiment was initially set to launch on March 24, 2020, but was delayed due to the start of the COVID-19 pandemic. Since we could observe the electronic sales of our potential sample in administrative data, we waited until average monthly sales recovered to pre-pandemic levels before launching the experiment; this occurred in August 2020.

### **4.3 Sample**

To maximize the absolute value of the offer, we selected the RCT sample to include the top quartile of the FinTech company’s approximately 130,000 users, as measured by their monthly sales in August 2020 (the month before our experiment launched). By using August 2020 sales, our filtering excluded firms that had closed or greatly reduced their sales due to COVID-19. This sample selection was informed by a randomized pilot we conducted in May 2019 with 11,755 firms throughout the sales distribution. We found that the take-up rate of the lower fee was increasing in baseline sales and that the elasticity of card sales with respect to the fee was statistically significant only for the fourth quartile.

After filtering out firms included in our randomized pilot, we identified the top quartile of users (34,010 firms) for the RCT. Our FinTech partner then filtered out users that were not in good standing administratively, which resulted in a sample of 33,978 firms. As expected by virtue of

randomization, the probability of being excluded for not being in good standing does not differ by treatment arm (Appendix-Table C.1).

## 5 Data

We use two main sources of data: administrative data on the 33,978 firms provided by our FinTech partner and survey data that we collected on a subsample of firms in the RCT.

### 5.1 Administrative Data

We have pre-experimental data on characteristics (such as managers' sex and age, business type, dates of registration and first transaction), and firm  $\times$  day level data on the number and peso amount of transactions via the FinTech payments technology starting in July 2019.

Appendix-Table C.2 shows baseline sample statistics. It also shows that the randomization is balanced across treatments; the numbers in each row of the table come from a regression of each firm characteristic on a set of treatment dummies: unannounced reminder, announced reminder, deadline, and 2.75% fee. Column (1) shows the intercept (and thus the control group mean), while columns (2)–(5) show the coefficients on the treatment arm dummies. Column (6) shows the omnibus  $F$ -statistic and corresponding  $p$ -value for the regression in that row. All of the variables are balanced across treatment arms.

As shown in Panel A of Appendix-Table C.2, the managers in our RCT are 44.1% female and on average 39 years old. Panel B shows that the most common business types are *small retailers* (corner stores and prepared food vendors; 26%) and *professionals* (medical services, dentists, and veterinarians, 23.9%). For the rest of the business types, the *beauty* category includes hair dressers, barber shops, beauty salons, and spas, *clothing* includes clothing, shoe, and accessory stores, *restaurants* are restaurants, cafes, and bars, and *other* includes auto shops, construction material wholesalers, and other business types. The average length of time that firms have been using the FinTech technology is 24 months.

We have experimental data on whether and when the firm (i) opened the email, (ii) clicked on the link in the email, (iii) filled out the form to activate the lower fee, and (iv) logged into their online account, from the first day of the experiment (September 29, 2020) through the final day that the lower fee was valid (March 31, 2021). We also have firm  $\times$  day level data on the number and peso volume of transactions on the FinTech payments technology through the end of the experiment in March 2021.

In the group with no interventions designed to increase take-up (i.e., the no-deadline, no-reminder group), overall take-up by day 8 was 25.4%. Take-up is slightly increasing in firm size: the take-up rate was 23.1% in the smallest quintile of firms (measured by baseline sales through

the technology), and 26.8% in the largest quintile.<sup>12</sup>

While these levels of take-up may appear surprisingly low for an opportunity that was worth 3% of profits for the median firm, they are substantially higher than for other (less-valuable) campaigns of our FinTech partner (1–2%) and are comparable to estimates from studies that use letters to offer high-value cost-saving opportunities to individuals.<sup>13</sup>

## 5.2 Survey Data

We conducted a survey on a subset of firms. The number of surveys was constrained by our FinTech partner (471 firms), with a response rate of 33.7%. Since one goal of the survey was to provide evidence on the mechanism explaining the larger treatment effect of the announced over the unannounced reminder, we oversampled firms in these two treatment arms, and also oversampled firms that accepted the offer. Appendix-Table C.3 shows that the survey subsample is balanced across treatments, and Appendix-Table C.4 shows that it is comparable to the non-surveyed sample on observables.

Our sample turns out to be fairly similar to the full set of all Mexican firms in terms of size distribution. Appendix-Figure C.1 shows the distribution of employees per firm, both in our survey sample and in the 2019 INEGI Economic Census data of all Mexican firms. The median number of employees in our sample is 3, and the average is 3.9. We see that the distributions look very similar except for the very largest firms in Mexico. In our survey, 87% of firms have one to five employees, compared to 90% of all firms in Mexico. Furthermore, the largest firm in our survey has 150 employees, which corresponds to the 99.7th percentile of the distribution of number of employees across all firms in Mexico.

In addition to eliciting the number of employees, the survey asks about profits, share of sales through the technology, how accurately managers know the fees they were charged and the value of transactions they made through the technology in the last week, and how much they expected to save by accepting the lower fee. The latter question enables us to measure whether knowledge about the size of the fee reduction translates into knowledge about how much they would save, or alternatively if they had trouble translating the fee reduction into a cost savings by multiplying it by their sales (Shue and Townsend, 2021).

The survey also includes questions about whether managers remember receiving the email and SMS, whether they read the SMS, whether they noticed the offer had a deadline (for those assigned a deadline), and what impact the lower fee had on their business. Depending on whether and

---

<sup>12</sup> The difference between the first and fifth quintiles is statistically significant at the 10% level ( $p = 0.063$ ).

<sup>13</sup> Take-up rates of high-value cost-savings opportunities include 20% mortgage refinancing (saving over \$1,200 per year) in response to a letter from the financial institution (Johnson et al., 2019) and 28% take-up of a Medicare Part D plan that covered more of the drugs an individual was prescribed (saving about \$100 per year), also in response to a letter (Kling et al., 2012).

when the manager adopted, we also ask questions about why they adopted on the first day, waited and adopted on a later day, or did not adopt. Finally, the survey includes general social survey (GSS) questions related to trust, procrastination, memory, and other biases, which we use for heterogeneous treatment effects analysis to understand mechanisms. On trust, survey respondents describe how much they agree with the statement “I trust advertised offers” on a 1 to 5 scale ranging from strongly disagree to strongly agree. (We describe the GSS questions in more detail in Section 7.)

We also asked managers how long they expected it to take and how long it actually took to fill out the form to activate the offer. Most firms report that they estimated it to take between six and ten minutes, and that it took between one and five minutes (Appendix-Figure C.2).<sup>14</sup>

Finally, we asked managers who did not adopt the profitable opportunity why they did not adopt. The most common answers (see Appendix-Figure C.3, also for the wording of the questions) fall into four groups. A first group of managers “ran out of time,” which would be consistent with procrastinating due to present bias. A second group “forgot,” consistent with limited memory. Managers who provided these two most-common responses may have been very busy (as captured by a high cost draw in our model), both when they got the initial email and, if applicable, when they got the reminder, and this high cost draw combined with their present bias or limited memory led them to never adopt.<sup>15</sup> A third group thought it would take too much time, did not consider it important, or were not sure if it would benefit them. Although expected cost savings from the lower fee are equal to 3% of median firm profits, there is heterogeneity in margins and sales transacted through the FinTech technology. Thus, while it is a profitable opportunity for a substantial fraction of firms in the experiment, some firms might still not consider it worth the perceived effort cost. Fourth, some managers reported that they did not trust the offer. We will analyze the roles of present bias, limited memory, and distrust in detail below.

## 6 Results

Our primary results use the following regression specification, estimated separately for each day of the experiment from  $t = 1$  to  $t = 8$ :

$$y_i^t = \lambda_{s(i)} + \sum_{k=2}^K \beta_k T_i^k + \varepsilon_i, \quad (5)$$

---

<sup>14</sup> It is possible that the expected time reported by managers included the time to perform other actions before adopting the lower fee such as reading the fine print, calculating how valuable the offer was, or discussing with someone else at the firm whether to accept the offer.

<sup>15</sup> The first bar in Appendix-Figure C.3, “opened email but did not remember doing so,” is the number of people who we observe opened the email but who did not recall opening the email when we surveyed them, so we could not ask them why they did not accept the offer.

where  $y_i^t$  measures cumulative take-up, i.e., is equal to 1 if firm  $i$  accepted the offer on or before day  $t$ . The specification includes randomization strata fixed effects  $\lambda_{s(i)}$  (which absorb the constant). The indicator  $T_i^k$  denotes assignment of firm  $i$  to treatment arm  $k$  (where the omitted category  $k = 1$  corresponds to the control group), and  $\varepsilon_i$  are heteroskedasticity-robust standard errors (not clustered since the randomization unit is the individual firm).

The model parameters map to our empirical setting as follows: A time period  $t$  corresponds to a day. The cost draw  $c_t$  reflects how a busy manager is on day  $t$ , as it takes a bit of time to click the link in the email or SMS message and to fill out the short form. (The manager may also want to check for fine print, calculate how valuable the offer is, or discuss with someone else at the firm before accepting the offer.) In fact, when asked why they adopted on the first day or why they delayed, 75.5% of managers who adopted on the first day report doing so because they had time that day, and 72.6% of managers who delayed adopting say they were too busy on the day they received the email (Appendix-Figure C.4).

The benefit  $y$  is the present value of cost savings from the lower merchant fee. As the lower fee was activated one day after a firm accepted the offer (as also indicated in the online form), the assumption that the cost to adopt  $c_t$  is borne in the current period but the benefit  $y$  is experienced with a one-period delay maps to our empirical setting.

For firms with a deadline,  $T = 8$ , except for the same-day deadline arm where  $T = 1$ ; and  $T \rightarrow \infty$  for those without a deadline.<sup>16</sup> We compare cumulative take-up on each of the eight days prior to the deadline, and in some figures also compare take-up for six months after the RCT (since treatment groups without deadline could adopt after day 8).

A reminder raises the probability of remembering,  $\rho_t$ , in the period it is sent. However, only a pre-announcement of receiving a reminder in a future period  $\tau$  increases the manager's current beliefs about the probability of remembering in that future period,  $\hat{\rho}_\tau$ .

Finally, trust in the offer  $\alpha_t$  could potentially be affected by any of the treatments. We are particularly interested in two cases. The first is the case where the manager might trust the offer more if they are told they will receive a reminder and then receive this reminder. The subscript  $t$  on  $\alpha_t$  allows trust to increase either upon receiving the initial pre-announcement, or only upon receiving the reminder that had previously been announced. The second is the case where the deadline might decrease trust in the offer.

We next turn to estimating the treatment effects of each of our treatments. We restate each theoretical prediction in terms of its mapping into the empirical setting.

---

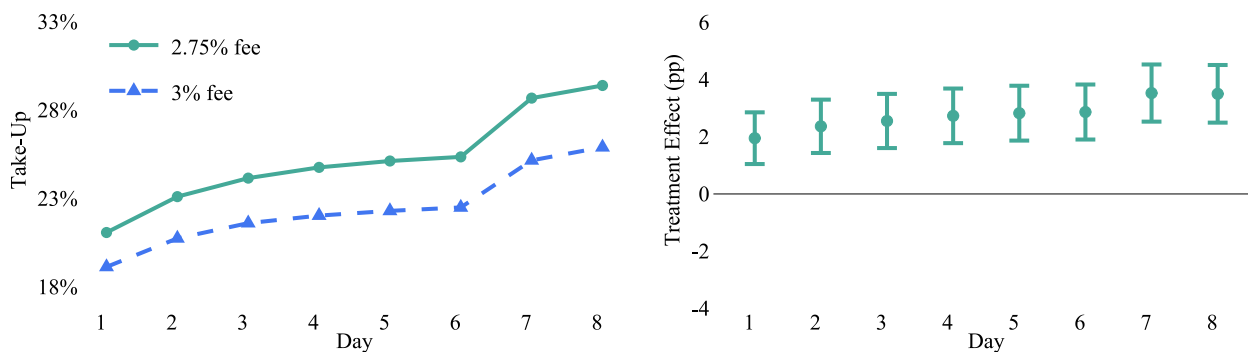
<sup>16</sup> For treatment arms without a deadline,  $T = 184$  is more accurate, which is the number of days between the initial offer and the fixed end date of the six-month lower fee. In Appendix A, we formally show that we can approximate the no-deadline case by a sufficiently long but finite  $T$ .

## 6.1 Offer Value and Take-Up

Prediction 1 states that take-up should be higher when the offer is more valuable. To evaluate this prediction, we exploit the random variation we introduced by offering either a 2.75% or 3% fee. Since the 2.75% fee reduces costs more, it should have a higher take-up.

Figure 3 shows cumulative take-up rates (left panel) and the treatment effects of the 2.75% offer on cumulative take-up and their 95% confidence intervals (right panel) each day of the experiment.<sup>17</sup> As predicted, take-up of the more profitable 2.75% offer is higher. On the first day of the experiment, it was 2 pp higher compared to a base of 19.1% take-up in the 3% offer group. By day 8, it was 3.5 pp higher compared to a base of 25.9% take-up in the 3% offer group. In relative terms, the 2.75% fee increased take-up by day 8 by 13.5%. These effects are statistically significant at the 1% level.

Figure 3: Effect of Offer Value on Take-up



This figure shows take-up rates by offer value and treatment effects of a more-valuable offer. The line graph on the left shows average cumulative take-up rates for the 2.75% and 3% fee groups. The coefficient graph on the right shows the treatment effects of a more-valuable offer (2.75% rather than 3%) on cumulative take-up in percentage points (pp), separately for each day of the experiment. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on an indicator of the fee group (2.75% rather than 3%), controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline), among 29,968 firms with 2.75% and 3% offers, excluding the pure control group.

We also note that the higher take-up of the 2.75% offer persists after the deadline. Appendix-Figure C.6a shows take-up over six months for firms that did not have a deadline. (Those with a deadline are excluded from the figure since they could not adopt anymore.) The gap between the 2.75% and 3% group persists and increases slightly to 4 pp. Our results confirm Prediction 1 and validate the basic incentive design of our experiment.

## 6.2 Reminder and Memory

We now turn to the role of reminders and memory. From Prediction 2, a reminder increases take-up if managers are forgetful and reminders increase the probability of remembering.

<sup>17</sup> Take-up rates for each treatment arm by the October 6 deadline are in Appendix-Figure C.5 and Appendix-Table C.5.

For this analysis, we pool all firms that received a reminder (unannounced or announced), since the prediction applies to both types of reminders. Figure 4a shows cumulative take-up rates for the reminder and no-reminder arms (left panel) and regression coefficients and 95% confidence intervals for the effect of the reminder (right panel) from day 1 to day 8. In both groups, take-up rates are close to 20% on day 1 and increase steadily to about 24% on day 6, the day before the reminder. There is no statistically significant difference on any day before the reminder was sent.<sup>18</sup> After the reminder was sent on day 7, the take-up rate in the reminder group was 4.1 pp higher than in the no-reminder group. On day 8, the difference in take-up is 4.7 pp. The effects of the reminder on cumulative take-up by days 7 and 8 are statistically significant at the 1% level. Appendix-Figure C.6b shows that—restricting to those without a deadline (who could still adopt after day 8)—the gap in take-up driven by the reminder persists for the six months after the experiment.<sup>19</sup>

**Lack of heterogeneity in reminder effect.** We fail to find any statistically significant heterogeneity of the reminder effect across any subgroups.

The effect of the reminder is similarly large in the 2.75% and 3% offer groups (Appendix-Figure C.8a): there is no statistically significant heterogeneous treatment effect of the reminder exploiting the random variation in how valuable the offer is. The same holds when we test for heterogeneity by the expected gain from adopting the offer, which is calculated as the change in the fee times the firm’s baseline sales. As expected, firms with an above-median expected gain have higher overall take-up, but the effect of the reminder is just as large for firms with above- and below-median expected gain (Appendix-Figure C.10a). Both of these heterogeneity tests were pre-specified. We also have a measure of the offer value from the survey, which is the percent of total sales that the firm transacts through the FinTech technology.<sup>20</sup> While there is substantial variation in this measure, we again do not find heterogeneous reminder effects (Appendix-Table C.7, column 3).

The effect of the reminder is also equally large for smaller and larger firms, measured by below- vs. above-median baseline sales (Appendix-Figure C.9a). We pre-specified the heterogeneity test using above- vs. below-median baseline sales as one of the main two variables we would test for heterogeneity on. The same holds when we use below- vs. above-median number of employees from our survey data as an alternative measure of size (Appendix-Table C.8, column 1). Note that the heterogeneity tests by number of employees use survey data since we do not observe number of

---

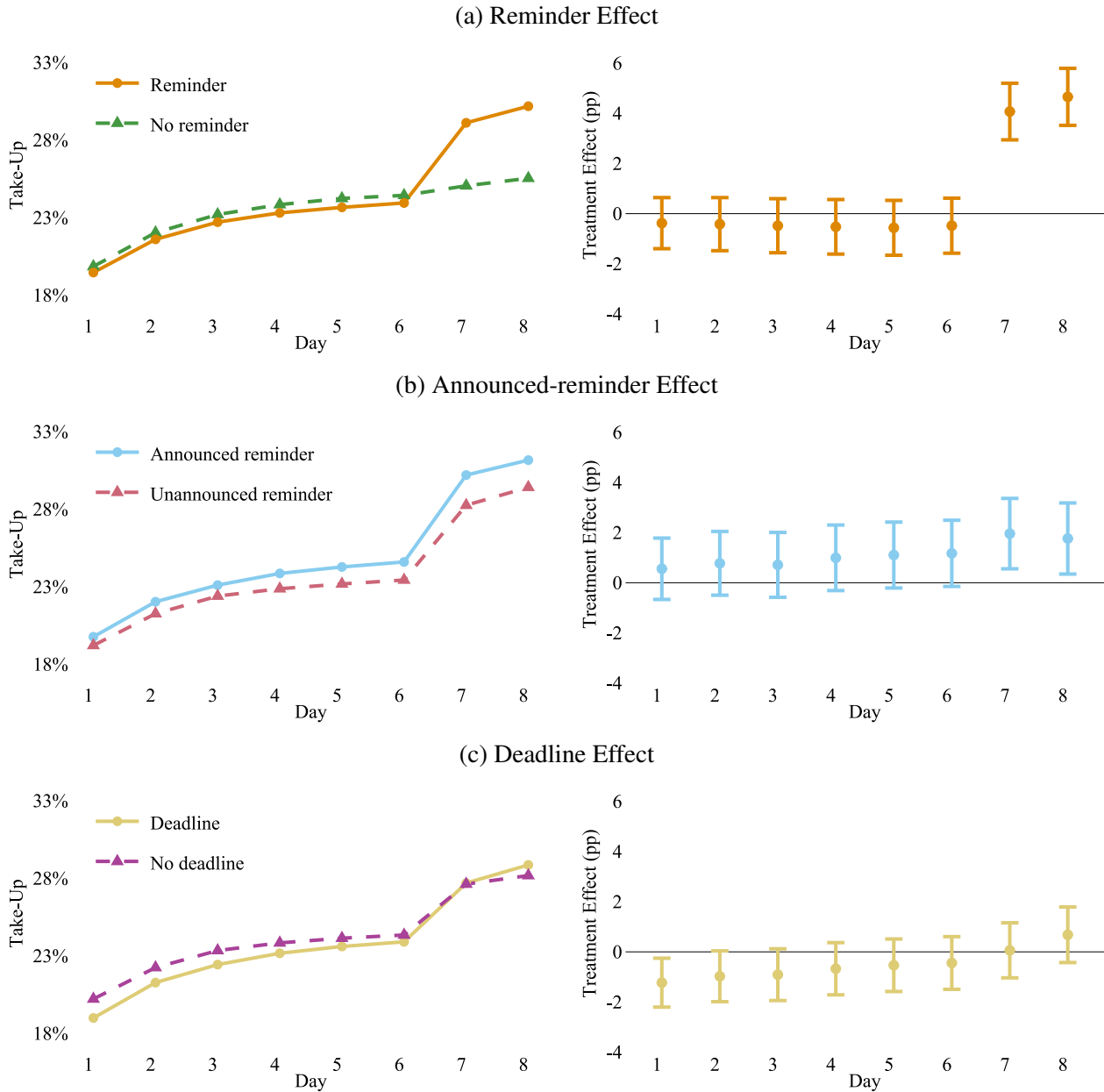
<sup>18</sup> The same holds when we pool across days 1–6 for increased power (Appendix-Table C.6).

<sup>19</sup> We also re-estimate the reminder effect restricting the sample to those receiving an unannounced reminder (as the announced reminder group received a different initial message) and again find no difference prior to the reminder and a large 3.2 pp effect on cumulative take-up after the unannounced reminder is sent on day 7 and a 3.9 pp effect on cumulative take-up by day 8 (Appendix-Figure C.7).

<sup>20</sup> No heterogeneity tests using survey measures were pre-specified, as we designed the survey after seeing the results on take-up over days 1–8 from the administrative data.



Figure 4: Effect of Reminder, Announced Reminder, and Deadline on Take-Up



This figure shows take-up rates by treatment arm (left panels) and treatment effects (right panels) of a reminder, an announced reminder, and a deadline. Line graphs on the left show average take-up rates by treatment group. Coefficient graphs on the right show the corresponding coefficient estimates for the differential take-up of the groups, separately for each day of the experiment. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on treatment, controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). Panel (a) includes 25,327 firms with reminders (both announced and unannounced) and no reminders, excluding the same-day deadline and pure control groups. Panel (b) includes 16,254 firms with announced and unannounced reminders, excluding firms without reminders. Panel (c) includes 25,327 firms with and without deadlines, excluding the same-day deadline and pure control groups.

employees in the administrative data. Because one-employee firms might be different than larger firms, we also test for heterogeneous treatment effects of the reminder for firms with one employee vs. more than one employee, and again do not find statistically significant heterogeneous treatment effects (Appendix-Table C.8, column 2).

The reminder effect also holds irrespective of (i) whether the manager who received the email was the owner of the firm or not (Appendix-Table C.7), (ii) firm type (Appendix-Table C.9), which was the other of the two main variables to test for heterogeneity on that we pre-specified, (iii) splits by the other manager and firm characteristics that we pre-specified, which are owner age, owner sex, and the month-over-month change in baseline sales from August to September 2020 (excluding the days of the experiment, September 29 and 30) to capture firm growth (Appendix-Table C.10), and (iv) whether or not the offer had a deadline (Appendix-Figure C.11). In sum, the reminder has a large effect on take-up across all of the types of firms in our experiment. These findings are consistent with a strong role of memory and forgetting, though they could also be attributed to alternative explanations outside of our model, which we explore next.

**Alternative explanations for reminder effect.** One possibility is that the reminder effect is not driven by managers knowing about the offer and then forgetting, but rather by them paying limited attention and thus not knowing about the offer. Indeed, limited or selective attention have been shown to lower profits in other contexts (Beaman et al., 2014; Hanna et al., 2014; List et al., 2023).

While we cannot rule out limited attention playing a role, we do observe whether managers opened the initial email prior to receiving the reminder. Given the design of the email with a large banner at the top showing the lower fee in large bold numbers and stating “offer to lower your merchant fee” (Figure 1), it is unlikely that managers opened the email without learning that it was an offer to lower their merchant fee. Overall, the rate of opening the initial email prior to receiving the reminder was 40.5%. This is substantially higher than the 23% open rate of marketing emails the FinTech company had sent to its users, and more than twice the 18% industry-wide open rate for retail. Appendix-Figure C.12a shows that conditional on opening the email prior to receiving the reminder, the reminder still has a 5.2 pp effect on take-up. We do not observe whether managers opened the SMS message, but we asked about this in our survey and 49.4% remembered receiving the SMS; of those, 89.7% reported reading the message.

Furthermore, when we asked managers why they did not accept the offer, 20.7% responded that they forgot (Appendix-Figure C.3). Another 23.4% had opened the email according to our administrative data but did not remember doing so, which could be either limited memory—forgetting not only to accept the offer but forgetting the offer even existed—or limited attention (e.g., opening the email but not reading it).

Taken together, these results suggest that limited attention cannot by itself explain the full effect

of the reminder. They do not rule out, though, that limited attention played a role for some firms and that the reminder could have also eased limited attention constraints.

We also asked firms directly why they thought the FinTech company sent them a reminder. The top three reasons that firms thought the FinTech sent them a reminder were to make sure they wouldn't forgo a valuable offer, that the FinTech company knew they would forget, or that it is a usual business practice (82.2% of firms answered one of these three responses). Only a small percentage (11.5%) were wary of the motives for being sent a reminder, answering that the reason was to increase the FinTech's profits or to make firms fall for a scam (Appendix-Figure C.13a).

### **6.3 Announced Reminder and Beliefs about Memory**

We next analyze the differential effects of the announced and the unannounced reminder. Prediction 3 specifies that the announcement reduces pre-reminder take-up if managers do not believe they have perfect memory, but has no differential effect on pre-reminder take-up if managers believe they have perfect memory. Figure 4b shows that there is not a *negative* effect of announcing the reminder on pre-reminder take-up; the point estimate is non-significant and *positive*. In the model, the lack of a negative pre-reminder effect of announcing the reminder suggests that managers are overconfident about their memory, assuming that the effect of the reminder itself indicates limited memory.

### **6.4 Announced Reminder and Trust**

From Prediction 4, an announced reminder will not increase final take-up compared to an unannounced reminder if firms inherently trust the offer, but will increase final take-up if some firms distrust the offer and their trust in the offer increases after receiving the reminder that had been pre-announced. As seen in Figure 4b, we find that take-up after the reminder is sent is nearly 1.8 pp higher in the announced reminder group than the unannounced reminder group (statistically significant at the 5% level). This gap in take-up between the announced and unannounced reminder groups persists over the six months after the reminder was sent (Appendix-Figure C.6c).

We note that the point estimates in Figure 4b are also positive prior to the reminder being sent, and we cannot reject that the treatment effect of the announced reminder relative to the unannounced reminder is larger on the day of the reminder (day 7) than the day before the reminder (day 6). Hence, there might be an effect of announcing the reminder when the reminder is first announced (in addition to the effect of the announced relative to unannounced reminder on the day the reminder is sent). However, when we pool across pre-reminder days and separately across post-reminder days for additional power, we *do* reject that the pre-reminder and post-reminder effects of the announced reminder are equal (Appendix-Table C.6). Specifically, the additional effect

of the announced reminder once the reminder is received is 1 pp (statistically significant at the 5% level).

**Lack of heterogeneous post-reminder announcement effects.** We do not find statistically significant heterogeneous treatment effects of the announced reminder on cumulative take-up by day 8 along any dimensions: the size of the fee reduction (Appendix-Figure C.8b), baseline sales (Appendix-Figure C.9b), the expected gain (Appendix-Figure C.10b), percent of total sales the firm transacts through the FinTech payments technology (Appendix-Table C.7), number of employees, whether the firm has more than one employee (Appendix-Table C.8), whether the owner was the recipient of the emails (Appendix-Table C.7), the business type (Appendix-Table C.9), the owner age, owner sex, business growth in the month prior to the experiment (Appendix-Table C.10), or whether there is a deadline (Appendix-Figure C.14c).<sup>21</sup>

In sum, the larger effect of the announced reminder relative to the unannounced reminder appears to hold regardless of which other treatments are included with it, and across all of the types of firms in our experiment.

The higher take-up in the announced reminder group is consistent with (some) managers not fully trusting the offer initially, but their trust increasing after the FinTech firm followed through with the reminder as it had promised in advance. As a result, the perceived value of the offer to these managers would increase and thus their take-up would increase, as outlined in Prediction 4.

We collected additional evidence to explore both the trust mechanism and potential alternative mechanisms, which we discuss in Section 7.

## 6.5 Deadline and Offer Value

Prediction 5 states that the deadline effect is decreasing in the offer value. We exploit again the random variation in the offer value as well as the randomization into the deadline treatment to test this prediction.

Appendix-Figure C.8c shows the day-by-day cumulative take-up rates among merchants in the deadline versus no-deadline groups (left panel) and the treatment effects of the deadline (right panel)—in both panels separately for merchants who received the more-valuable (2.75%) versus the less-valuable (3%) offer. There is no effect of the deadline on cumulative take-up by day 8 for the more-valuable offer (2.75% fee), but for the less-valuable offer (3% fee), the deadline increases take-up by 2 pp. The difference in the estimated treatment effect of a deadline for the 2.75% and 3% fee groups is statistically significant at the 10% level ( $p = 0.083$ ).

---

<sup>21</sup> The treatment effect of the announced reminder is not statistically significant from zero for some subgroups, likely due to a loss of power from splitting the sample. Nevertheless, we can never reject that the treatment effect of the announced reminder relative to the unannounced reminder is equal across groups in these heterogeneity tests.

We also note that, for the less-valuable offer, the no-deadline group catches up to the deadline group within 13 days after the deadline (Appendix-Figure C.15). This is consistent with the model, reflecting the variance in cost draws over time. Intuitively, those with a lower value of the offer might postpone adopting if they have relatively high cost draws (are fairly busy) during the first 8 days and face no deadline. The chances of getting a sufficiently low cost draw (being sufficiently not busy) to adopt are higher with an additional 13 days of cost draws.

Overall, these results confirm that Prediction 5 holds in our experiment and validate the incentive design also within the deadline vs. no-deadline arms.

## 6.6 Deadline and Take-Up

From Prediction 6, a deadline cannot have a negative effect on cumulative take-up on any date  $t \leq T$  unless the deadline affects trust, memory, or perceived memory. In fact, in a large class of present-bias models (without trust or overconfidence about memory), it is not possible for a deadline to lead to a reduction in cumulative take-up in any period up until the deadline.

Figure 4c shows that on day 1, take-up is 1.2 pp lower in the deadline than in the no-deadline group ( $p = 0.014$ ). This indicates that the deadline must affect either trust, memory, or perceived memory (or something outside of our model). It is unlikely that the deadline would affect memory in a way that could reduce first-day take-up, as it would need to somehow *reduce* first-day memory. The deadline could affect perceived memory in a similar way as an announced reminder (i.e., the manager thinks that they will now remember on the day of the deadline, or marks the deadline on their calendar), but given that the announced reminder did not have a negative effect on first-day take-up, it is unlikely that the mechanism through which the negative effect of the deadline on first-day take-up operates is by changing the perceived probability of remembering on the day of the deadline,  $\hat{p}_T$ . The remaining potential explanation is that the deadline affects trust, and in Section 7, we provide suggestive evidence that the mechanism behind the negative deadline effect on first-day take-up is indeed through the deadline reducing trust in the offer.

## 6.7 Deadline and Present Bias

We now turn to Prediction 7, which states that the more present-biased a manager is, the larger is the treatment effect of the deadline on cumulative adoption by the deadline. Vice versa, the lack of a deadline effect on cumulative adoption by the deadline would imply limited or no present bias in this context.

Figure 4c shows that by day 8, there is no statistically significant difference in cumulative take-up between the deadline and no-deadline groups. The point estimate of the effect of a deadline on cumulative take-up by the deadline is positive but not statistically significant, at 0.7 pp ( $p = 0.228$ ).

Going beyond the deadline, Appendix-Figure C.6d shows that the no-deadline group catches up to the deadline group within 3 days after the deadline, which is expected (and consistent with the model) if firms are experiencing different cost draws each day in terms of how busy they are. Beyond 3 days after the deadline, take-up in the no-deadline group surpasses that of the deadline group, and six months after the deadline there is 2 pp higher take-up in the no-deadline group.

The lack of a statistically significant deadline effect on cumulative take-up by the deadline implies that, on average, managers display insufficient present bias for it to affect their adoption of the profitable opportunity. However, this lack of a statistically significant average treatment effect on cumulative take-up by the deadline masks substantial heterogeneity by firm size, which is one of the main two variables that we pre-specified using for heterogeneity tests in our pre-analysis plan.

**Heterogeneity in deadline effect by firm size.** Appendix-Figure C.9c shows that there is a statistically significant heterogeneous treatment effect of the deadline by firm size. Among larger firms, i.e., those with above-median baseline sales, the deadline has no effect, which by Prediction 7 indicates that larger firms exhibit little or no present bias. Among smaller firms, instead, the deadline increases cumulative take-up by the deadline by 2 pp. The difference in the estimated treatment effect of a deadline by firm size is statistically significant at the 5% level ( $p = 0.014$ ).

The positive effect of a deadline for smaller firms, compared to the lack of an effect for larger firms, suggests that either decision-making in smaller firms exhibits more present bias (by Prediction 7 if other factors are similar between smaller and larger firms), or other differences between smaller and larger firms lead the deadline to only have an effect for smaller firms. As for the first possibility, differences in present bias across the firm size distribution could be due to smaller firms being run by more present-biased individuals, or it could be that managers are equally present-biased in smaller and larger firms but smaller firms lack organizational structures to mitigate the effects of the manager's present bias. Smaller firms potentially being run by more present-biased individuals could be due to selection (e.g., larger firms appoint less present-biased individuals to run the firm) or due to firm growth being a function of present bias (e.g., firms run by less present-biased individuals take advantage of more profitable opportunities and thus grow more).

As for the second possibility, other explanatory factors within the model might include differences across small and large firms in the relative value of the offer, in the relative opportunity cost of adopting, or in the day-to-day variance in the opportunity cost of adopting. Small and large firms could also differ in other factors that affect whether a deadline has an effect but that are not in the model.

**Lack of heterogeneity in deadline effect by other factors.** To assess the potential role of other explanatory factors, we conducted a large array of heterogeneity tests. First, we test for heterogeneity based on firm type—which is the other one of the two main variables to test for heterogeneity on that we pre-specified (Appendix-Table C.9, column 3). Second, we test for heterogeneous effects in the other variables we pre-specified, which are owner age, owner sex, and the month-over-month change in baseline sales from August to September 2020 (excluding the days of the experiment, September 29 and 30) to capture firm growth (Appendix-Table C.10). Third, we test for heterogeneity by percent sales using the technology (Appendix-Table C.7) or whether the owner is the recipient of the emails (Appendix-Table C.7). Fourth, we test for heterogeneous effects of the deadline on cumulative adoption by the deadline across the three reminder groups: announced reminder, unannounced reminder, and no reminder (Appendix-Figure C.14d). We do not find heterogeneity in the deadline effect for any of these variables.

**Alternative explanations of deadline effect.** We also investigated other (unmodeled) potential mechanisms behind the effects or lack of effects of the deadline treatment.

First, to the extent that we do estimate a deadline effect (for smaller firms and less-valuable offers), an alternative mechanism might be that deadlines create a sense of scarcity. Thus, we asked managers a survey question on why they thought the offer had a deadline. Only 11.5% of managers gave responses consistent with the deadline creating scarcity. The vast majority of managers thought that the offer had a deadline because deadlines are a usual business practice or common marketing tool (Appendix-Figure C.13b). Furthermore, given that the deadline only had an effect on firms randomized into the treatment arm with *less-valuable* offers, it is unlikely that a deadline for a less-valuable offer would create a sense of scarcity whereas a deadline for a more-valuable offer would not.

Second, to the extent that we do *not* estimate a deadline effect (for larger firms and more-valuable offers), the mechanism might be that managers did not see the sentence of the email that includes the deadline, even though “fill the form by October 6” was printed in bold red letters to make it easily noticeable (Figure 1). In the survey, we asked directly whether managers in the deadline arm noticed that the offer had a deadline, and 67.7% said yes. Furthermore, given that the deadline did have an effect for the less-valuable offer, it seems implausible that managers with a *more-valuable* offer would pay *less* attention to the email and thus be more likely to miss that it had a deadline. We also note that the announcement about a future reminder in the announced-reminder arm was even further down the email (also in bold red lettering), and the announced reminder did have an effect so managers must have seen it.

## 7 Mechanisms: The Role of Trust

In this section we explore the hypotheses that the larger effect of the announced relative to the unannounced reminder can be linked to increasing how much managers trust the offer from the FinTech firm, and that the deadline’s negative effect on first-day take-up can be explained by reducing trust.

### 7.1 Mechanisms Behind Announced Reminder Effect

**Perception of Offer Value.** We first test whether the announcement and subsequent receipt of the announced reminder changed managers’ perceptions of the offer. To that end, we asked managers who received a reminder, either announced or unannounced, the question: “Did receiving the reminder change your perception of the value of the offer?” The comparison of their answers in Appendix-Figure C.16 reveals that receiving the announced reminder caused a large 16.1 pp increase in the likelihood that the manager responded that the reminder changed their perception of the offer’s value (statistically significant at the 5% level), relative to a base of 23.1% responding yes to this question in the unannounced-reminder group.

In addition, we asked the open-ended follow-up question “Why did the reminder change your perception of the offer’s value?” Comparing responses in the announced- and unannounced-reminder groups, we find that there were more responses related to trust in the announced reminder group—such as “I had doubts and didn’t trust whether it was from [FinTech company]” and “[the reminder] gave it credibility.”

**Heterogeneity in Trust.** Motivated by the mentions of trust in the answers to the open-ended question, we next leverage the general social survey (GSS) measures that we collected in the survey. Specifically, survey respondents are asked to describe how much they agree with the following statements on a 1 to 5 scale, where 5 is *strongly agree*, 4 is *agree*, 3 is *neither agree nor disagree*, 2 is *disagree*, and 1 is *strongly disagree*. The survey questions measure managers’ general levels of trust in advertised offers (“I trust advertised offers”), reciprocity (“I am more inclined to do business with people who live up to their promises”), procrastination (“I tend to postpone tasks, even when I know it is better to do them immediately”), memory (“I tend to have good memory about pending tasks that I have to do and complete”), overconfidence about memory (“I tend to think my memory is better than it really is”), and attention (“I can focus completely when I have to finish a task”).

We use the answers to these questions to test for heterogeneous treatment effects of the announced relative to the unannounced reminder across firms with different characteristics. For each survey measure, we create an indicator variable  $\mathbb{1}(\text{High survey measure})_i$ , which we set equal to 1



if the respondent agrees or strongly agrees with the respective statement. We then estimate the following regression combining the administrative and survey data on the sample that received either an announced or unannounced reminder:

$$y_i = \alpha + \beta_1 \mathbb{1}(\text{High survey measure})_i + \beta_2 \mathbb{1}(\text{Announced reminder})_i + \beta_3 \mathbb{1}(\text{High survey measure})_i \times \mathbb{1}(\text{Announced reminder})_i + \varepsilon_i, \quad (6)$$

where  $y_i$  is an indicator variable equal to one if firm  $i$  accepts the offer any time until March 31. The coefficient  $\beta_3$  estimates the heterogeneous treatment effect of the announced reminder by survey measure. For example, for the first survey measure, the estimate  $\hat{\beta}_3$  reveals whether the announced reminder has a differential treatment effect for firms that trust advertised offers more.

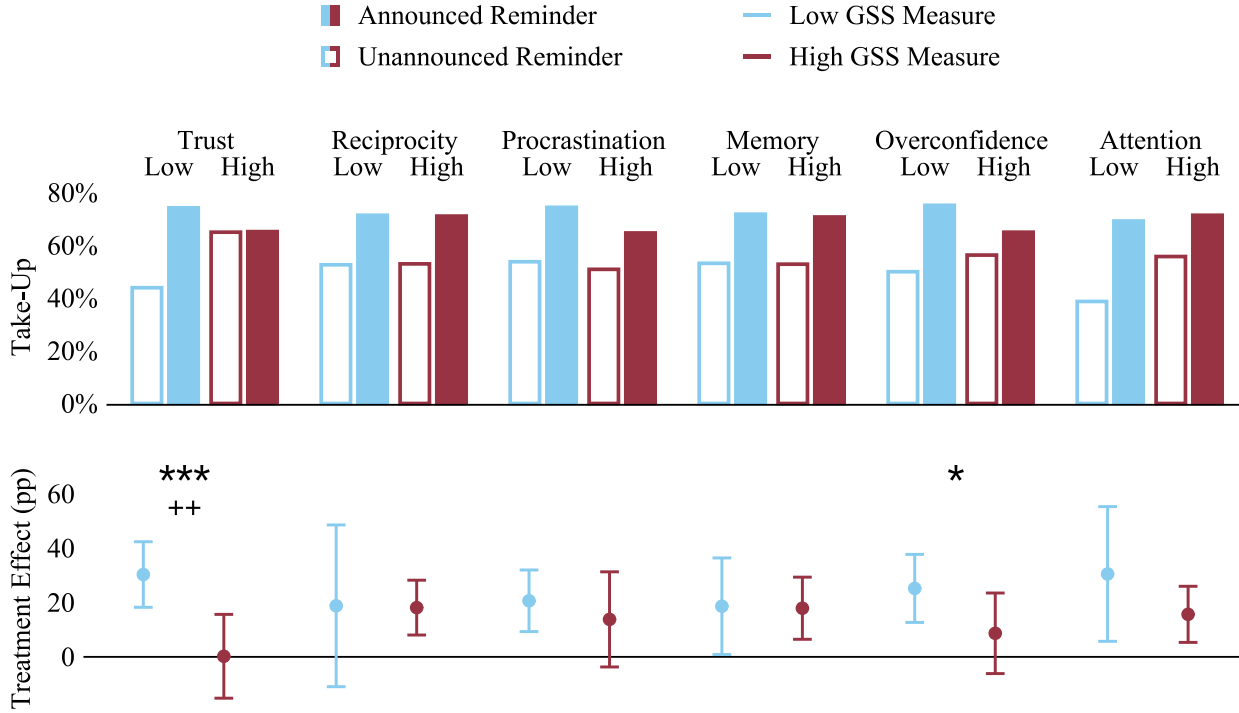
The upper panel of Figure 5 includes four bars for each GSS survey measure, which show take-up separately by treatment arm,  $\mathbb{1}(\text{Announced reminder})_i \in \{0, 1\}$  (empty versus filled bars), and by value of the survey measure,  $\mathbb{1}(\text{High survey measure})_i \in \{0, 1\}$  (blue versus red bars). Note that take-up rates in the survey are higher than in the administrative data because we oversampled firms that accepted the offer to be better powered.

The panel shows that take-up is higher among managers who received the announced reminder rather than the unannounced reminder for almost all GSS sample splits. The exception is the high-trust group. Here, take-up rates are identical and high in both the announced- and the unannounced-reminder groups, indicating that the announced reminder did not have an effect on take-up for firms with high general trust. In contrast, comparing the two bars in the low-trust column, it is clear that the announced reminder did have a large effect on take-up for firms with low general trust.

The lower panel of Figure 5 shows the corresponding estimated effect of announced reminders by survey measure, where the coefficients in the “Low” columns correspond to  $\hat{\beta}_2$  from estimating equation (6) and the coefficients in the “High” columns correspond to  $\hat{\beta}_2 + \hat{\beta}_3$ . Above each pair of treatment effects, we use stars to show the statistical significance of  $\hat{\beta}_3$  without adjusting for multiple hypothesis testing and plus signs to show statistical significance using the Romano and Wolf (2005) multiple hypothesis correction across all of the regressions we are using to estimate heterogeneous treatment effects (where there is one regression per GSS measure). The  $\hat{\beta}_3$  coefficient estimates, standard errors, and Romano-Wolf adjusted  $p$ -values are in Appendix-Table C.11.

For the split by high and low trust, we see that the estimated treatment effect of announced reminders is entirely concentrated among firms with low general trust. In contrast, for firms with high general trust, the treatment effect of announced reminders relative to unannounced reminders is virtually zero and not statistically significant. Appendix-Table C.11 shows that the heterogeneous treatment effect for high vs. low levels of trust is statistically significant at the 1% level before adjusting for multiple hypothesis testing ( $p = 0.003$ ) and at the 5% level after adjusting ( $p = 0.029$ ). In contrast, we do not see statistically significant coefficients on heterogeneity tests

Figure 5: Heterogeneous Effects of Announced Reminder by GSS Measures



This figure shows heterogeneous take-up rates and treatment effects of the announced relative to the unannounced reminder, separately by each of the six GSS survey sample splits. Note that take-up rates in the survey are higher than in the administrative data because we oversampled firms that accepted the offer to be better powered. The lower graph shows the corresponding estimated effect of announced reminders by GSS measure, where the coefficients in the “Low” columns correspond to  $\hat{\beta}_2$  from estimating equation (6) and the coefficients in the “High” columns correspond to  $\hat{\beta}_2 + \hat{\beta}_3$ . Data include firms with announced and unannounced reminders in the survey sample, excluding those who did not respond to the GSS questions ( $N = 388$ ). Above each pair of treatment effects we show the statistical significance of  $\hat{\beta}_3$  (the  $\hat{\beta}_3$  coefficient estimates and standard errors are in Appendix-Table C.11), with stars for heteroskedasticity-robust  $p$ -values (\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ) and plus signs for Romano-Wolf multiple-hypothesis-adjusted  $p$ -values (+  $p < 0.1$ , ++  $p < 0.05$ , +++  $p < 0.01$ ). GSS = general social survey.

for the other survey measures.<sup>22</sup> We also note that we did not find any heterogeneous treatment effects of the announced reminder, relative to the unannounced reminder, across any of the firm and owner characteristics explored in Section 6.3.

For completeness, we confirm that there is no heterogeneous treatment effect of reminders in and of themselves (without prior announcement). When we repeat the exercise of splitting the sample by high and low survey measures, but compare the unannounced-reminder group to the *no-reminder* group, we do not find heterogeneous treatment effects among firms with high vs. low levels of general trust. After adjusting for multiple hypothesis testing, there are no statistically significant heterogeneous treatment effects of the unannounced reminder relative to no reminder by GSS measures (Appendix-Table C.12).

<sup>22</sup>Prior to adjusting for multiple hypothesis testing, there is a heterogeneous treatment effect for overconfidence that is statistically significant only at the 10% level ( $p = 0.096$ ), while after adjusting it is not significant ( $p = 0.411$ ).

These results corroborate the notion that the estimated effect of the announced reminder relative to the unannounced reminder reflects its effectiveness in generating trust that the offer represents a valuable business opportunity. Managers who likely trusted the offer already are not affected by the announcement, while managers who likely did not trust the offer initially are the ones who are more likely to adopt in response to the announcement.

**Heterogeneity in Length of Business Relationship.** Finally, we probe the trust channel in the administrative data, where we have a much larger sample but no direct measure of trust. Instead, we employ the length of a firm’s business relationship with the FinTech company—i.e. the number of months the firm has been using the payment technology—as an indirect proxy, as trust may have been fostered over time through interfirm interactions.

In Figure 6, we plot take-up rates against the length of the business relationship, separately for the announced- and unannounced-reminder arms. As expected, firms that have been using the payments technology for longer have higher overall *levels* of take-up. The treatment effect of the announced reminder relative to the unannounced reminder, instead, is concentrated among firms that have been using the technology for less time, as the difference between the blue and red lines indicates.

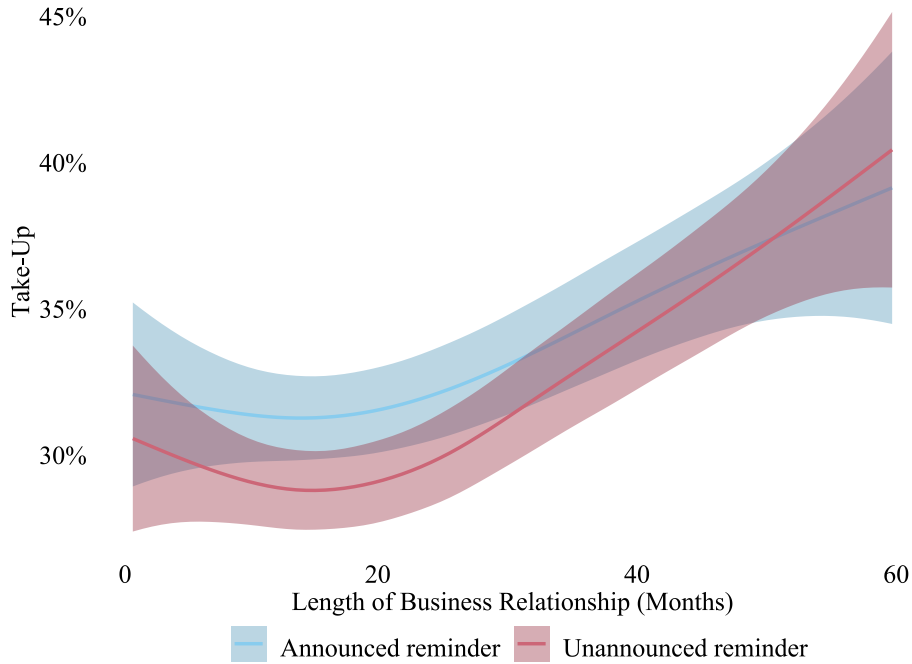
To formally test whether the treatment effect of the announced reminder relative to the unannounced reminder is decreasing in the length of a firm’s business relationship with the FinTech company, we follow a procedure described in Appendix D. Briefly, since our trust proxy, length of the business relationship, could be correlated with other factors, we control flexibly for the other covariates in the administrative data. First, we assume that take-up is a linear function of two (non-linear) polynomials: (i) a function of the length of the business relationship and treatment, where treatment here refers to the announced reminder rather than the unannounced reminder; and (ii) a function of other covariates and treatment. We then use a machine learning algorithm to estimate predicted individual treatment effects of the announced reminder. Using our estimates from this procedure, we formally test the negative relationship between the length of the business relationship and the treatment effect of the announced reminder seen in Figure 6. We use a bootstrapping procedure to test the statistical significance of this negative relationship, and find that the relationship is negative in 92.5% of the bootstrap samples.<sup>23</sup>

**Alternative Explanations.** The results presented so far corroborate the hypothesized role of trust. Nevertheless, we consider three alternative mechanisms for the effect of announced re-

---

<sup>23</sup>Because the polynomial that minimizes root mean squared error turns out to be non-linear (in particular, quadratic) in the interaction between the length of the business relationship and treatment, we evaluate whether the relationship is negative at the median value of the length of the business relationship. The relationship is also negative in 93.3%, 90.5%, and 81.3% of the bootstrap samples when evaluated at the 25th percentile, mean, and 75th percentile.

Figure 6: Effect of Announced Reminder by Length of Business Relationship



This figure plots take-up rates of the offer against the number of months firms have been using the FinTech technology, split by reminder type. Take-up is measured from September 29 to March 31. Colored lines are local polynomial regression fits, and shaded regions are 95% confidence intervals. The blue line shows take-up in the announced-reminder arm and the red line shows take-up in the unannounced-reminder arm. Thus, the difference between the blue and red lines is the treatment effect of the announced reminder relative to the unannounced reminder.  $N = 16,254$  firms with announced or unannounced reminders, including firms above the 95th percentile of months using the technology which are omitted from the graph for legibility.

minders on take-up.

First, we consider differences in information acquisition. We ask whether the announcement of a future reminder may induce firms to check how valuable or profitable the offer is to them, knowing they can adopt when they get the reminder. For example, managers may not know their current merchant fee (which we decided not to include in the email to avoid adding confusion by including too many numbers in the email); if so, they might take the time between the initial message and the reminder to log into their account and check their current merchant fee.

We address this alternative mechanism using both survey and administrative data. In the survey, we asked managers “What was your fee with [the FinTech provider] the week before you received the offer?” We compare their response to the correct answer, and find that managers are fairly accurate (Appendix-Figure C.17): 23% of managers reported their fee precisely, and the vast majority who were not perfectly accurate reported that their fee was 4% which could be due to rounding up the 3.5–3.75% fee or including the value-added tax that is charged on the fee. Thus, the vast majority of managers either accurately reported their fee or slightly overreported it, which if anything would lead them to think the offer was even more profitable than it was.

In addition, we use the administrative data on whether firms logged in to their accounts to check their current fee or sales. We create outcome indicator variables if anyone in the firm logged into their account or checked the amount of deposits from electronic sales in the days between when the initial offer was sent and before the reminder was sent. As shown in Appendix-Table C.13, we find that firms that were told about a future announced reminder were not more likely to check their online accounts over the course of the experiment compared to firms in the unannounced reminder group.

Finally, we return to the survey data and analyze the answers to two questions. In one question, we asked managers who did not accept the offer prior to the reminder day and received an announced reminder, “Did you do anything between receiving the initial email and receiving the reminder so that you would know whether to take up the offer when you received the reminder?” Nearly all managers (92.4%) reported not taking any particular action to evaluate the offer between the time they received the initial email and the reminder. Among the remaining 7.6% who do report taking some action between receiving the initial offer and receiving the announced reminder, only 2 managers reported calculating whether they should accept the offer. Another question asked “Why did you wait until {days to accept} days later [to accept the offer]?”<sup>24</sup> As shown in Appendix-Figure C.4, only 12.3% of managers replied they needed to discuss or think about the offer first, and there is no statistically significant difference in this proportion between the announced- and unannounced-reminder arms ( $p = 0.361$ ).

We conclude that the possibility of the announced reminder leading firms to take additional steps to evaluate the offer prior to the date on which they knew they would receive the reminder is not a plausible explanation for our findings.

As a second alternative mechanism, we consider the possibility that managers in the unannounced reminder group may have felt annoyed when they received the reminder or ashamed that they had not yet adopted the profitable opportunity. As a result, they may have been less likely to adopt than if they had been told in advance that they would receive the reminder. Feeling ashamed could represent an “ostrich effect” where receiving the unannounced reminder made the decision maker “stick their head in the sand” and avoid making a decision (as in Olafsson and Pagel, 2017).

To test for these or other negative responses to an unannounced reminder, we asked managers who received a reminder an open-ended question to tell us how they felt when they received the reminder (see Appendix-Figure C.18). Only 2.5% of managers responded that they were annoyed by the reminder, and there is no statistically significant difference between the announced- and unannounced-reminder groups ( $p = 0.494$ ). Instead, the most common responses indicated that

---

<sup>24</sup> The full survey question is: “We sent you the emails and SMS to let you know about this offer on September 29, but we see that you filled the form on {activation date}. Why did you wait until {days to accept} day(s) later?” This survey question was asked to firms with a deadline, who accepted after the first day of the offer, and recalled accepting or clicking on the offer.

the reminder made managers feel important as a client.

A third possibility is that knowing a reminder will come on a particular day helps solve organizational frictions: for example, knowing when the reminder will come makes it easier to schedule a meeting where the decision of whether to adopt the profitable opportunity will be discussed. It is worth noting, however, that if the firm schedules a meeting about it they are unlikely to forget to adopt, so they should not need to know a reminder is coming to schedule such a meeting. In addition, if the mechanism were organizational frictions we would expect differential effects depending on the size of the firm (in terms of number of employees and baseline sales) and depending on whether the firm is a single-person firm or has more than one employee, as organizational frictions are more likely in larger firms. However, in Section 6.3 we did not see any heterogeneous treatment effects for these variables. Finally, the deadline also provides a particular date to make the decision by, so if this were the mechanism behind the announced reminder effect, the deadline should have a similar effect among larger firms (again because organizational frictions are more likely in larger firms)—but it did not.

## **7.2 Mechanisms Behind Negative First-Day Deadline Effect**

Next, we examine whether trust also plays a role in the negative treatment effect of the deadline on take-up on the day the initial offer emails and SMS messages were sent. Our model predicts that the deadline could only reduce take-up on any day up until (and including) the deadline date if it affected trust, memory, or perceived memory—and not otherwise. Section 6.6 explains why it is unlikely that the deadline affected memory or perceived memory, leaving trust as the only other potential explanation (in our model).

The main piece of evidence that the negative first-day deadline effect is due to the deadline reducing trust comes from Figure C.14d. In particular, the deadline only has a negative first-day effect on take-up in the no-reminder and unannounced-reminder arms (which received the same initial email), whereas the treatment effect of the deadline is very close to zero and not statistically significant in the announced-reminder arm. In other words, when the email with a deadline also included the sentence announcing a future reminder (which we have found likely increased trust in the offer), this offset the negative effect of the deadline on first-day take-up. While this is not direct evidence that trust is also the channel through which the deadline reduced first-day take-up, it is suggestive.

In summary, interfirm trust emerges as a plausible and important determinant of companies' willingness to adopt a profitable business opportunity with their business partner.

## 8 Elasticity of Electronic Sales

We also evaluate the impact of a lower merchant fee on electronic sales through the payment technology. We use the following estimating equation:

$$y_{it} = \beta \cdot Treated_i \times Post_t + \gamma_i + \delta_t + \varepsilon_{it}, \quad (7)$$

where  $y_{it}$  is an outcome measuring sales using the payments technology,  $i$  denotes a firm,  $t$  denotes a month,  $\gamma_i$  are firm fixed effects, and  $\delta_t$  are time fixed effects. Standard errors are clustered at the firm level.  $Treated_i$  is an indicator for a firm that received a lower merchant fee offer, i.e., a firm in any treatment arm except the control group, and  $Post_t$  is an indicator that equals one during any time period after we sent the offers. Our main coefficient of interest  $\beta$  measures the intent-to-treat (ITT) effect of receiving an offer on use of the FinTech payments technology. To estimate the treatment on the treated (TOT), i.e., the effect on the firms that adopted the lower merchant fee, we replace  $Treated_i$  with  $Accepted_i$  in specification (7) and instrument  $Accepted_i$  with  $Treated_i$ .

Panel A of Table C.14 shows the ITT effect of the lower merchant fee on electronic sales through the payment technology. The first two columns of Panel A show regression results with intensive measures of payment usage: log sales volume in pesos and log number of transactions (both plus 1). Firms that received the offer increased the average sales volume and number of payments they transacted with the payment technology by 10.7% and 3%, respectively.<sup>25</sup> The third column of Panel A shows the regression results with the extensive measure of payment usage: an indicator if the firm made at least one transaction on or after the current month. Firms that received the offer increased their probability of continuing to use the payment technology by 1.3 pp.

Panel B of the same table shows the TOT effect of the lower merchant fee on electronic sales through the payment technology. Firms that accepted the offer increased the sales volume and number of payments they transacted with the payment technology by 42% and 10.8%, respectively. Firms that accepted the offer also increased their probability of continuing to use the payment technology by 4.3 pp. The control mean of the probability of continuing to use the payment technology is 84.7%. This means that firms that accepted the offer were, in relative terms, 5.1% more likely to continue to use the payment technology on or after a given month compared to the control mean. Because the increase in sales processed through the FinTech technology by firms that accepted the lower merchant fee (42%) was larger than the decrease in our FinTech partner's revenues from these firms paying a lower fee on sales they would have made anyway (up to  $(3.75 - 2.75)/3.75 = 27\%$ ), offering the lower merchant fee turned out to increase the profits of our FinTech partner.

Our findings suggest that lowering merchant fees can increase payment usage on both the

---

<sup>25</sup> These percent changes are calculated as  $(\exp(\beta) - 1) \times 100\%$ .

intensive and extensive margins. This raises the question of how firms encouraged customers to use cards more, or discouraged them from using cards before the treatment. It is possible that some firms previously preferred cash payments due to the cost of accepting card payments and used various methods to steer consumers towards cash transactions. For example, firms can surcharge a percentage or fixed amount to card-paying customers, set a minimum threshold for paying by card, hide the POS terminal in a drawer and only bring it out if they perceive that they will lose a sale by not accepting card payments, or say that the POS terminal is not working. Cash has various indirect costs, however (Alvarez et al., 2022; Bachas et al., 2020), so with a lower fee per transaction, some firms may have switched from preferring cash payments to preferring card payments.

One way for a firm to incentivize more card payments by customers is to eliminate surcharges. Surcharging is the practice of passing through the merchant fee only to customers who pay by card, and it is prevalent in this context: Higgins (2024) finds that 63% of small retailers with POS terminals surcharge their customers who pay by card in Mexico. To explore the impact of lower fees on firms, we asked those who accepted the lower fee to respond to the open-ended question, “Is this offer working for your business? What impact has it had?” Despite the open-ended nature of this question, 24.1% of managers replied that prior to receiving the lower fee, they would surcharge customers who paid by card, while after receiving the lower fee they stopped surcharging and absorbed the cost (Appendix-Figure C.19). For example, one manager replied, “[The effect is] very good, [we] don’t charge the fee to customers anymore.”

## **9 Conclusion**

The analyses in this paper reveal a significant role of non-standard determinants in explaining firm decision-making. Well-known behavioral determinants of individuals failing to act in non-managerial settings appear to also bind in the context of managerial decision-making within firms. In particular, imperfect memory and distorted beliefs about the probability of remembering in the future emerge as significant determinants in our setting, while present bias does not appear to play an important role for larger firms, though it potentially does play a role for smaller firms. Beyond those three factors, we provide evidence of distrust as a key explanatory variable. Distrust is likely an important friction in many firm-to-firm relationships, and we provide evidence on an intervention that increased trust in a particular firm-to-firm interaction, which thereby led to increased adoption of a profitable opportunity.

While the role of these determinants in inhibiting decision-making has been much discussed in the consumer-level literature, they have received less attention when studying firms and impediments to their profit maximization and growth. Our findings suggest that the analysis of firm decision-making would benefit from researchers considering mechanisms beyond the traditional



economic frictions that explain non-adoption of profitable opportunities.

## References

- Abel, Andrew and Janice Eberly (1994). “A Unified Model of Investment Under Uncertainty.” *American Economic Review* 84(5), 1369–1384.
- Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P Vasquez (2022). “The Effects of Joining Multi-national Supply Chains: New Evidence from Firm-to-Firm Linkages.” *Quarterly Journal of Economics* 137(3), 1495–1552.
- Alvarez, Fernando, David Argente, Rafael Jimenez, and Francesco Lippi (2022). “Cash: A Blessing or a Curse?” *Journal of Monetary Economics* 125, 85–128.
- Ashraf, N., D. Karlan, and W. Yin (2006). “Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines.” *Quarterly Journal of Economics* 121(2), 635–672.
- Atkin, David, Azam Chaudhry, Shamyala Chaudry, Amit K Khandelwal, and Eric Verhoogen (2017). “Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan.” *Quarterly Journal of Economics* 132(3), 1101–1164.
- Bachas, Pierre, Paul Gertler, Sean Higgins, and Enrique Seira (2021). “How Debit Cards Enable the Poor to Save More.” *Journal of Finance* 76(4), 1913–1957.
- Bachas, Pierre, Sean Higgins, and Anders Jensen (2020). “Towards a Cashless Economy? Evidence from the Elasticity of Cash Deposits of Mexican Firms.”
- Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan (2024a). “Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs?”
- Banerjee, Abhijit, Greg Fischer, Dean Karlan, Matt Lowe, and Benjamin N Roth (2024b). “Do Collusive Norms Maximize Profits? Evidence from a Vegetable Market Experiment in India.”
- Banerjee, Abhijit V. and Esther Duflo (2000). “Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry.” *Quarterly Journal of Economics* 115(3), 989–1017.
- Beaman, Lori, Jeremy Magruder, and Jonathan Robinson (2014). “Minding Small Change among Small Firms in Kenya.” *Journal of Development Economics* 108, 69–86.
- Bernstein, Shai, Emanuele Colonnelli, Mitchell Hoffman, and Benjamin Iverson (2024). “Life After Death: A Field Experiment with Small Businesses on Information Frictions, Stigma, and Bankruptcy.”
- Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts (2013). “Does Management Matter? Evidence from India.” *Quarterly Journal of Economics* 128(1), 1–51.
- Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen (2014). “The New Empirical Empirics of Management.” *Journal of the European Economic Association* 12(4), 835–876.
- Bloom, Nicholas, Raffaella Sadun, and John Van Reenen (2012). “The Organization of Firms Across Countries.” *Quarterly Journal of Economics* 127(4), 1663–1705.
- Bruhn, Miriam, Dean Karlan, and Antoinette Schoar (2018). “The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico.” *Journal of Political Economy* 126(2), 635–687.
- Bryan, Gharad, Dean Karlan, and Scott Nelson (2010). “Commitment Devices.” *Annual Review of Economics* 2(1), 671–698.

- Cai, Jing and Adam Szeidl (2018). “Interfirm Relationships and Business Performance.” *Quarterly Journal of Economics* 133(3), 1229–1282.
- Calzolari, Giacomo and Mattia Nardotto (2017). “Effective Reminders.” *Management Science* 63(9), 2915–2932.
- Camerer, Colin and Dan Lovallo (1999). “Overconfidence and Excess Entry: An Experimental Approach.” *American Economic Review* 89(1), 306–318.
- Camerer, Colin and Ulrike Malmendier (2007). “Behavioral Economics of Organizations.” *Behavioral Economics and Its Applications*. Ed. by Peter Diamond and Hannu Vartiainen. Princeton University Press.
- Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky (2022). “Who Chooses Commitment? Evidence and Welfare Implications.” *Review of Economic Studies* 89(3), 1205–1244.
- Celhay, Pablo A., Paul J. Gertler, Paula Giovagnoli, and Christel Vermeersch (2019). “Long-Run Effects of Temporary Incentives on Medical Care Productivity.” *American Economic Journal: Applied Economics* 11(3), 92–127.
- Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery (2013). “Barriers to Household Risk Management: Evidence from India.” *American Economic Journal: Applied Economics* 5(1), 104–135.
- D’Acunto, Francesco, Marcel Prokopczuk, and Michael Weber (2019). “Historical Antisemitism, Ethnic Specialization, and Financial Development.” *Review of Economic Studies* 86(3), 1170–1206.
- D’Acunto, Francesco, Jin Xie, and Jiaquan Yao (2024). “Trust and Contracts: Empirical Evidence.”
- Daft, Richard L. (2015). *Organization Theory and Design*. 12th edition. Boston, MA: Cengage Learning.
- Dai, Hengchen, Silvia Saccardo, Maria A. Han, Lily Roh, Naveen Raja, Sitaram Vangala, Hardikumar Modi, Shital Pandya, Michael Sloyan, and Daniel M. Croymans (2021). “Behavioural Nudges Increase COVID-19 Vaccinations.” *Nature* 597(7876), 404–409.
- de Mel, Suresh, David McKenzie, and Christopher Woodruff (2010). “Who Are the Microentrepreneur Owners? Evidence from Sri Lanka on Tokman versus De Soto.” *International Differences in Entrepreneurship*. Ed. by Josh Lerner and Antoinette Schoar. University of Chicago Press.
- DellaVigna, Stefano and Matthew Gentzkow (2019). “Uniform Pricing in U.S. Retail Chains.” *Quarterly Journal of Economics* 134(4), 2011–2084.
- DellaVigna, Stefano and Ulrike Malmendier (2004). “Contract Design and Self-Control: Theory and Evidence.” *Quarterly Journal of Economics* 119(2), 353–402.
- DellaVigna, Stefano and Ulrike Malmendier (2006). “Paying Not to Go to the Gym.” *American Economic Review* 96(3), 694–719.
- Duflo, Esther, Michael Kremer, and Jonathan Robinson (2011). “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya.” *American Economic Review* 101(6), 2350–2390.
- Ericson, Keith M. Marzilli (2011). “Forgetting We Forget: Overconfidence and Memory.” *Journal of the European Economic Association* 9(1), 43–60.
- Ericson, Keith M. Marzilli (2014). “Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange.” *American Economic Journal: Economic Policy* 6(1), 38–64.

- Ericson, Keith M. Marzilli (2017). “On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation.” *Journal of the European Economic Association* 15(3), 692–719.
- Feigenberg, B., E. Field, and R. Pande (2013). “The Economic Returns to Social Interaction: Experimental Evidence from Microfinance.” *Review of Economic Studies* 80(4), 1459–1483.
- Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2013). “Human Capital and Regional Development.” *Quarterly Journal of Economics* 128(1), 105–164.
- Giné, Xavier, Dean Karlan, and Jonathan Zinman (2010). “Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation.” *American Economic Journal: Applied Economics* 2(4), 213–235.
- Giorcelli, Michela (2019). “The Long-Term Effects of Management and Technology Transfers.” *American Economic Review* 109(1), 121–152.
- Graham, John R. and Campbell R. Harvey (2001). “The Theory and Practice of Corporate Finance: Evidence from the Field.” *Journal of Financial Economics* 60(2), 187–243.
- Graham, John R., Campbell R. Harvey, and Manju Puri (2013). “Managerial Attitudes and Corporate Actions.” *Journal of Financial Economics* 109(1), 103–121.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2004). “The Role of Social Capital in Financial Development.” *American Economic Review* 94(3), 526–556.
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2008). “Trusting the Stock Market.” *Journal of Finance* 63(6), 2557–2600.
- Handel, Benjamin R (2013). “Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts.” *American Economic Review* 103(7), 2643–2682.
- Hanna, Rema, Sendhil Mullainathan, and Joshua Schwartzstein (2014). “Learning Through Noticing: Theory and Evidence from a Field Experiment.” *Quarterly Journal of Economics* 129(3), 1311–1353.
- Hardy, Morgan and Jamie McCasland (2023). “Are Small Firms Labor Constrained? Experimental Evidence from Ghana.” *American Economic Journal: Applied Economics* 15(2), 253–284.
- Higgins, Sean (2024). “Financial Technology Adoption: Network Externalities of Cashless Payments in Mexico.” *American Economic Review* 114(11), 3469–3512.
- Jagannathan, Ravi, David A. Matsa, Iwan Meier, and Vefa Tarhan (2016). “Why Do Firms Use High Discount Rates?” *Journal of Financial Economics* 120(3), 445–463.
- Johnson, Eric J, Stephan Meier, and Olivier Toubia (2019). “What’s the Catch? Suspicion of Bank Motives and Sluggish Refinancing.” *Review of Financial Studies* 32(2), 467–495.
- Kaplan, Steven N. and Morten Sorensen (2021). “Are CEOs Different?” *Journal of Finance* 76(4), 1773–1811.
- Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman (2016a). “Getting to the Top of Mind: How Reminders Increase Saving.” *Management Science* 62(12), 3393–3411.
- Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl (2009). “Trust and Social Collateral.” *Quarterly Journal of Economics* 124(3), 1307–1361.
- Karlan, Dean, Melanie Morten, and Jonathan Zinman (2016b). “A Personal Touch in Text Messaging Can Improve Microloan Repayment.” *Behavioral Science & Policy* 1(2), 25–31.
- Kling, J. R., S. Mullainathan, E. Shafir, L. C. Vermeulen, and M. V. Wrobel (2012). “Comparison Friction: Experimental Evidence from Medicare Drug Plans.” *Quarterly Journal of Economics* 127(1), 199–235.

- Kremer, Michael, Jean Lee, Jonathan Robinson, and Olga Rostapshova (2013). “Behavioral Biases and Firm Behavior: Evidence from Kenyan Retail Shops.” *American Economic Review* 103(3), 362–368.
- Kuchler, Theresa and Michaela Pagel (2021). “Sticking to Your Plan: The Role of Present Bias for Credit Card Paydown.” *Journal of Financial Economics* 139(2), 359–388.
- Laibson, D. (1997). “Golden Eggs and Hyperbolic Discounting.” *Quarterly Journal of Economics* 112(2), 443–478.
- Levine, Ross and Yona Rubinstein (2017). “Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?” *Quarterly Journal of Economics* 132(2), 963–1018.
- List, John A, Ian Muir, Devin Pope, and Gregory Sun (2023). “Left-Digit Bias at Lyft.” *Review of Economic Studies* 90(6), 3186–3237.
- Malmendier, Ulrike and Geoffrey Tate (2005). “CEO Overconfidence and Corporate Investment.” *Journal of Finance* 60(6), 2661–2700.
- Malmendier, Ulrike and Geoffrey Tate (2008). “Who Makes Acquisitions? CEO Overconfidence and the Market’s Reaction.” *Journal of Financial Economics* 89(1), 20–43.
- McKenzie, David and Christopher Woodruff (2017). “Business Practices in Small Firms in Developing Countries.” *Management Science* 63(9), 2967–2981.
- McMillan, John and Christopher Woodruff (1999). “Interfirm Relationships and Informal Credit in Vietnam.” *Quarterly Journal of Economics* 114(4), 1285–1320.
- Mishra, Prachi, Nagpurnanand Prabhala, and Raghuram G Rajan (2022). “The Relationship Dilemma: Why Do Banks Differ in the Pace at Which They Adopt New Technology?” *Review of Financial Studies* 35(7), 3418–3466.
- Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny (1991). “The Allocation of Talent: Implications for Growth.” *Quarterly Journal of Economics* 106(2), 503.
- O’Donoghue, Ted and Matthew Rabin (1999). “Doing It Now or Later.” *American Economic Review* 89(1).
- Olafsson, Arna and Michaela Pagel (2017). “The Ostrich in Us: Selective Attention to Financial Accounts, Income, Spending, and Liquidity.”
- Osili, Una Okonkwo and Anna L. Paulson (2008). “Institutions and Financial Development: Evidence from International Migrants in the United States.” *Review of Economics and Statistics* 90(3), 498–517.
- Rigol, Natalia and Benjamin Roth (2024). “Intrinsic Motivation and Referrals Within Firms: Evidence from a Large Microfinance Institution.”
- Romano, Joseph P. and Michael Wolf (2005). “Stepwise Multiple Testing as Formalized Data Snooping.” *Econometrica* 73(4), 1237–1282.
- Schoar, Antoinette (2010). “The Divide between Subsistence and Transformational Entrepreneurship.” *Innovation Policy and the Economy* 10, 57–81.
- Shue, Kelly and Richard R. Townsend (2021). “Can the Market Multiply and Divide? Non-Proportional Thinking in Financial Markets.” *Journal of Finance* 76(5), 2307–2357.
- Strulov-Shlain, Avner (2023). “More Than a Penny’s Worth: Left-Digit Bias and Firm Pricing.” *Review of Economic Studies* 90(5), 2612–2645.
- Tasoff, Joshua and Robert Letzler (2014). “Everyone Believes in Redemption: Nudges and Overoptimism in Costly Task Completion.” *Journal of Economic Behavior & Organization* 107, 107–122.

Verhoogen, Eric (2023). “Firm-Level Upgrading in Developing Countries.” *Journal of Economic Literature* 61(4), 1410–1464.

# Supplemental Appendix

## A Model Proofs

This appendix includes proofs of some properties of the equilibrium described in Section 2.2, as well as the model predictions in Section 2.3.

### A.1 Equilibrium Properties

**Proposition 1** (Optimal Strategy). *The optimal strategy is to adopt in period  $t$  if and only if the cost draw is below  $c_t^*$  as defined by equations (3) and (4).*

*Proof.* Denote the manager's decision of whether to adopt in period  $t$  as  $s_t$  and their perceived future strategies (as well as their perception of their successor's future strategies) as of time  $t$  as  $\{\hat{s}_{t+k}^t\}_{k=1}^\infty$ . We use  $u_\tau^t(\hat{s}_\tau^t)$  to denote the instantaneous utility the manager believes, as of time  $t$ , to receive at time  $\tau$  for taking the action they perceive they are going to take at time  $\tau$ ,  $\hat{s}_\tau^t$ . Note that for  $\tau > t$ , this differs from the time- $\tau$  instantaneous utility they would actually receive by taking action  $\hat{s}_\tau^t$  in period  $\tau$  if they are present biased, hence the need for the superscripts in  $u_\tau^t$ . We can then write the manager's expected utility of the perceived actions at time  $t$  as:

$$V_t = u_t^t(s_t) + \beta \mathbb{E}_t \left[ \sum_{k=1}^{\infty} \left( \prod_{j=1}^k \hat{\rho}_{t+j} \right) \delta^k u_{t+k}^t(\hat{s}_{t+k}^t) \right]. \quad (8)$$

In a perception-perfect equilibrium, the manager chooses the optimal  $s_t$  to maximize  $V_t$  under a dynamically consistent belief that  $\{\hat{s}_{t+k}^t\}_{k=1}^\infty$  is the perceived strategies of the future selves, i.e., there exists a sequence of future beliefs  $\{\{\hat{s}_{t+k}^t\}_{k=1}^\infty\}_{\tau>t}$  such that they are:

1. internally consistent, i.e., for all  $\tau > t$ ,

$$\hat{s}_\tau^t \in \arg \max_a u_\tau^\tau(a) + \hat{\beta} \mathbb{E}_t \left[ \sum_{k=1}^{\infty} \left( \prod_{j=1}^k \hat{\rho}_{t+j} \right) \delta^k u_{\tau+k}^\tau(\hat{s}_{\tau+k}^t) \right],$$

where  $\hat{\beta}$  enters instead of  $\beta$  since the manager believes at time  $t$  that all future selves in  $\tau > t$  will have present bias level  $\hat{\beta}$ ; and

2. externally consistent, i.e., for all  $t < t' < \tau$ ,  $\hat{s}_\tau^t = \hat{s}_\tau^{t'}$ .

The external consistency allows us to drop the superscript in the notation  $\hat{s}_{t+k}^t$ , so we now simply write  $\hat{s}_{t+k}$ . Similarly, note that the instantaneous utility functions are also externally consistent, i.e.,  $u_\tau^t = u_\tau^{t'}$  for all  $t < t' < \tau$ , so we may also drop the superscripts on  $u_\tau^t$  for  $t < \tau$  and write it as  $\hat{u}_\tau$ , where we add a hat to denote the perception as of time  $t$  about future instantaneous utility. For  $\tau = t$ , we may also suppress the superscript and write  $u_\tau^t = u_\tau^\tau$  as  $u_\tau$  since it is the instantaneous

utility in the current period. By a similar argument, we can suppress the superscript on  $\hat{V}_\tau^t$  defined in equation (2).

The internal consistency condition exhibits a recursive structure. To see this, note that we can write the perceived continuation value as

$$\hat{V}_{t+1} = \hat{u}_{t+1}(\hat{s}_{t+1}) + \mathbb{E}_{t+1} \left[ \sum_{k=1}^{\infty} \left( \prod_{j=1}^k \hat{\rho}_{t+1+j} \right) \delta^k \hat{u}_{t+1+k}(\hat{s}_{t+1+k}) \right],$$

where  $\beta$  and  $\hat{\beta}$  do not enter because  $t+1$  is already in the future. Then we can rewrite the internal consistency condition as:

$$\hat{s}_t \in \arg \max_a u_t(a) + \hat{\beta} \hat{\rho}_{t+1} \delta \mathbb{E}_t[\hat{V}_{t+1}].$$

Plugging in the utility form, this condition states that the future self at time  $\tau > t$  is *perceived* to receive a utility of  $\hat{\beta} \delta \alpha_\tau y - c_\tau$  if she adopts, and  $\hat{\beta} \delta \hat{\rho}_{\tau+1} \mathbb{E}_\tau[\hat{V}_{\tau+1}]$  otherwise, and to act optimally. Therefore, the optimal threshold in the perceived strategy of the future self in period  $\tau > t$  is

$$\hat{c}_\tau^* = \hat{\beta} \delta (\alpha_\tau y - \hat{\rho}_{\tau+1} \mathbb{E}_\tau[\hat{V}_{\tau+1}]), \quad (9)$$

which has a hat on  $\hat{c}_\tau^*$  since it is a direct function of  $\hat{\beta}$ .

Similarly, the optimality condition for the manager at time  $t$  is

$$s_t \in \arg \max_a u_t(a) + \beta \hat{\rho}_{t+1} \delta \mathbb{E}_t[\hat{V}_{t+1}].$$

At time  $t$ , the manager receives a utility of  $\beta \delta \alpha_t y - c_t$  if she adopts, and  $\beta \hat{\rho}_{t+1} \delta \mathbb{E}_t[\hat{V}_{t+1}]$  otherwise. Therefore, the optimal threshold for the manager at time  $t$  deciding whether to adopt is

$$c_t^* = \beta \delta (\alpha_t y - \hat{\rho}_{t+1} \mathbb{E}_t[\hat{V}_{t+1}]), \quad (10)$$

which has no hat on the  $c_t^*$  since it is a direct function of  $\beta$  rather than  $\hat{\beta}$ .

As of time  $t$ , we can then write the continuation value for period  $t+1$  as follows (where  $\beta$  and  $\hat{\beta}$  do not enter because the future payoffs are already being discounted by  $\beta$  by the present self at time  $t' \leq t$ ):

$$\hat{V}_{t+1} = \begin{cases} \delta \alpha_{t+1} y - c_{t+1} & \text{if } c_{t+1} \leq \hat{c}_{t+1}^*, \\ \delta \hat{\rho}_{t+2} \mathbb{E}_{t+1}[\hat{V}_{t+2}] & \text{if } c_{t+1} > \hat{c}_{t+1}^*. \end{cases} \quad (11)$$

Taking expectations of equation (11) over cost draws,

$$\mathbb{E}_t[\hat{V}_{t+1}] = \int_{\underline{c}}^{\hat{c}_{t+1}^*} (\delta \alpha_{t+1} y - c) dF(c) + \int_{\hat{c}_{t+1}^*}^{\bar{c}} \delta \hat{\rho}_{t+2} \mathbb{E}_{t+1}[\hat{V}_{t+2}] dF(c), \quad (12)$$

which is equivalent to equation (4) by the fundamental theorem of calculus.  $\square$

The following lemma will be useful in many of the subsequent proofs. For ease of notation, from now on we assume a constant  $\alpha$  and  $\hat{\rho}$  over time, but allow  $\rho_t$  to vary over time as the probability of remembering in a particular period will be affected by receiving a reminder in that period when we turn to the model predictions.

**Lemma 1** (Monotonicity of Continuation Values). *For a sufficiently large  $\hat{\beta}$ , a higher continuation value in period  $t + 1$  implies a higher continuation value in period  $t$ .*

*Proof.* Taking the partial derivative of equation (4) with respect to  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$ ,

$$\begin{aligned} \frac{\partial \mathbb{E}_t[\hat{V}_{t+1}]}{\partial \mathbb{E}_{t+1}[\hat{V}_{t+2}]} &= (\delta \alpha y - \hat{c}_{t+1}^*) f(\hat{c}_{t+1}^*) \frac{d\hat{c}_{t+1}^*}{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]} - f(\hat{c}_{t+1}^*) \frac{d\hat{c}_{t+1}^*}{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]} \delta \hat{\rho} \mathbb{E}_{t+1}[\hat{V}_{t+2}] \\ &\quad + (1 - F(\hat{c}_{t+1}^*)) \delta \hat{\rho}. \end{aligned} \quad (13)$$

Next, taking derivatives of equation (9) for  $\tau = t + 1$  with respect to  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$ , we have that  $d\hat{c}_{t+1}^*/d\mathbb{E}_{t+1}[\hat{V}_{t+2}] = -\beta \delta \hat{\rho}$ . Rearranging equation (9) with  $\tau = t + 1$  to express  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$  as a function of  $\hat{c}_{t+1}^*$ , then plugging this and  $d\hat{c}_{t+1}^*/d\mathbb{E}_{t+1}[\hat{V}_{t+2}] = -\beta \delta \hat{\rho}$  into the right-hand side of equation (13) and rearranging terms:

$$\frac{\partial \mathbb{E}_t[\hat{V}_{t+1}]}{\partial \mathbb{E}_{t+1}[\hat{V}_{t+2}]} = \left[ (1 - F(\hat{c}_{t+1}^*)) - (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*) \right] \delta \hat{\rho}. \quad (14)$$

Thus  $\partial \mathbb{E}_t[\hat{V}_{t+1}]/\partial \mathbb{E}_{t+1}[\hat{V}_{t+2}] \geq 0$  if  $\kappa \equiv (1 - F(\hat{c}_{t+1}^*)) - (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*) \geq 0$ . This holds for  $\hat{\beta} = 1$ , and by continuity it also holds for sufficiently large  $\hat{\beta} < 1$ .  $\square$

**Proposition 2** (Existence and Uniqueness). *An equilibrium of the model uniquely exists (for  $\hat{\beta}$  sufficiently large that monotonicity of continuation values holds).*

*Proof.* By Proposition 1, an optimal strategy of the future self does not deviate from an optimal threshold strategy at a positive probability, and thus it suffices to only consider the threshold strategies for the perceived actions of future selves. Now, the claim to be proven is that an optimal threshold strategy, characterized by the sequence of optimal thresholds  $\{c_t^*\}$  with  $\hat{c}_t^* \in [\underline{c}, \bar{c}]$  for each  $t$  uniquely exists for each future self.

We first construct such a sequence of thresholds by approximating the given no-deadline problem with problems of sufficiently long deadlines and then prove that the corresponding actions satisfy the given dynamic and satisfy it uniquely. Formally, denote  $\{\mathbb{E}_t[\hat{V}_{t+1}^T]\}_{t \leq T}$  as the sequence of continuation values defined by equation (12) when the deadline is at date  $T$ , with the terminal



condition that  $\hat{\rho}_{T+1}\mathbb{E}_T[\hat{V}_{T+1}^T] = 0$ . Denote the optimal cost thresholds induced by these values as  $\hat{c}_t^T$  and  $c_t^T$ ; the superscripts denote that the deadline is at date  $T$ .

As  $T \rightarrow \infty$ , the continuation values  $\mathbb{E}_t[V_{t+1}^T]$  converge. This is because for a given date  $t$  and sufficiently large  $T_1 < T_2$ , it holds that

$$\left| \mathbb{E}_t[\hat{V}_{t+1}^{T_1}] - \mathbb{E}_t[\hat{V}_{t+1}^{T_2}] \right| \leq \delta \left| \mathbb{E}_{t+1}[\hat{V}_{t+2}^{T_1}] - \mathbb{E}_{t+1}[\hat{V}_{t+2}^{T_2}] \right| \leq \delta^{T_1-t-1} \left| \mathbb{E}_{T_1}[\hat{V}_{T_1+1}^{T_2}] \right|,$$

which can be arbitrarily small since the last expectation is bounded by  $\delta y$ . Note that the first inequality is based on the assumption that  $\hat{\beta}$  is sufficiently large that monotonicity of continuation values holds (Lemma 1), so that

$$0 \leq \frac{\partial \mathbb{E}_t[\hat{V}_{t+1}]}{\partial \mathbb{E}_{t+1}[\hat{V}_{t+2}]} = \left( 1 - F(\hat{c}_{t+1}^*) - (1 - \hat{\beta})\hat{c}_{t+1}^* f(\hat{c}_{t+1}^*) \right) \delta \hat{\rho}_{t+2} \leq \delta.$$

Now, by Cauchy's sequence criterion, the sequence  $\{\mathbb{E}_t[\hat{V}_{t+1}^T]\}$  converges, and so do the sequences  $\{\hat{c}_t^T\}$  and  $\{c_t^T\}$ . Therefore, we can take their limits as  $\{\mathbb{E}_t[\hat{V}_{t+1}]\}$ ,  $\{\hat{c}_t^*\}$ , and  $\{c_t^*\}$ .

Note that both sides of equations (9) and (12) are continuous functions. Therefore, it follows that  $\{\mathbb{E}_t[\hat{V}_{t+1}]\}$ ,  $\{\hat{c}_t^*\}$ , and  $\{c_t^*\}$  satisfy (9) and (12) by taking the limits of each side.  $\square$

**Corollary 1.** *As the deadline increases to infinity, the equilibrium thresholds will converge to the no-deadline case.*

This corollary implies that we can approximate the no-deadline case by a deadline case with a sufficiently long deadline.

## A.2 Proofs of Model Predictions

A sufficiently large  $\hat{\beta}$  such that monotonicity of continuation values holds (Lemma 1) is a condition for all predictions except Prediction 2. For conciseness we do not write out this condition in the statement of each prediction.

We also introduce a second lemma that will be used in some of the proofs.

**Lemma 2** (Optimal Threshold and Take-Up). *A higher optimal threshold  $c_t^*$  in period  $t$  increases the (cumulative) take-up by any date  $\tau \geq t$ .*

*Proof.* The cumulative take-up by any date  $\tau \geq t$  is

$$P(\{c_t^*\}_{t \leq \tau}) = \sum_{t=1}^{\tau} F(c_t^*) \prod_{j=1}^t \rho_j (1 - F(c_{j-1}^*)).$$

Our goal is to show that this cumulative take-up is increasing in the optimal threshold  $c_t^*$  of any

date  $t \leq \tau$ .

First, note that  $F(c_t^*)$  measures conditional take-up at date  $t$ , i.e., the proportion of those who have not yet adopted by  $t - 1$  who adopt at date  $t$ . For a larger  $c_t^*$  in period  $t \leq \tau$ , conditional take-up  $F(c_t^*)$  is larger, which follows because the distribution  $F(\cdot)$  is weakly increasing. Next, we show that cumulative take-up is increasing in conditional take-up by examining the partial derivative:

$$\begin{aligned}
\frac{\partial P}{\partial F(c_t^*)} &= \prod_{j=1}^t \rho_j (1 - F(c_{j-1}^*)) \cdot \left( 1 - \rho_{t+1} \sum_{\tau'=t+1}^{\tau} F(c_{\tau'}^*) \prod_{j=t+2}^{\tau'} (1 - F(c_{j-1}^*)) \rho_j \right) \\
&\geq \prod_{j=1}^t \rho_j (1 - F(c_{j-1}^*)) \cdot \left( 1 - \left[ \sum_{\tau'=t+1}^{\tau} F(c_{\tau'}^*) \prod_{j=t+2}^{\tau'} (1 - F(c_{j-1}^*)) \right] \right) \\
&= \prod_{j=1}^t \rho_j (1 - F(c_{j-1}^*)) \cdot \prod_{j=t+1}^{\tau} (1 - F(c_j^*)) > 0.
\end{aligned} \tag{15}$$

This completes the proof.  $\square$

Intuitively, define conditional take-up as above and unconditional take-up as the proportion of all managers who adopt at date  $t$  (not cumulative). Then the increase in  $c_t^*$  leads to an increase in conditional take-up at date  $t$ , which increases the unconditional take-up at date  $t$  but decreases the unconditional (non-cumulative) take-up in the future dates after  $t$ . Overall, however, the former effect always dominates the latter effect because the cumulative take-up rate between date  $t + 1$  and date  $\tau$ , conditional on not having adopted by date  $t$ —i.e., the term in square brackets in (15)—is always lower than one.

We now turn to the model predictions.

**Prediction 1** (Offer Value and Take-Up). *A higher value of the offer (higher  $y$ ) increases take-up.*

*Proof.* By Lemma 2, cumulative take-up is higher if the optimal cost threshold is higher in each period, so we seek to prove that  $dc_t^*/dy > 0$ .

Taking derivatives of equation (3) with respect to  $y$ ,

$$\frac{dc_t^*}{dy} = \beta \delta \left( \alpha - \hat{\rho} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{dy} \right). \tag{16}$$

Thus a sufficient condition for  $dc_t^*/dy > 0$  is that  $d\mathbb{E}_t[\hat{V}_{t+1}]/dy < \alpha/\hat{\rho}$ , which is what we will show.

Taking derivatives of equation (4) with respect to  $y$ ,

$$\begin{aligned} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{dy} &= (\delta\alpha y - \hat{c}_{t+1}^*) f(\hat{c}_{t+1}^*) \frac{d\hat{c}_{t+1}^*}{dy} + \delta\alpha F(\hat{c}_{t+1}^*) \\ &\quad - f(\hat{c}_{t+1}^*) \frac{d\hat{c}_{t+1}^*}{dy} \delta\hat{\rho} \mathbb{E}_{t+1}[\hat{V}_{t+2}] + (1 - F(\hat{c}_{t+1}^*)) \delta\hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy}. \end{aligned} \quad (17)$$

Rearranging equation (9) with  $\tau = t + 1$  to express  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$  as a function of  $\hat{c}_{t+1}^*$ , plugging this into equation (17), and rearranging terms:

$$\frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{dy} = \left(\hat{\beta}^{-1} - 1\right) \frac{d\hat{c}_{t+1}^*}{dy} f(\hat{c}_{t+1}^*) \hat{c}_{t+1}^* + \delta\alpha F(\hat{c}_{t+1}^*) + (1 - F(\hat{c}_{t+1}^*)) \delta\hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy}. \quad (18)$$

Taking derivatives of equation (9) for  $\tau = t + 1$  with respect to  $y$ , and plugging  $d\hat{c}_{t+1}^*/dy$  into equation (18):

$$\begin{aligned} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{dy} &= \left(\hat{\beta}^{-1} - 1\right) \left(\hat{\beta} \delta \left(\alpha - \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy}\right)\right) f(\hat{c}_{t+1}^*) \hat{c}_{t+1}^* \\ &\quad + \delta\alpha F(\hat{c}_{t+1}^*) + (1 - F(\hat{c}_{t+1}^*)) \delta\hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy} \\ &= \left(F(\hat{c}_{t+1}^*) + (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*)\right) \delta\alpha \\ &\quad + \left(1 - F(\hat{c}_{t+1}^*) - (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*)\right) \delta\hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy} \\ &= (1 - \kappa) \delta\alpha + \kappa \delta\hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy}, \end{aligned} \quad (19)$$

where  $\kappa \equiv 1 - F(\hat{c}_{t+1}^*) - (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*)$ . Note that  $\kappa \leq 1$  always, and for sufficiently large  $\hat{\beta}$  such that monotonicity of continuation values holds (Lemma 1),  $0 \leq \kappa \leq 1$ . At the deadline  $t = T$ ,  $d\mathbb{E}_T[\hat{V}_{T+1}]/dy = 0$ , as there is no continuation value since the task expires. Plugging this into the right side of equation (19) for  $t = T - 1$ ,  $d\mathbb{E}_{T-1}[\hat{V}_T]/dy = (1 - \kappa) \delta\alpha \leq \delta\alpha$ . By induction, we can prove that  $d\mathbb{E}_t[\hat{V}_{t+1}]/dy \leq \delta\alpha$ :

$$\begin{aligned} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{dy} &= (1 - \kappa) \delta\alpha + \kappa \delta\hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{dy} \\ &\leq (1 - \kappa) \delta\alpha + \kappa \delta\hat{\rho} \delta\alpha \\ &\leq (1 - \kappa) \delta\alpha + \kappa \delta\alpha = \delta\alpha. \end{aligned}$$

Thus we have:

$$\hat{\rho} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{dy} \leq \hat{\rho} \delta \alpha \leq \delta \alpha \leq \alpha,$$

so  $d\mathbb{E}_t[\hat{V}_{t+1}]/dy \leq \alpha/\hat{\rho}$ , with strict inequality as long as at least one of  $\delta < 1$ ,  $\hat{\rho} < 1$ , or  $\kappa > 0$  holds strictly.  $\square$

**Prediction 2** (Reminder and Memory). *If a reminder (unannounced or announced) increases the probability of remembering, it will also increase take-up.*

*Proof.* From equation (3),  $c_t^*$  only depends on  $\hat{\rho}_\ell$  for  $\ell > t$  and not on  $\rho_t$ . Hence  $\rho_t$  impacts the probability of adopting at time  $t$  only through its impact on the probability of the task being active in period  $t$ . For any period  $t$ ,

$$\Pr(\text{task active at } t) = \prod_{j=1}^t \rho_j \prod_{k=1}^{t-1} (1 - F(c_k^*)),$$

so  $d\Pr(\text{task active at } t)/d\rho_t > 0$ . By assumption, an unannounced reminder received in period  $t$  increases  $\rho_t$  and therefore increase the probability of the task being active and the probability of adopting.  $\square$

**Prediction 3** (Announced Reminder and Beliefs about Memory). *The announced reminder (a) reduces take-up at  $t = 1$ , compared to the unannounced reminder, if managers do not believe they have perfect memory, ( $\hat{\rho}_t < 1$  for  $t > 1$ ), and (b) has no differential effect on take-up at  $t = 1$  if managers believe they have perfect memory ( $\hat{\rho}_t = 1$  for  $t > 1$ ).*

*Proof.* Take-up at  $t = 1$  will be lower if the optimal cost threshold is lower, so we seek to show that  $dc_t^*/d\hat{\rho} \leq 0$ . Taking derivatives of equation (3) with respect to  $\hat{\rho}$ ,

$$\frac{dc_t^*}{d\hat{\rho}} = -\beta \delta \mathbb{E}_t[\hat{V}_{t+1}] - \beta \delta \hat{\rho} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{d\hat{\rho}}. \quad (20)$$

Thus a sufficient condition for  $dc_t^*/d\hat{\rho} < 0$  is  $d\mathbb{E}_t[\hat{V}_{t+1}]/d\hat{\rho} > 0$ .

Taking derivatives of equation (4) with respect to  $\hat{\rho}$ ,

$$\begin{aligned} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{d\hat{\rho}} &= (\delta \alpha y - \hat{c}_{t+1}^*) f(\hat{c}_{t+1}^*) \frac{d\hat{c}_{t+1}^*}{d\hat{\rho}} - f(\hat{c}_{t+1}^*) \frac{d\hat{c}_{t+1}^*}{d\hat{\rho}} \delta \hat{\rho} \mathbb{E}_{t+1}[\hat{V}_{t+2}] \\ &\quad + (1 - F(\hat{c}_{t+1}^*)) \delta \mathbb{E}_{t+1}[\hat{V}_{t+2}] + (1 - F(\hat{c}_{t+1}^*)) \delta \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{d\hat{\rho}}. \end{aligned} \quad (21)$$

Rearranging equation (9) with  $\tau = t + 1$  to express  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$  as a function of  $\hat{c}_{t+1}^*$  and plugging

this into the second term on the right-hand side of (21) (but leaving  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$  in the third term), and rearranging terms:

$$\begin{aligned} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{d\hat{\rho}} &= \left(\hat{\beta}^{-1} - 1\right) \frac{d\hat{c}_{t+1}^*}{d\hat{\rho}} \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*) \\ &\quad + \left(1 - F(\hat{c}_{t+1}^*)\right) \delta \mathbb{E}_{t+1}[\hat{V}_{t+2}] + \left(1 - F(\hat{c}_{t+1}^*)\right) \delta \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{d\hat{\rho}}. \end{aligned} \quad (22)$$

Taking derivatives of equation (9) for  $\tau = t + 1$  with respect to  $\hat{\rho}$ , plugging in  $d\hat{c}_{t+1}^*/d\hat{\rho}$  in equation (22), and rearranging terms:

$$\begin{aligned} \frac{d\mathbb{E}_t[\hat{V}_{t+1}]}{d\hat{\rho}} &= \left(\hat{\beta}^{-1} - 1\right) \left[ -\hat{\beta} \delta \left( \mathbb{E}_{t+1}[\hat{V}_{t+2}] + \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{d\hat{\rho}} \right) \right] \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*) \\ &\quad + \left(1 - F(\hat{c}_{t+1}^*)\right) \delta \mathbb{E}_{t+1}[\hat{V}_{t+2}] + \left(1 - F(\hat{c}_{t+1}^*)\right) \delta \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{d\hat{\rho}} \\ &= \left(1 - F(\hat{c}_{t+1}^*) - (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*)\right) \delta \left( \mathbb{E}_{t+1}[\hat{V}_{t+2}] + \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{d\hat{\rho}} \right) \\ &= \kappa \delta \left( \mathbb{E}_{t+1}[\hat{V}_{t+2}] + \hat{\rho} \frac{d\mathbb{E}_{t+1}[\hat{V}_{t+2}]}{d\hat{\rho}} \right), \end{aligned} \quad (23)$$

where  $\kappa \equiv 1 - F(\hat{c}_{t+1}^*) - (1 - \hat{\beta}) \hat{c}_{t+1}^* f(\hat{c}_{t+1}^*)$ . Note that  $\kappa \leq 1$  always, and for sufficiently large  $\hat{\beta}$  such that monotonicity of continuation values holds (Lemma 1),  $0 \leq \kappa \leq 1$ . At the deadline  $t = T$ ,  $d\mathbb{E}_T[\hat{V}_{T+1}]/d\hat{\rho} = 0$  as there is no continuation value since the task expires. Plugging this into the right-hand side of equation (23) for  $t = T - 1$ , the second term inside the parentheses will equal 0. The first term inside the parentheses is positive, so  $d\mathbb{E}_{T-1}[\hat{V}_T]/d\hat{\rho} \geq 0$ . Recursively, for all  $t < T$ ,  $d\mathbb{E}_t[\hat{V}_{t+1}]/d\hat{\rho} \geq 0$  from equation (23), which is what we sought to prove.

Announced reminders increase  $\hat{\rho}_\tau$  for the period  $\tau$  in which the manager is told the reminder will arrive, as long as  $\hat{\rho}_\tau < 1$ . On the other hand, if  $\hat{\rho}_\tau = 1$  even in the absence of the announced reminder,  $\hat{\rho}_\tau$  cannot be increased. Thus, (a) if managers do not believe they have perfect memory, the announced reminder lowers the cost threshold  $c_t^*$  since  $dc_t^*/d\hat{\rho} < 0$ , and (b) if managers believe they have perfect memory,  $\hat{\rho}_\tau$  does not increase and hence there is no change in  $c_t^*$  or take-up.  $\square$

**Prediction 4** (Announced Reminder and Trust). *The announced reminder (a) does not affect final take-up, compared to the unannounced reminder, if firms inherently trust the offer ( $\alpha_t = 1$ ); and (b) increases final take-up if some firms distrust the offer ( $\alpha_t < 1$ ) and their trust increases after receiving the announced reminder.*

*Proof.* By Lemma 2, in order to show that the announced reminder increases final (cumulative) take-up, it suffices to prove that  $dc_t^*/d\alpha > 0$ . The proof that  $dc_t^*/d\alpha > 0$  is analogous to the proof from Prediction 1 that  $dc_t^*/dy > 0$ , given the way that  $\alpha_t y$  enters equations (3) and (4). If (a)  $\alpha_t = 1$

inherently, then  $\alpha_t$  cannot increase from receiving the announced reminder, and hence there is no change in  $c_t^*$  and no difference in take-up between receiving an announced or an unannounced reminder. If (b)  $\alpha_t < 1$  and the announced reminder increases  $\alpha_t$  by increasing trust in the offer, then the threshold  $c_t^*$  increases. The higher optimal cost threshold after receiving the reminder leads to an increase in take-up.  $\square$

We next turn to predictions about deadlines. In a slight abuse of notation, we will now define  $T$  as the date after which firms in the deadline group can no longer adopt, whereas firms in the no-deadline group can still adopt when  $t > T$ .

**Prediction 5** (Deadline and Offer Value). *A higher value of the offer (higher  $y$ ) implies a lower treatment effect of a deadline on adoption by the deadline (for  $y$  sufficiently large that the manager without a deadline still adopts with positive probability, and  $\rho_t$  and  $T$  sufficiently large).*

*Proof.* Denote  $\mathbb{E}_{t-1}[\hat{V}_t]$  by  $v_t$ . Equation (12) can then be written as

$$v_t = \int_{\underline{c}}^{\hat{c}_t^*} (\delta\alpha y - c) dF(c) + \int_{\hat{c}_t^*}^{\bar{c}} \delta\hat{\rho} v_{t+1} dF(c),$$

where

$$\hat{c}_t^* = \hat{\beta} \delta (\alpha y - \hat{\rho} v_{t+1}).$$

Assume that  $\rho_t = 1$ ; for cases with  $\rho_t$  sufficiently large (i.e., sufficiently close to 1), the following arguments hold by continuity in  $\rho_t$ . In this special case with  $\rho_t = 1$ , the cumulative adoption rate by period  $T$  is

$$P(\{c_t^*\}_{t=1}^T) = 1 - \prod_{t=1}^T (1 - F(c_t^*)),$$

where

$$c_t^* = \beta \delta (\alpha y - \hat{\rho} v_{t+1}).$$

We now introduce the notation of a tilde indicating the no-deadline case. When there is a deadline at date  $T$ ,  $v_{T+1} = 0$ , while in general, without a deadline,  $\tilde{v}_{T+1} > 0$ . The deadline effect can be written as

$$P|_{v_{T+1}=0} - P|_{v_{T+1}=\tilde{v}_{T+1}} = - \int_0^{\tilde{v}_{T+1}} \frac{\partial P}{\partial v_{T+1}} dv_{T+1}.$$

We are thus interested in the sign of

$$\frac{\partial}{\partial y} (P|_{v_{T+1}=0} - P|_{v_{T+1}=\tilde{v}_{T+1}}) = - \int_0^{\tilde{v}_{T+1}} \frac{\partial^2 P}{\partial y \partial v_{T+1}} dv_{T+1}. \quad (24)$$

To this aim, we seek to determine the sign of  $\partial^2 P / \partial y \partial v_{T+1}$ . Define  $R \equiv \ln(1 - P)$ , or equivalently  $P = 1 - e^R$ . Then,

$$\begin{aligned} \frac{\partial^2 P}{\partial y \partial v_{T+1}} &= \frac{\partial}{\partial y} \left( \frac{\partial P}{\partial v_{T+1}} \right) \\ &= \frac{\partial}{\partial y} \left( -e^R \frac{\partial R}{\partial v_{T+1}} \right) \\ &= -e^R \left( \frac{\partial R}{\partial y} \frac{\partial R}{\partial v_{T+1}} + \frac{\partial^2 R}{\partial y \partial v_{T+1}} \right). \end{aligned} \quad (25)$$

Next, define  $\phi(c) \equiv \ln(1 - F(c))$ . We have that

$$R = \sum_{t=1}^T \ln(1 - F(c_t^*)) = \sum_{t=1}^T \phi(c_t^*).$$

Thus,

$$\begin{aligned} \frac{\partial R}{\partial y} &= \sum_{t=1}^T \phi'(c_t^*) \frac{\partial c_t^*}{\partial y}, \\ \frac{\partial R}{\partial v_{T+1}} &= \sum_{t=1}^T \phi'(c_t^*) \frac{\partial c_t^*}{\partial v_{T+1}}, \\ \frac{\partial^2 R}{\partial y \partial v_{T+1}} &= \sum_{t=1}^T \left( \phi''(c_t^*) \frac{\partial c_t^*}{\partial y} \frac{\partial c_t^*}{\partial v_{T+1}} + \phi'(c_t^*) \frac{\partial^2 c_t^*}{\partial y \partial v_{T+1}} \right), \end{aligned}$$

where

$$\begin{aligned} \phi'(c) &= -\frac{f(c)}{1 - F(c)}, \\ \phi''(c) &= \phi'(c) \left( \frac{f'(c)}{f(c)} + \frac{f(c)}{1 - F(c)} \right) \end{aligned}$$

from the definition of  $\phi(c)$ . Define  $\lambda(c) \equiv F(c) + (1 - \hat{\beta})cf(c)$ . Then, we have the following

derivatives:

$$\begin{aligned}
\frac{\partial v_t}{\partial v_{t+1}} &= (\delta\alpha y - \hat{c}_t^*) f(\hat{c}_t^*) \left(-\hat{\beta} \delta\hat{\rho}\right) - \hat{\rho} \delta v_{t+1} f(\hat{c}_t^*) \left(-\hat{\beta} \delta\hat{\rho}\right) + (1 - F(\hat{c}_t^*)) \delta\hat{\rho}_{t+1} \\
&= \left(1 - F(\hat{c}_t^*) - (1 - \hat{\beta}) \hat{c}_t^* f(\hat{c}_t^*)\right) \delta\hat{\rho} \\
&= (1 - \lambda(\hat{c}_t^*)) \delta\hat{\rho}, \\
\frac{\partial v_t}{\partial y} &= (\delta\alpha y - \hat{c}_t^*) f(\hat{c}_t^*) \hat{\beta} \delta\alpha - \hat{\rho} v_{t+1} f(\hat{c}_t^*) \hat{\beta} \delta\alpha + \delta\alpha F(\hat{c}_t^*) \\
&= \left(F(\hat{c}_t^*) + (1 - \hat{\beta}) \hat{c}_t^* f(\hat{c}_t^*)\right) \delta\alpha \\
&= \lambda(\hat{c}_t^*) \delta\alpha.
\end{aligned}$$

Therefore, recursively we can write that

$$\begin{aligned}
dv_t &= \frac{\partial v_t}{\partial v_{t+1}} dv_{t+1} + \frac{\partial v_t}{\partial y} dy \\
&= (1 - \lambda(\hat{c}_t^*)) \delta\hat{\rho} dv_{t+1} + \lambda(\hat{c}_t^*) \delta\alpha dy \\
&= (1 - \lambda(\hat{c}_t^*)) \delta\hat{\rho} (1 - \lambda(\hat{c}_{t+1}^*)) \delta\hat{\rho} dv_{t+2} + (1 - \lambda(\hat{c}_t^*)) \delta\hat{\rho} \lambda(\hat{c}_{t+1}^*) \delta\alpha dy + \lambda(\hat{c}_t^*) \delta\alpha dy \\
&= \left( (\delta\hat{\rho})^{T-t+1} \prod_{s=t}^T (1 - \lambda(\hat{c}_s^*)) \right) dv_{T+1} + \delta\alpha \sum_{s=t}^T \left( (\delta\hat{\rho})^{s-t} \prod_{u=t}^{s-1} (1 - \lambda(\hat{c}_u^*)) \right) \lambda(\hat{c}_s^*) dy.
\end{aligned}$$

Consequently, the partial derivatives of  $v_t$  with respect to  $v_{T+1}$  and  $y$  are

$$\begin{aligned}
\frac{\partial v_t}{\partial v_{T+1}} &= (\delta\hat{\rho})^{T-t+1} \prod_{s=t}^T (1 - \lambda(\hat{c}_s^*)), \\
\frac{\partial v_t}{\partial y} &= \delta\alpha \sum_{s=t}^T \left( (\delta\hat{\rho})^{s-t} \prod_{u=t}^{s-1} (1 - \lambda(\hat{c}_u^*)) \right) \lambda(\hat{c}_s^*).
\end{aligned}$$

The cross-partial derivative of interest can then be found using logarithmic differentiation as

$$\begin{aligned}
\frac{\partial^2 v_t}{\partial y \partial v_{T+1}} &= \frac{\partial}{\partial y} \left( \frac{\partial v_t}{\partial v_{T+1}} \right) \\
&= \frac{\partial v_t}{\partial v_{T+1}} \sum_{s=t}^T \frac{\partial}{\partial y} \ln(1 - \lambda(\hat{c}_s^*)) \\
&= \frac{\partial v_t}{\partial v_{T+1}} \sum_{s=t}^T \frac{-\lambda'(\hat{c}_s^*)}{1 - \lambda(\hat{c}_s^*)} \frac{\partial \hat{c}_s^*}{\partial y}.
\end{aligned}$$



Also note that

$$\begin{aligned}\frac{\partial c_t^*}{\partial y} &= \beta \delta \alpha - \beta \delta \hat{\rho} \frac{\partial v_{t+1}}{\partial y}, \\ \frac{\partial c_t^*}{\partial v_{T+1}} &= -\beta \delta \hat{\rho} \frac{\partial v_{t+1}}{\partial v_{T+1}}, \\ \frac{\partial^2 c_t^*}{\partial y \partial v_{T+1}} &= -\beta \delta \hat{\rho} \frac{\partial^2 v_{t+1}}{\partial y \partial v_{T+1}}.\end{aligned}$$

Plugging in the partial derivatives of  $v_{t+1}$  leads to

$$\begin{aligned}\frac{\partial c_t^*}{\partial y} &= \beta \delta \alpha \left( 1 - \delta \hat{\rho} \sum_{s=t+1}^T \left( (\delta \hat{\rho})^{s-t-1} \prod_{u=t+1}^{s-1} (1 - \lambda(\hat{c}_u^*)) \right) \lambda(\hat{c}_s^*) \right), \\ \frac{\partial c_t^*}{\partial v_{T+1}} &= -\beta (\delta \hat{\rho})^{T-t+1} \prod_{s=t+1}^T (1 - \lambda(\hat{c}_s^*)), \\ \frac{\partial^2 c_t^*}{\partial y \partial v_{T+1}} &= -\beta \delta \hat{\rho} \frac{\partial v_{t+1}}{\partial v_{T+1}} \sum_{s=t+1}^T \frac{-\lambda'(\hat{c}_s^*)}{1 - \lambda(\hat{c}_s^*)} \frac{\partial \hat{c}_s^*}{\partial y} \\ &= \frac{\partial c_t^*}{\partial v_{T+1}} \sum_{s=t+1}^T \frac{-\lambda'(\hat{c}_s^*)}{1 - \lambda(\hat{c}_s^*)} \frac{\partial \hat{c}_s^*}{\partial y}.\end{aligned}$$

Thus, we find that

$$\begin{aligned}\frac{\partial^2 R}{\partial y \partial v_{T+1}} &= \sum_{t=1}^T \phi''(c_t^*) \frac{\partial c_t^*}{\partial y} \frac{\partial c_t^*}{\partial v_{T+1}} + \phi'(c_t^*) \frac{\partial^2 c_t^*}{\partial y \partial v_{T+1}} \\ &= \sum_{t=1}^T \phi''(c_t^*) \frac{\partial c_t^*}{\partial y} \frac{\partial c_t^*}{\partial v_{T+1}} + \phi'(c_t^*) \frac{\partial c_t^*}{\partial v_{T+1}} \sum_{s=t+1}^T \frac{-\lambda'(\hat{c}_s^*)}{1 - \lambda(\hat{c}_s^*)} \frac{\partial \hat{c}_s^*}{\partial y}.\end{aligned}$$

Next, we seek to control the size of this cross-partial derivative. The first step is to assume that there is some  $\underline{y}$  such that for every  $y \geq \underline{y}$  and  $v_{T+1} \in [0, \tilde{v}_{T+1}]$ , managers always adopt at a positive probability, i.e.,  $c_t \in [\underline{c}, \bar{c}] \subset (\underline{c}, \bar{c})$ . Therefore,  $\lambda(c_t) \in [\underline{\lambda}, \bar{\lambda}]$ . For sufficiently large  $\hat{\beta}$ , it also holds that  $\bar{\lambda} < 1$ . To find the lower bound of  $\partial c_t^* / \partial y$ , note that

$$\begin{aligned}& 1 - \sum_{s=t+1}^T \left( (\delta \hat{\rho})^{s-t-1} \prod_{u=t+1}^{s-1} (1 - \lambda(\hat{c}_u^*)) \right) \lambda(\hat{c}_s^*) \\ & \geq 1 - \sum_{s=t+1}^T \left( \prod_{u=t+1}^{s-1} (1 - \lambda(\hat{c}_u^*)) \right) \lambda(\hat{c}_s^*) \\ & = \prod_{s=t+1}^T (1 - \lambda(\hat{c}_s^*)) > 0.\end{aligned}$$

Therefore, for each  $t > 0$ ,

$$\frac{\partial c_t^*}{\partial y} \geq \beta \delta \alpha (1 - \delta \hat{\rho}) \equiv \underline{\omega} > 0.$$

The upper bound can be obtained as

$$\frac{\partial c_t^*}{\partial y} \leq \beta \delta \alpha \equiv \bar{\omega} < 1.$$

We conclude that  $\partial c_t^* / \partial y \in [\underline{\omega}, \bar{\omega}]$ . On the other hand,

$$-\beta \delta \hat{\rho} (\delta \hat{\rho} (1 - \underline{\lambda}))^{T-t} \leq \frac{\partial c_t^*}{\partial v_{T+1}} \leq -\beta \delta \hat{\rho} (\delta \hat{\rho} (1 - \bar{\lambda}))^{T-t}.$$

Together these imply that

$$\begin{aligned} \left| \frac{\partial^2 R}{\partial y \partial v_{T+1}} \right| &\leq \sum_{t=1}^T |\phi''(c_t^*)| \left| \frac{\partial c_t^*}{\partial y} \right| \left| \frac{\partial c_t^*}{\partial v_{T+1}} \right| + |\phi'(c_t^*)| \left| \frac{\partial c_t^*}{\partial v_{T+1}} \right| \sum_{s=t+1}^T \frac{|\lambda'(\hat{c}_s^*)|}{|1 - \lambda(\hat{c}_s^*)|} \left| \frac{\partial \hat{c}_s^*}{\partial y} \right| \\ &\leq \sum_{t=1}^T M_2 \bar{\omega} \beta \delta \hat{\rho} (\delta \hat{\rho} (1 - \underline{\lambda}))^{T-t} + M_1 \beta \delta \hat{\rho} (\delta \hat{\rho} (1 - \underline{\lambda}))^{T-t} \sum_{s=t+1}^T \frac{M_\lambda}{1 - \bar{\lambda}} \bar{\omega} \\ &\leq M_2 \bar{\omega} \beta \delta \hat{\rho} \left( \sum_{t=1}^T (\delta \hat{\rho} (1 - \underline{\lambda}))^{T-t} \right) + M_1 \beta \delta \hat{\rho} \frac{M_\lambda}{1 - \bar{\lambda}} \bar{\omega} \left( \sum_{t=1}^T (T-t) (\delta \hat{\rho} (1 - \underline{\lambda}))^{T-t} \right) \\ &\leq \frac{M_2 \bar{\omega} \beta \delta \hat{\rho}}{1 - \delta \hat{\rho} (1 - \underline{\lambda})} + M_1 \beta \delta \hat{\rho} \frac{M_\lambda}{1 - \bar{\lambda}} \bar{\omega} \frac{\delta \hat{\rho} (1 - \underline{\lambda})}{(1 - \delta \hat{\rho} (1 - \underline{\lambda}))^2} \equiv M, \end{aligned}$$

where  $M_1 \equiv \max_{c \in [\underline{c}, \bar{c}]} |\phi'(c)|$ ,  $M_2 \equiv \max_{c \in [\underline{c}, \bar{c}]} |\phi''(c)|$ , and  $M_\lambda \equiv \max_{c \in [\underline{c}, \bar{c}]} |\lambda'(c)|$ . They exist since all the functions are continuous over the given compact interval. This result indicates that as  $T \rightarrow \infty$ , the cross-partial derivative is bounded from above. At the same time, if we define  $L \equiv \min_{c \in [\underline{c}, \bar{c}]} |\phi'(c)| > 0$ , then we have  $\phi(c_t^*) < -L$  and thus

$$\begin{aligned} \frac{\partial R}{\partial y} &= \sum_{t=1}^T \phi'(c_t^*) \frac{\partial c_t^*}{\partial y} \leq -\sum_{t=1}^T L \underline{\omega} = -TL \underline{\omega} < 0, \\ \frac{\partial R}{\partial v_{T+1}} &= \sum_{t=1}^T \phi'(c_t^*) \frac{\partial c_t^*}{\partial v_{T+1}} \geq \sum_{t=1}^T L \beta \delta \hat{\rho} (\delta \hat{\rho} (1 - \bar{\lambda}))^{T-t} \geq L \beta \delta \hat{\rho} > 0. \end{aligned}$$

Therefore,

$$\frac{\partial R}{\partial y} \frac{\partial R}{\partial v_{T+1}} + \frac{\partial^2 R}{\partial y \partial v_{T+1}} \leq -TL \underline{\omega} L \beta \delta \hat{\rho} + M.$$

For sufficiently large  $T$ , this is negative, so from equation (25),

$$\frac{\partial^2 P}{\partial y \partial v_{T+1}} > 0.$$

Thus the sign of  $\partial(P|_{v_{T+1}=0} - P|_{v_{T+1}=\tilde{v}_{T+1}})/\partial y$  in equation (24) is negative. In conclusion, the treatment effect of a deadline is decreasing in  $y$  under the conditions stated in the prediction.  $\square$

**Prediction 6** (Deadline and Take-Up). *The treatment effect of a deadline on cumulative take-up by any date up until the deadline is always positive if the deadline does not affect trust, memory, or perceived memory.*

*Proof.* If the deadline does not affect  $\alpha_t$ ,  $\rho_t$ , or  $\hat{\rho}_t$ , the only exogenous change caused by the deadline is a decrease in the perceived value of delay at the deadline date  $T$ , i.e., a decrease in  $\mathbb{E}_T[\hat{V}_{T+1}]$ . In the deadline group,  $\mathbb{E}_T[\hat{V}_{T+1}] = 0$ , as there is no possibility of adopting after date  $T$ , while in the no deadline group, the value of delay at date  $T$  is  $\mathbb{E}_T[\hat{V}_{T+1}] \geq 0$ .

For a sufficiently large  $\hat{\beta}$  such that monotonicity of continuation values holds (Lemma 1), this implies that the perceived value of delay  $\mathbb{E}_t[\hat{V}_{t+1}]$  at any date  $t \leq T$  is smaller in the deadline case than in the no deadline case. Formally, this can be shown by induction on  $t \leq T$ . Assume that  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$  is smaller for the deadline group than the no-deadline group in period  $t + 1$ . Then this must be true in period  $t$  by Lemma 1, which states that  $\mathbb{E}_t[\hat{V}_{t+1}]$  is an increasing function of  $\mathbb{E}_{t+1}[\hat{V}_{t+2}]$ .

The optimal threshold  $c_t^*$  in period  $t$ , as shown in (3), is a decreasing function in  $\mathbb{E}_t[\hat{V}_{t+1}]$ . Therefore, a smaller value of delay in period  $t$  translates to a higher optimal threshold in the same period. Finally, by Lemma 2, this implies a higher cumulative take-up, i.e., a positive treatment effect of the deadline on cumulative take-up.  $\square$

**Prediction 7** (Deadline and Present Bias). *The more present-biased a manager is, the larger is the treatment effect of a deadline on cumulative adoption by the deadline (for  $\beta$  sufficiently large that the manager without a deadline still adopts with positive probability, and  $\rho_t$  and  $T$  sufficiently large).*

*Proof.* Denote the optimal cost thresholds for the deadline and no-deadline groups as  $c_t^*$  and  $\tilde{c}_t^*$ , respectively. Then let  $\{c_t^*\}_{t=1}^T$  denote the optimal thresholds at each point in time up to the deadline period  $T$  in the deadline case and  $\{\tilde{c}_t^*\}_{t=1}^T$  denote those in the no-deadline case (still up to the deadline case's deadline in period  $T$ ). In the deadline period  $t = T$ ,  $\tilde{c}_T^* < c_T^*$  since  $\mathbb{E}_T[\hat{V}_{T+1}] = 0$  in the deadline group and  $\mathbb{E}_T[\hat{V}_{T+1}] > 0$  in the no-deadline group. It will also hold that  $\tilde{c}_t^* < c_t^*$  for all  $t$  if  $\hat{\beta}$  is sufficiently large that monotonicity of continuation values holds (Lemma 1).

Define the cumulative adoption probability at time  $T$  as

$$P(\{c_t^*\}_{t=1}^T) = \sum_{t=1}^T F(c_t^*) \prod_{j=1}^t \rho_j (1 - F(c_{j-1}^*)).$$

Formally, we then want to prove that the gap in cumulative adoption rates by time  $T$ ,

$$P(\{c_t^*\}_{t=1}^T) - P(\{\tilde{c}_t^*\}_{t=1}^T),$$

is weakly decreasing in  $\beta$  (since higher values of  $\beta$  indicate less present bias) over an interval such that  $0 < F(\tilde{c}_t^*) < F(c_t^*) < 1$ . We will show this for  $\rho \in [\rho_0, 1]$  and  $T > T_0$ , where  $\rho_0$  and  $T_0$  are constants.

The proposition to be proven is thus

$$\frac{\partial}{\partial \beta} (P(\{c_t^*\}_{t=1}^T) - P(\{\tilde{c}_t^*\}_{t=1}^T)) \leq 0.$$

First, define that

$$\gamma_t^* \equiv c_t^*/\beta = \delta(\alpha_t y - \hat{\rho} \mathbb{E}_t[\hat{V}_{t+1}]),$$

which is independent of  $\beta$  (as  $\mathbb{E}_t[\hat{V}_{t+1}]$  is a function of  $\hat{\beta}$ , not  $\beta$ ) and, similarly,  $\tilde{\gamma}_t^* \equiv \tilde{c}_t^*/\beta$ . With these definitions, we can write  $P(\{c_t^*\}_{t=1}^T)$  as  $\bar{P}(\beta, \{\gamma_t^*\}_{t=1}^T)$ . Moreover, since  $\tilde{c}_t^* < c_t^*$  for all  $t$ , it must hold that  $\tilde{\gamma}_t^* < \gamma_t^*$ , so the difference in cumulative take-up can be written as a path integral:

$$\frac{\partial}{\partial \beta} (P(\{c_t^*\}_{t=1}^T) - P(\{\tilde{c}_t^*\}_{t=1}^T)) = \sum_{t=1}^T \int_{\tilde{\gamma}_t^*}^{\gamma_t^*} \frac{\partial^2}{\partial \beta \partial \gamma_t} \bar{P}(\beta, \gamma_1^*, \dots, \gamma_{t-1}^*, \gamma, \tilde{\gamma}_{t+1}^*, \dots, \tilde{\gamma}_T^*) d\gamma_t.$$

Therefore, it suffices to show that

$$\frac{\partial^2 \bar{P}}{\partial \beta \partial \gamma_t} \leq 0$$

whenever  $0 < F(c_t) < 1$  for all  $t$ .

By continuity, it suffices to prove this for  $\rho_t = 1$  as it follows that it holds for sufficiently large  $\rho_t$ . When  $\rho_t = 1$ , we can rewrite the cumulative adoption rate by period  $T$  as

$$P(\{c_t^*\}_{t=1}^T) = 1 - \prod_{j=1}^T (1 - F(c_j^*)). \quad (26)$$

We want to calculate  $\partial^2 \bar{P} / \partial \beta \partial \gamma_t$ . First, we have that

$$\frac{\partial \bar{P}}{\partial \gamma_t} = \beta f(c_t^*) \prod_{j \neq t} (1 - F(c_j^*)), \quad (27)$$

since the only factor in the second term on the right-hand side of equation (26) that contains  $\gamma_t$  is  $(1 - F(c_t^*))$ , and since  $\partial c_t^*/\partial \gamma_t = \beta$ . Next, taking logarithms of both sides of equation (27), then taking the partial derivative with respect to  $\beta$ , rearranging, and plugging back in  $\partial \bar{P}/\partial \gamma_t$  for the term on the right-hand side of equation (27),

$$\begin{aligned}
\frac{\partial^2 \bar{P}}{\partial \beta \partial \gamma_t} &= \frac{\partial}{\partial \beta} \left( \frac{\partial \bar{P}}{\partial \gamma_t} \right) = \frac{\partial \bar{P}}{\partial \gamma_t} \frac{\partial}{\partial \beta} \left( \ln \frac{\partial \bar{P}}{\partial \gamma_t} \right) \\
&= \frac{\partial \bar{P}}{\partial \gamma_t} \frac{\partial}{\partial \beta} \left( \ln \beta + \ln f(c_t^*) + \sum_{j \neq t} \ln(1 - F(c_j^*)) \right) \\
&= \frac{\partial \bar{P}}{\partial \gamma_t} \left( \frac{1}{\beta} + \frac{c_t^* f'(c_t^*)}{\beta f(c_t^*)} - \sum_{j \neq t} \frac{c_j^* f(c_j^*)}{\beta (1 - F(c_j^*))} \right) \\
&= f(c_t^*) \prod_{j \neq t} (1 - F(c_j^*)) \left( 1 + c_t^* \frac{f'(c_t^*)}{f(c_t^*)} - \sum_{j \neq t} \frac{c_j^* f(c_j^*)}{1 - F(c_j^*)} \right) \\
&= - \left( \prod_{j=1}^T (1 - F(c_j^*)) \right) \frac{f(c_t^*)}{1 - F(c_t^*)} \left( \sum_{j \neq t} \frac{c_j^* f(c_j^*)}{1 - F(c_j^*)} - 1 - c_t^* \frac{f'(c_t^*)}{f(c_t^*)} \right), \quad (28)
\end{aligned}$$

where we have used the fact that  $\partial c_t^*/\partial \beta = c_t^*/\beta = \gamma_t^*$  because  $c_t^*$  is linear in  $\beta$ .

The right hand side of (28) has three components and a preceding negative sign. The first two components are positive, so a sufficient condition for  $\partial^2 \bar{P}/\partial \beta \partial \gamma_t \leq 0$  is that the third component is positive. Since  $F(c)$  has a continuously differentiable positive density function over its support  $[\underline{c}, \bar{c}]$  and  $0 < F(c_j) < 1$  by assumption, it must hold that for every  $j \neq t$ , the continuous function

$$\frac{c_j^* f(c_j^*)}{1 - F(c_j^*)} \leq \min_x \frac{x f(x)}{1 - F(x)} \equiv M$$

has a strictly positive minimum value that we define as  $M$ . Similarly, the function

$$1 + c_t^* \frac{f'(c_t^*)}{f(c_t^*)} \leq L$$

is uniformly bounded from above. Therefore, we can bound the third component on the right-hand side of (28):

$$\sum_{j \neq t} \frac{c_j^* f(c_j^*)}{1 - F(c_j^*)} - 1 - c_t^* \frac{f'(c_t^*)}{f(c_t^*)} \geq (T - 1)M - L, \quad (29)$$

since the first term on the left-hand side of (29) sums over all periods except  $t$ . Thus, for large enough  $T \geq T_0 = (L/M) + 1$ , the third component on the right-hand side of (28) is positive, in which case  $\partial^2 \bar{P}/\partial \beta \partial \gamma_t \leq 0$  for all  $t$ .  $\square$

## B Model Simulations

We simulate the model to test when the formal restrictions required for the predictions to hold are binding. For simplicity, we assume that memory and trust are constant over time and thus drop the  $t$  subscript on  $\alpha$ ,  $\rho$ , and  $\hat{\rho}$ . Unless otherwise specified, in all simulations we set  $y = 1$ ,  $\alpha = 1$ ,  $T = 8$ ,  $F(c) \sim \text{Uniform}([0.4, 0.6])$ ,  $\delta = 0.99$ ; allow  $\rho$ ,  $\hat{\rho}$ , and  $\beta$  to vary over  $[0, 1]$  in increments of 0.01; and allow  $\hat{\beta}$  to vary over  $[\beta, 1]$  in increments of 0.01.

Predictions 1, 3, 4, and 6 only have the restriction that  $\hat{\beta}$  is sufficiently large such that monotonicity of continuation values holds (Lemma 1), while the proof of Prediction 2 does not impose any restrictions on the parameters of the model, so we do not include simulations for that prediction. Appendix-Figure B.1a shows the results of these simulations, where we plot for each value of  $\hat{\beta}$  the proportion of simulations (combinations of the parameters  $\rho$ ,  $\hat{\rho}$ , and  $\beta$ ) where the prediction holds.

To simulate Prediction 1, we allow  $y$  to vary from 0.5 to 1.5 in increments of 0.01 and test whether cumulative take-up of the offer by  $T$  is weakly monotonically increasing as we increase  $y$  over that range. Prediction 1 nearly always holds regardless of the level of  $\hat{\beta}$ : The lowest proportion of simulations in which this holds occurs when  $\hat{\beta} = 0.42$ , but even then the prediction holds in 99.6% of simulations. For Prediction 3, we test whether cumulative take-up of the offer by  $T$  is weakly monotonically increasing as we increase  $\hat{\rho}$  over the range  $[0, 1]$ . This prediction holds in 100% of simulations as long as  $\hat{\beta} > 0.42$ , and even when  $\hat{\beta} \leq 0.42$ , the prediction holds in a minimum of 98.0% of simulations for each value of  $\hat{\beta}$ . For Prediction 4, we allow  $\alpha$  to vary from 0 to 1 in increments of 0.01 and test whether cumulative take-up of the offer by  $T$  is weakly monotonically increasing as we increase  $\alpha$  over that range. This prediction holds in a minimum of 99.8% of simulations for each value of  $\hat{\beta}$ . For Prediction 6, we compare cumulative take-up in a deadline group where  $T = 8$  with take-up in a no-deadline group where  $T = 184$ , as in our experimental setting. (By Corollary 1, the no deadline case can be approximated by a sufficiently long deadline.) We then test whether the cumulative take-up by any day  $t \leq 8$  is weakly larger in the deadline group than in the no-deadline group. This prediction holds in a minimum of 84.4% of simulations for each value of  $\hat{\beta}$ .

Predictions 5 and 7 have additional restrictions in addition to  $\hat{\beta}$  being sufficiently large that Lemma 1 holds. Specifically, the proofs impose the following additional restrictions, where in an abuse of notation,  $T$  now denotes only the deadline group's deadline: (i) parameter values must be such that managers in the no-deadline arm adopt with positive probability, (ii)  $\rho$  must be sufficiently large, and (iii)  $T$  must be sufficiently large. We again test how binding these restrictions are in simulations of the model, and we exclude cases where managers in the no-deadline arm never adopt. Panels (b) and (c) of Appendix-Figure B.1 show the results of these simulations, where we now plot a set of heatmaps showing the proportion of simulations where the prediction holds, for

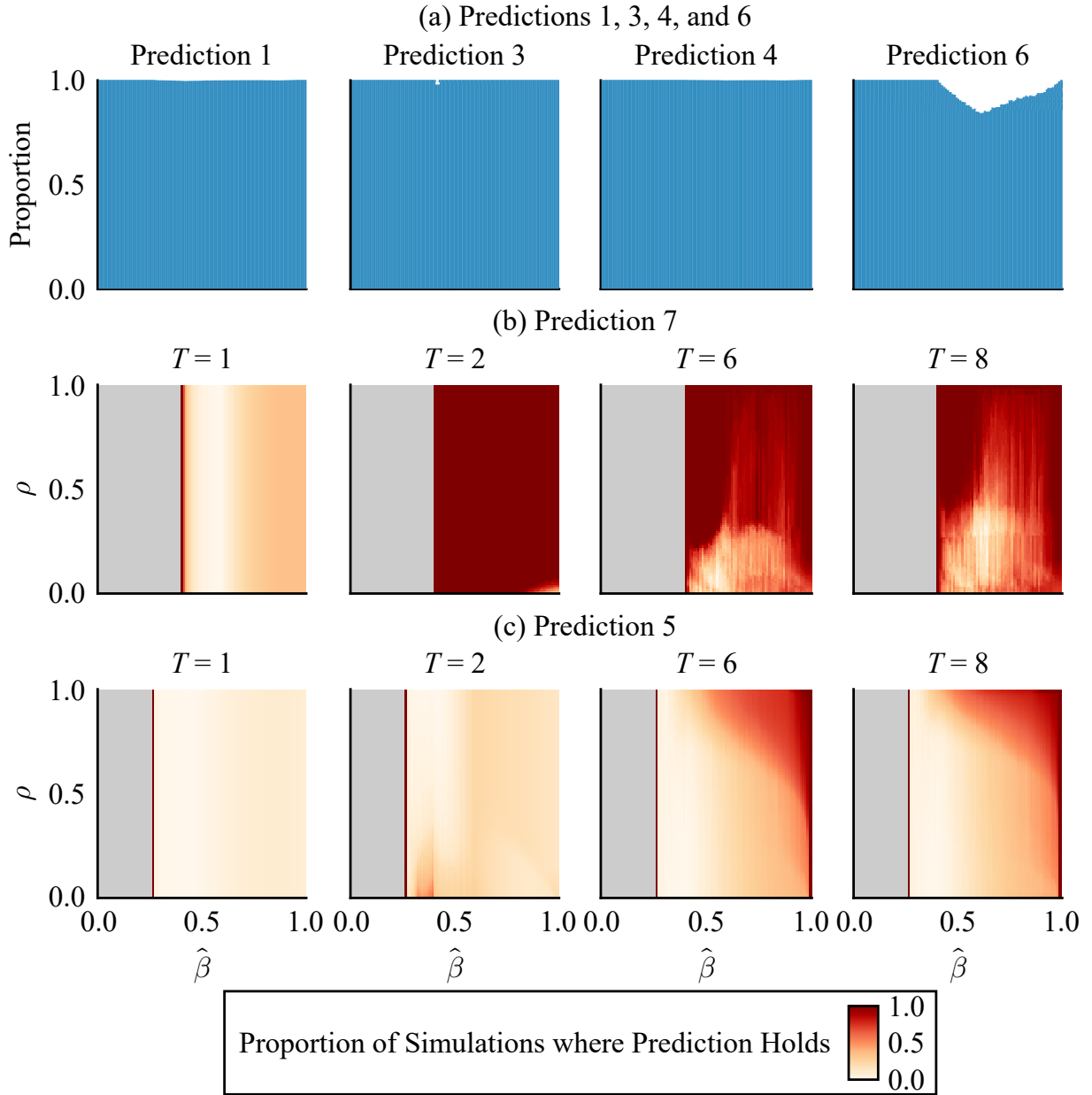
various  $\rho$ ,  $\hat{\beta}$ , and  $T$  (since these are the parameters with restrictions for the predictions to hold). Each heatmap is for a different value of  $T$ , and within each heatmap  $\rho$  and  $\hat{\beta}$  are the two parameters that vary on the y- and x-axes, respectively.

We begin with Prediction 7 since its simulation is more straightforward (as we allow one fewer parameter to vary than in the simulations of Prediction 5). We test whether, for each combination of values of  $\rho$ ,  $\hat{\rho}$ ,  $\beta$ ,  $\hat{\beta}$ , and  $T$ , the difference in cumulative take-up by  $t = T$  between the deadline and no-deadline arms weakly decreases in  $\beta$ . For values of  $\hat{\beta} < 0.41$ , which also implies  $\beta < 0.41$  due to the restriction that  $\hat{\beta} \geq \beta$ , the first restriction that the parameters are such that the manager without a deadline adopts with positive probability no longer holds. Thus for these very low values of  $\hat{\beta}$  and  $\beta$ , the proportion of simulations in which the prediction holds is not defined, hence that portion of the figure appears in gray. We begin with the  $T = 8$  case for the deadline group (as in our experiment). For reasonable values of  $\hat{\beta}$  and  $\rho$  (e.g.,  $\hat{\beta} \geq 0.7$  and  $\rho \geq 0.7$ ), Prediction 7 nearly always holds. To test when the restriction that  $T$  is sufficiently large binds for Prediction 7, we repeat the same process for various values of  $T$ , and find that the prediction continues to hold for reasonable  $\hat{\beta}$  and  $\rho$  values holds as long as  $T \geq 2$ .

For Prediction 5, we follow the same procedure as for Prediction 7, but we now additionally allow  $y$  to vary over  $[0.5, 1.5]$  in increments of 0.01 and test whether the treatment effect of the deadline on cumulative adoption by  $t = T$  is weakly decreasing in  $y$ . The restrictions for this prediction to hold are more binding: when  $T = 8$  as in our experiment, the prediction still holds in the majority of cases when  $\hat{\beta}$  and  $\rho$  are both sufficiently large, but there are a number of cases in which the prediction does not hold. For example, the prediction tends to (but does not always) hold in the region where  $\hat{\beta} \geq 0.8$  and  $\rho \geq 0.8$ . (Furthermore, for values of  $\beta \leq \hat{\beta} < 0.27$ , the parameter values are such that the manager without a deadline never adopts, and thus the proportion of simulations in which the prediction holds is undefined and those portions of the heatmap are shaded gray.) We next test the restriction that  $T$  must be sufficiently large by repeating the same process for various values of  $T$ , and find that results are similar to the  $T = 8$  case when  $T = 6$ , but that the prediction tends to fail to hold for lower values of  $T$ .

We conclude that, with the exception of Prediction 5, the formal restrictions on our model are inconsequential as the predictions nearly always hold even without these restrictions. In the case of Prediction 5, the restrictions of a sufficiently large  $\hat{\beta}$ ,  $\rho$ , and  $T$  do bind, but the prediction tends to hold when  $T = 8$  as in our experiment,  $\hat{\beta} \geq 0.8$ , and  $\rho \geq 0.8$ .

Figure B.1: Proportion of Simulations where Predictions Hold



This figure shows the proportion of simulations where each prediction holds across various parameter values. Panel (a) shows the proportion of simulations where Predictions 1, 3, 4, and 6 hold for various values of  $\hat{\beta}$  since the proofs of these predictions require restricting  $\hat{\beta}$  to be large enough that Lemma 1 holds. Panels (b) and (c) show the proportion of simulations where Prediction 7 and 5 hold for various values of  $\hat{\beta}$ ,  $\rho$ , and  $T$ , since these predictions impose the same restriction on  $\hat{\beta}$  as well as the restriction that  $\rho$  and  $T$  be sufficiently large. Other parameter values for the simulations are described in Appendix B.



## C Tables and Figures

Table C.1: Balance in Probability of Exclusion

	Excluded (1)
Intercept	0.0006*** (0.0002)
Unannounced reminder	-0.0001 (0.0004)
Announced reminder	0.0010* (0.0005)
Deadline	0.0003 (0.0004)
Same-day deadline	0.0005 (0.0006)
2.75% Fee	-0.0001 (0.0004)
<i>F</i> -statistic	1.289 [0.265]
Number of observations	34,010

This table tests for differential exclusion by treatment arm. The dependent variable equals 1 if the firm was excluded from the RCT sample due to the FinTech partner's filtering of firms that were not in good standing administratively. The unit of observation is a firm, and the regression includes all firms prior to exclusions ( $N = 34,010$ ). Heteroskedasticity-robust standard errors are in parentheses. The  $p$ -value of the  $F$ -statistic for the joint test of whether exclusion is predicted by the various treatments is in square brackets. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.2: Baseline Treatment Balance (Full Sample)

	Intercept	Unannounced reminder	Announced reminder	Deadline	Same-day deadline	2.75% Fee	Joint test <i>F</i> -stat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<b>Panel A: Manager characteristics</b>							
Female	0.440*** (0.005)	-0.001 (0.007)	0.003 (0.007)	-0.002 (0.006)	0.005 (0.009)	0.002 (0.006)	0.256 [0.937]
Age	39.457*** (0.12)	0.186 (0.16)	0.244 (0.17)	-0.039 (0.14)	-0.198 (0.20)	-0.011 (0.13)	1.051 [0.385]
<b>Panel B: Business characteristics</b>							
<i>Business type</i>							
Small retailers	0.259*** (0.005)	-0.001 (0.006)	-0.001 (0.007)	0.001 (0.005)	0.003 (0.008)	0.000 (0.005)	0.037 [0.999]
Professionals	0.239*** (0.004)	-0.001 (0.006)	-0.001 (0.006)	0.001 (0.005)	0.000 (0.008)	0.000 (0.005)	0.014 [1.000]
Beauty	0.087*** (0.003)	0.000 (0.004)	0.000 (0.004)	0.002 (0.003)	0.000 (0.005)	0.000 (0.003)	0.065 [0.997]
Clothing	0.090*** (0.003)	0.000 (0.004)	0.000 (0.004)	0.000 (0.003)	-0.002 (0.005)	0.000 (0.003)	0.027 [1.000]
Restaurants	0.122*** (0.003)	0.002 (0.005)	0.001 (0.005)	0.000 (0.004)	0.000 (0.006)	-0.001 (0.004)	0.038 [0.999]
Other	0.203*** (0.004)	0.000 (0.006)	0.002 (0.006)	-0.003 (0.005)	-0.001 (0.007)	0.001 (0.004)	0.115 [0.989]
<i>Pre-treatment sales variables</i>							
Months since first transaction	24.101*** (0.17)	0.120 (0.24)	0.101 (0.25)	-0.074 (0.20)	0.022 (0.30)	0.120 (0.19)	0.173 [0.973]
% months business made sales	0.817*** (0.002)	-0.001 (0.003)	0.001 (0.003)	0.001 (0.003)	0.003 (0.004)	0.001 (0.003)	0.210 [0.958]
Log monthly card sales volume	8.783*** (0.012)	0.014 (0.015)	-0.011 (0.016)	0.012 (0.013)	0.024 (0.020)	-0.002 (0.012)	0.806 [0.545]
Log monthly card transactions	2.057*** (0.015)	-0.003 (0.020)	-0.010 (0.021)	0.003 (0.017)	-0.015 (0.025)	0.006 (0.016)	0.137 [0.984]

This table reports differences in characteristics of the manager (typically the firm owner) and of the business by treatment group. The unit of observation is a firm, and each regression includes all firms in the experiment ( $N = 33,978$ ). Each row shows coefficients from a regression of that row's characteristic on an intercept (column 1) and indicator variables for unannounced reminder (column 2), announced reminder (column 3), deadline (column 4), same-day deadline (column 5), and 2.75% fee (column 6). Column 7 shows the  $F$ -statistic and corresponding  $p$ -value from an omnibus  $F$ -test of the coefficients on all treatment group dummies in that row's regression. Manager and business characteristics are defined when the user signs up for the technology. Pre-treatment sales variables include only card sales and are an average over all months from July 2019 to August 2020. Heteroskedasticity-robust standard errors are included in parentheses, and  $p$ -values for the  $F$ -statistics are in square brackets. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.3: Baseline Treatment Balance (Survey Subsample)

	Intercept	Unannounced reminder	Announced reminder	Deadline	2.75% Fee	Joint test <i>F</i> -stat
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Panel A: Manager characteristics</b>						
Female	0.396*** (0.083)	-0.029 (0.084)	-0.053 (0.083)	0.060 (0.046)	0.103* (0.046)	1.832 [0.121]
Age	41.067*** (1.48)	-0.698 (1.58)	-1.216 (1.53)	0.193 (1.00)	-1.084 (1.00)	0.456 [0.768]
<b>Panel B: Business characteristics</b>						
<i>Business type</i>						
Small retailers	0.339*** (0.077)	-0.103 (0.078)	-0.126 (0.077)	0.012 (0.041)	0.050 (0.041)	1.151 [0.332]
Professionals	0.234*** (0.075)	0.049 (0.074)	0.013 (0.073)	-0.001 (0.042)	0.057 (0.042)	0.694 [0.597]
Beauty	0.133** (0.053)	-0.069 (0.053)	-0.078 (0.052)	-0.025 (0.023)	0.023 (0.023)	1.483 [0.206]
Clothing	0.032 (0.029)	0.068* (0.030)	0.061* (0.029)	0.001 (0.025)	-0.023 (0.025)	0.789 [0.533]
Restaurants	0.121*** (0.045)	0.030 (0.048)	0.028 (0.047)	-0.006 (0.029)	-0.064* (0.030)	1.283 [0.276]
Other	0.141** (0.056)	0.025 (0.057)	0.103* (0.059)	0.020 (0.036)	-0.044 (0.036)	1.814 [0.125]
<i>Pre-treatment sales variables</i>						
Months since first transaction	21.257*** (2.37)	2.620 (2.47)	1.091 (2.41)	2.617 (1.64)	-0.833 (1.64)	1.021 [0.396]
% months business made sales	0.866*** (0.029)	-0.036 (0.031)	-0.039 (0.030)	0.004 (0.020)	-0.014 (0.020)	0.449 [0.773]
Log monthly card sales volume	8.616*** (0.168)	0.178 (0.172)	0.120 (0.169)	-0.066 (0.100)	0.078 (0.100)	0.554 [0.696]
Log monthly card transactions	2.043*** (0.200)	0.037 (0.206)	-0.082 (0.203)	0.083 (0.127)	0.007 (0.128)	0.320 [0.864]

This table reports differences in characteristics of the manager (typically the firm owner) and of the business by treatment group. The unit of observation is a firm, and each regression includes all firms in the survey sample ( $N = 471$ ). Each row shows coefficients from a regression of that row's characteristic on an intercept (column 1) and indicator variables for unannounced reminder (column 2), announced reminder (column 3), deadline (column 4), and 2.75% fee (column 5). Column 6 shows the  $F$ -statistic and corresponding  $p$ -value from an omnibus  $F$ -test of all coefficient estimates in the regression. Manager and business characteristics are defined when the user signs up for the technology. Pre-treatment sales variables include only card sales and are an average over all months from July 2019 to August 2020. Heteroskedasticity-robust standard errors are in parentheses, and  $p$ -values for the  $F$ -statistics are in square brackets. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.4: Balance Between Survey Sample and Non-survey Sample

	Non-survey sample (1)	Survey sample (2)	Difference (3)	<i>p</i> -value (4)
<b>Panel A: Manager characteristics</b>				
Female	0.441 (0.497)	0.444 (0.497)	0.003 (0.023)	[0.907]
Age	39.509 (11.03)	39.732 (10.44)	0.223 (0.50)	[0.656]
<b>Panel B: Business characteristics</b>				
<i>Business type</i>				
Small retailers	0.260 (0.438)	0.268 (0.443)	0.008 (0.021)	[0.705]
Professionals	0.238 (0.426)	0.291 (0.455)	0.053 (0.021)	[0.012]**
Beauty	0.088 (0.283)	0.066 (0.248)	-0.022 (0.012)	[0.059]*
Clothing	0.090 (0.285)	0.079 (0.269)	-0.011 (0.012)	[0.381]
Restaurants	0.123 (0.329)	0.110 (0.314)	-0.013 (0.015)	[0.379]
Other	0.202 (0.402)	0.187 (0.390)	-0.015 (0.018)	[0.400]
<i>Pre-treatment sales variables</i>				
Months since first transaction	24.187 (16.94)	23.804 (17.75)	-0.383 (0.82)	[0.641]
% months business made sales	0.818 (0.227)	0.826 (0.220)	0.008 (0.010)	[0.446]
Log monthly card sales volume	8.791 (1.113)	8.757 (1.077)	-0.034 (0.050)	[0.492]
Log monthly card transactions	2.055 (1.423)	2.066 (1.379)	0.011 (0.064)	[0.863]
<i>F</i> -stat of joint test			1.09	[0.365]
Number of observations	33,507	471	33,978	33,978

This table reports differences in characteristics of the manager (typically the firm owner) and of the business by sample. The unit of observation is a firm. Column 1 contains the average mean and standard deviation of the non-survey sample, and column 2 of the survey sample. Column 3 shows the difference between columns 2 and 1 and standard error, and column 4 reports the associated *p*-value of the difference test. Data is from July 2019 to August 2020 and includes all firms in the experiment ( $N = 33,978$ ). \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.5: Treatment Effects on Cumulative Adoption by Day 8

	Firm accepted offer
<u>Panel A: Offer value</u>	
2.75% fee	0.294*** (0.004)
3.00% fee	0.259*** (0.004)
<u>Panel B: Treatment arm</u>	
No deadline, no reminder	0.254*** (0.007)
No deadline, unannounced reminder	0.290*** (0.007)
No deadline, announced reminder	0.305*** (0.008)
Deadline, no reminder	0.256*** (0.006)
Deadline, unannounced reminder	0.298*** (0.007)
Deadline, announced reminder	0.318*** (0.008)
Same-day deadline, no reminder	0.229*** (0.006)
Number of firms	33,978

This table reports the estimated effect of being assigned to a treatment group, separately for the fee type in panel A and the reminder and deadline type in panel B on the probability of take-up. The unit of observation is a firm. The omitted dummy in the regression is that of the control group. Data include all firms in the experiment ( $N = 33,978$ ), and the dependent variable is an indicator equal to 1 if a firm accepts the offer by day 8. The regression includes strata fixed effects. Heteroskedasticity-robust standard errors are in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.6: Effects of Reminder, Announced Reminder, and Deadline on Days 1–6 vs. 7–8

Treatment:	Firm accepted offer		
	Reminder (1)	Announced reminder (2)	Deadline (3)
Treatment	-0.005 (0.005)	0.009 (0.006)	-0.008 (0.005)
Treatment $\times$ Day 7 or 8	0.048*** (0.002)	0.010** (0.004)	0.012*** (0.003)
Number of observations	202,616	130,032	202,616
Number of firms	25,327	16,254	25,327
Cumulative take-up on day 6	0.244	0.234	0.243

This table reports treatment effects of an announced or unannounced reminder relative to no reminder (column 1), announced reminder relative to unannounced reminder (column 2), and deadline relative to no deadline (column 3), comparing take-up on days 1–6 (before the reminder was sent for those with a reminder) against take-up on days 7–8 (after the reminder was sent for those with a reminder). The unit of observation is a firm by day. Columns 1 and 3 exclude the same-day deadline and pure control groups, while column 2 excludes firms without reminders. “Day 7 or 8” is an indicator equal to 1 if the observation is from day 7 or 8. “Cumulative take-up by day 6” gives the cumulative take-up by day 6 in the relevant comparison group, i.e., the no-reminder group in column 1, the unannounced reminder group in column 2, and the no-deadline group in column 3. Regressions include day and strata fixed effects. Standard errors clustered at the firm level are included in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.7: Heterogeneous Effects by Owner Receiving Emails

Treatment:	Firm accepted offer					
	Reminder		Announced reminder		Deadline	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.429*** (0.109)	0.556*** (0.166)	0.379*** (0.064)	0.409*** (0.105)	0.494*** (0.057)	0.500*** (0.095)
Treatment	0.087 (0.117)	0.035 (0.182)	0.246*** (0.086)	0.364*** (0.138)	0.020 (0.082)	0.180 (0.133)
Above-median % sales using technology	0.238 (0.175)		0.171** (0.085)		0.072 (0.081)	
Above-median % sales using technology × Treatment	-0.159 (0.185)		-0.145 (0.119)		0.047 (0.114)	
Owner was recipient of emails		-0.123 (0.185)		0.047 (0.112)		-0.020 (0.101)
Owner was recipient of emails × Treatment		0.083 (0.201)		-0.181 (0.147)		-0.063 (0.142)
Number of firms	306	471	273	425	306	471

This table reports heterogeneous treatment effects of an announced or unannounced reminder relative to no reminder (columns 1–2), announced reminder relative to unannounced reminder (columns 3–4), and deadline relative to no deadline (columns 5–6), by the percentage of sales volume using the FinTech payments technology (columns 1, 3, and 5) or by whether the owner is the recipient of the FinTech company’s emails (columns 2, 4, and 6). The outcome variable measures whether the firm accepted the offer by day 8 (which was the deadline for firms in the deadline arm). In columns 1, 3, and 5, the sample excludes those who did not answer the survey questions used to construct the “% sales using technology” variable, which measures the percent of sales volume the firm transacts through the FinTech payment technology. Column 3 further restricts the sample to the announced and unannounced reminder groups. In columns 2, 4, and 6, “Owner was recipient of emails” is an indicator constructed from answers to the survey question, “Is the owner the person that receives emails from [redacted]?”, which was asked to all survey respondents. The answers indicate that the owner receives the emails in 88.7% of firms. Columns 2 and 6 include all firms in the survey, while column 4 excludes firms in the no-reminder arm. Heteroskedasticity-robust standard errors are in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.8: Heterogeneous Effects by Number of Employees

Treatment:	Firm accepted offer					
	Reminder		Announced reminder		Deadline	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0.478*** (0.105)	0.571*** (0.133)	0.472*** (0.053)	0.459*** (0.082)	0.485*** (0.051)	0.538*** (0.080)
Treatment	0.050 (0.111)	-0.036 (0.146)	0.112 (0.075)	0.158 (0.117)	0.073 (0.071)	0.005 (0.109)
Above-median # of employees	-0.024 (0.149)		-0.043 (0.071)		0.008 (0.068)	
Above-median # of employees × Treatment	0.073 (0.157)		0.168* (0.097)		0.074 (0.093)	
More than 1 employee		-0.152 (0.160)		-0.014 (0.091)		-0.060 (0.088)
More than 1 employee × Treatment		0.178 (0.173)		0.062 (0.129)		0.134 (0.120)
Number of firms	462	462	417	417	462	462
Mean heterogeneity variable	0.565	0.816	0.573	0.830	0.565	0.816

This table reports heterogeneous treatment effects of an announced or unannounced reminder relative to no reminder (columns 1–2), announced reminder relative to unannounced reminder (columns 3–4), and deadline relative to no deadline (columns 5–6), by number of employees. We use two measures for the number of employees: an indicator for above-median number of employees, defined as firms with  $\geq 3$  employees, and an indicator for more than one employee. The outcome variable measures whether the firm accepted the offer by day 8 (which was the deadline for firms in the deadline arm). All 471 firms in the survey were asked, “How many employees work in your business, including yourself?” The median number of employees is 3, and 9 firms that did not answer the question are excluded from the sample. Columns 1, 2, 5, and 6 include all firms that answered the question, while columns 3 and 4 restrict the sample to announced and unannounced-reminder groups. Heteroskedasticity-robust standard errors are included in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



Table C.9: Heterogeneous Effects by Firm Business Type

Treatment:	Firm accepted offer		
	Reminder (1)	Announced reminder (2)	Deadline (3)
Treatment	0.039*** (0.013)	0.019 (0.016)	0.015 (0.012)
Small retailers	0.002 (0.017)	0.016 (0.021)	-0.006 (0.017)
× Treatment			
Professionals	0.012 (0.018)	-0.008 (0.022)	0.012 (0.017)
× Treatment			
Beauty	0.014 (0.023)	-0.024 (0.028)	-0.029 (0.022)
× Treatment			
Clothing	0.004 (0.023)	-0.011 (0.029)	-0.023 (0.022)
× Treatment			
Restaurants	0.016 (0.020)	-0.006 (0.026)	-0.041** (0.020)
× Treatment			
Number of firms	25,327	16,254	25,327

This table reports heterogeneous treatment effects of an announced or unannounced reminder relative to no reminder (column 1), announced reminder relative to unannounced reminder (column 2), and deadline relative to no deadline (column 3), by business type of the firm. The omitted business-type category is “Other.” The outcome variable measures whether the firm accepted the offer by day 8. Columns 1 and 3 exclude the same-day deadline and pure control groups from the full sample of firms, while column 2 excludes firms without reminders. Regressions include strata fixed effects, which absorb non-interacted business-type dummies. Heteroskedasticity-robust standard errors are in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.10: Heterogeneous Effects by Baseline Characteristics

Treatment:	Firm accepted offer								
	Reminder			Announced reminder			Deadline		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.048*** (0.009)	0.046*** (0.008)	0.038*** (0.008)	0.005 (0.011)	0.011 (0.010)	0.020** (0.010)	-0.000 (0.008)	0.001 (0.008)	0.015* (0.008)
Age	-0.002 (0.010)			-0.016 (0.010)			-0.010 (0.008)		
Age × Treatment	-0.002 (0.012)			0.028* (0.015)			0.016 (0.012)		
Female		-0.027*** (0.010)			-0.031*** (0.010)			-0.032*** (0.008)	
Female × Treatment		0.003 (0.012)			0.016 (0.015)			0.014 (0.012)	
Growth			0.024*** (0.009)			0.043*** (0.010)			0.043*** (0.008)
Growth × Treatment			0.017 (0.012)			-0.005 (0.014)			-0.016 (0.011)
Number of firms	23,614	23,617	25,327	15,138	15,141	16,254	23,614	23,617	25,327

This table reports heterogeneous treatment effects of an announced or unannounced reminder relative to no reminder (columns 1–3), announced reminder relative to unannounced reminder (columns 4–6), and deadline relative to no deadline (columns 7–9), by (i) age of the owner, (ii) whether the owner is female, and (iii) firm growth. Firm growth is measured as the month-over-month change in baseline sales from August to September 2020 (excluding the days of the experiment, September 29 and 30). The outcome variable measures whether the firm accepted the offer by day 8 (which was the deadline for firms in the deadline arm). Columns 1–3 and 7–9 exclude firms in the same-day deadline and pure control groups, while columns 4–6 also exclude firms in the no-reminder group. In addition, columns 1, 5, and 7 exclude firms for which age of the owner is missing in administrative data, and columns 2, 5, and 8 exclude firms for which sex of the owner is missing. Heteroskedasticity-robust standard errors are in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.11: Heterogeneous Effects of Announced Reminder by GSS Measures

	Firm accepted offer					
	Trust (1)	Reciprocity (2)	Procrastination (3)	Memory (4)	Overconfidence (5)	Attention (6)
Intercept	0.439*** (0.048)	0.526*** (0.115)	0.538*** (0.044)	0.532*** (0.073)	0.500*** (0.050)	0.387*** (0.088)
Survey measure	0.211*** (0.072)	0.003 (0.121)	-0.029 (0.081)	-0.003 (0.085)	0.063 (0.073)	0.171* (0.097)
Announced reminder	0.303*** (0.062)	0.188 (0.152)	0.206*** (0.058)	0.186** (0.091)	0.252*** (0.064)	0.305** (0.127)
Survey measure × Announced reminder	-0.301*** (0.100)	-0.007 (0.160)	-0.069 (0.106)	-0.007 (0.108)	-0.165* (0.099)	-0.149 (0.137)
Romano-Wolf $p$ -values	[0.029] <sup>++</sup>	[0.998]	[0.888]	[0.998]	[0.411]	[0.703]
Number of firms	388	388	388	388	388	388
Prop. survey measure = 1	0.367	0.895	0.313	0.682	0.421	0.841

This table reports heterogeneous treatment effects of the announced reminder by general social survey (GSS) measures collected in our survey. The survey questions asked respondents how much they agree on a scale from 1 to 5 with the following six statements: (1) *Trust*: I trust advertised offers. (2) *Reciprocity*: I am more inclined to do business with people who live up to their promises. (3) *Procrastination*: I tend to postpone tasks, even when I know it is better to do them immediately. (4) *Memory*: I tend to have good memory about pending tasks that I have to do and complete. (5) *Overconfidence*: I tend to think my memory is better than it really is. (6) *Attention*: I can focus completely when I have to finish a task. “Survey measure” is a dummy variable that codes responses of 4 and 5 (agree and completely agree) to these questions as 1 and 1–3 (completely disagree, disagree and neither agree nor disagree) as 0. The sample includes firms with announced and unannounced reminders in the survey sample. The outcome variable measures whether the firm accepted the offer by March 31. All firms in the survey were asked these questions. 43 firms that did not answer these questions were excluded from the sample. Heteroskedasticity-robust standard errors are in parentheses. Romano-Wolf adjusted  $p$ -values for the interaction term are in square brackets. Stars are based on heteroskedasticity-robust  $p$ -values with \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Plus signs represent Romano-Wolf  $p$ -values adjusted for multiple hypothesis correction with <sup>+</sup>  $p < 0.1$ , <sup>++</sup>  $p < 0.05$ , <sup>+++</sup>  $p < 0.01$ . GSS = general social survey; prop. = proportion.

Table C.12: Heterogeneous Effects of Unannounced Reminder by GSS Measures

	Firm accepted offer					
	Trust (1)	Reciprocity (2)	Procrastination (3)	Memory (4)	Overconfidence (5)	Attention (6)
Intercept	0.406*** (0.088)	0.600*** (0.221)	0.586*** (0.092)	0.278*** (0.107)	0.370*** (0.094)	0.273** (0.135)
Survey measure	0.344* (0.178)	-0.143 (0.237)	-0.404*** (0.149)	0.359** (0.149)	0.322** (0.160)	0.279* (0.164)
Unannounced reminder	0.033 (0.100)	-0.074 (0.249)	-0.048 (0.102)	0.254* (0.129)	0.130 (0.106)	0.114 (0.162)
Survey measure × Unannounced reminder	-0.133 (0.192)	0.146 (0.266)	0.376** (0.170)	-0.362** (0.171)	-0.259 (0.176)	-0.108 (0.191)
Romano-Wolf <i>p</i> -values	[0.847]	[0.847]	[0.237]	[0.237]	[0.468]	[0.847]
Number of firms	227	227	227	227	227	227
Prop. survey measure = 1	0.367	0.895	0.313	0.682	0.421	0.841

This table reports heterogeneous treatment effects of the unannounced reminder by general social survey (GSS) measures collected in our survey. The survey questions asked respondents how much they agree on a scale from 1 to 5 with the following six statements: (1) *Trust*: I trust advertised offers. (2) *Reciprocity*: I am more inclined to do business with people who live up to their promises. (3) *Procrastination*: I tend to postpone tasks, even when I know it is better to do them immediately. (4) *Memory*: I tend to have good memory about pending tasks that I have to do and complete. (5) *Overconfidence*: I tend to think my memory is better than it really is. (6) *Attention*: I can focus completely when I have to finish a task. The indicator *High survey measures* codes responses of 4 and 5 (agree and completely agree) as 1 and 1-3 (completely disagree, disagree and neither agree nor disagree) as 0. The unit of observation is a firm. Data include firms with unannounced and no reminders in the survey sample and take-up from September 29, 2020, to March 31, 2021. All firms in the survey were asked these questions. 43 firms that did not answer these questions were excluded from the sample. Heteroskedasticity-robust standard errors are in parentheses. Romano-Wolf adjusted *p*-values for the interaction term are in square brackets. Stars are based on heteroskedasticity-robust *p*-values with \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Plus signs represent Romano-Wolf *p*-values adjusted for multiple hypothesis correction with +  $p < 0.1$ , ++  $p < 0.05$ , +++  $p < 0.01$ . GSS = general social survey; prop. = proportion.

Table C.13: Account Log ins by Treatment

	Firm logged in (1)	Firm viewed deposits (2)
Intercept	0.092*** (0.003)	0.037*** (0.002)
Unannounced reminder	0.000 (0.004)	0.000 (0.003)
Announced reminder	-0.003 (0.004)	0.000 (0.003)
Deadline	-0.001 (0.003)	0.000 (0.002)
Same-day deadline	0.000 (0.005)	0.002 (0.003)
2.75% offer	0.006* (0.003)	-0.001 (0.002)
Number of firms	33,978	33,978

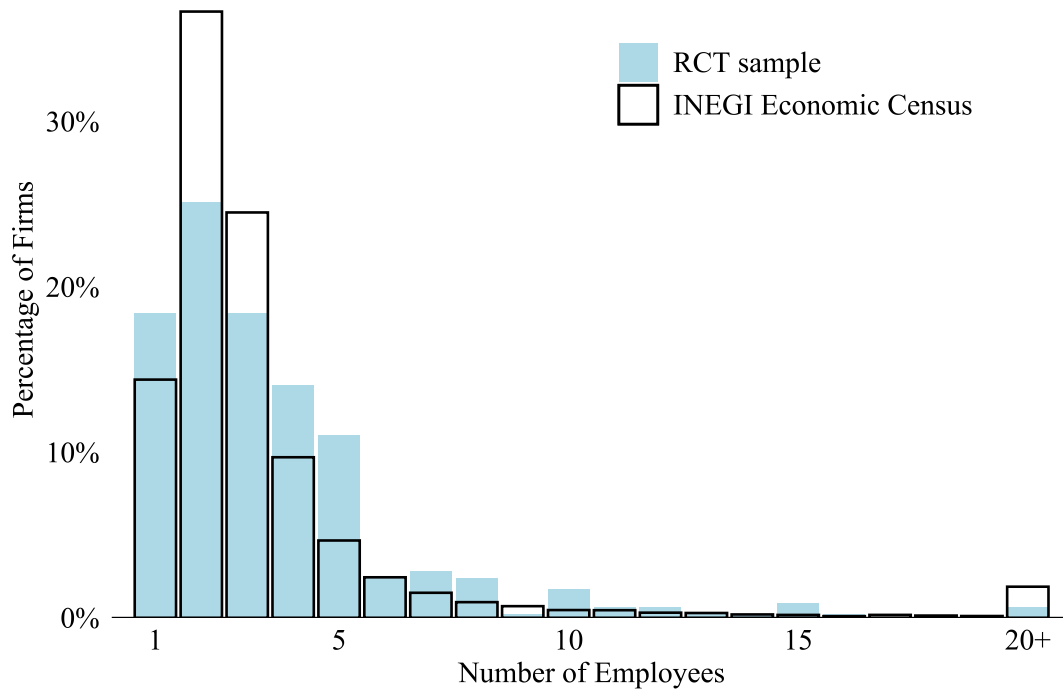
This table reports differences in the probability of a firm logging into the FinTech platform (column 1) and viewing deposits (column 2) by treatment, at any point between day 1 and day 8 of the experiment. The sample includes all firms in the experiment. Heteroskedasticity-robust standard errors are in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table C.14: Monthly Sales Elasticity: Intent to Treat and Treatment on the Treated

	Log(sales + 1) (1)	Log(# transactions + 1) (2)	Continued using technology (3)
<u>Panel A: Intent to Treat</u>			
Post * Treated	0.101** (0.046)	0.030* (0.016)	0.013** (0.005)
<u>Panel B: Treatment on the Treated</u>			
Post * Accepted	0.351** (0.161)	0.102* (0.055)	0.043** (0.017)
Number of observations	662,162	662,162	662,162
Number of firms	33,978	33,978	33,978
Cluster std. errors	Firm	Firm	Firm
Fixed effects	Firm & month	Firm & month	Firm & month
Control mean (levels)	24,471	30.02	0.847
Control mean (levels, winsorized)	12,178	19.52	0.847

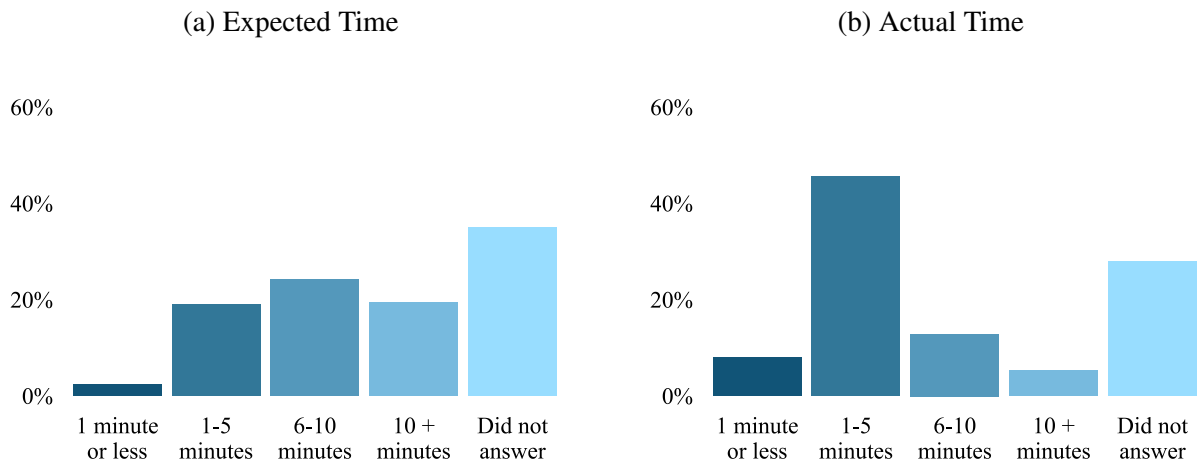
This table reports sales elasticities of the treated group (who received the offer) in Panel A and of the compliers (who accepted the offer) in Panel B. Data is from July 2019 to March 2021, includes Sep 29 and Sep 30 as part of October, and contains all firms in the experiment. The unit of observation is a firm-month. Post \* Treated is an interaction term of Post and Treated, where 'Post' is an indicator for the time period is after the treated firms received the offer and 'Treated' is an indicator for firms who received the offer. Post \* Accepted is an interaction term of Post and Adopted, where 'Accepted' is an indicator for firms that accepted the offer. Post \* Accepted is instrumented by Post \* Treated. Control means include data from the treatment period. Log average monthly sales volume and log average monthly transactions transform sales volume and transactions after winsorizing at the 95th percentile. Regressions include firm and month fixed effects. Clustered standard errors at the firm level are in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure C.1: Number of Employees



This figure shows the histograms of the number of employees by firm from our survey ( $N = 462$ ) and from the 2019 INEGI Economic Census ( $N = 5,360,215$ ). Our survey was conducted on a random sample of managers in the experiment ( $N = 471$  responses) and included the question: “How many employees work in your business, including yourself?” 9 managers that did not answer the survey question are excluded from the figure. We top-code the figure at 20 employees. 99.8% of firms in our sample and 98.2% of firms in the Economic Census have  $\leq 20$  employees. In the RCT sample, the average number of employees is 3.9, the median is 3, the maximum 150, and the standard deviation 7.4.

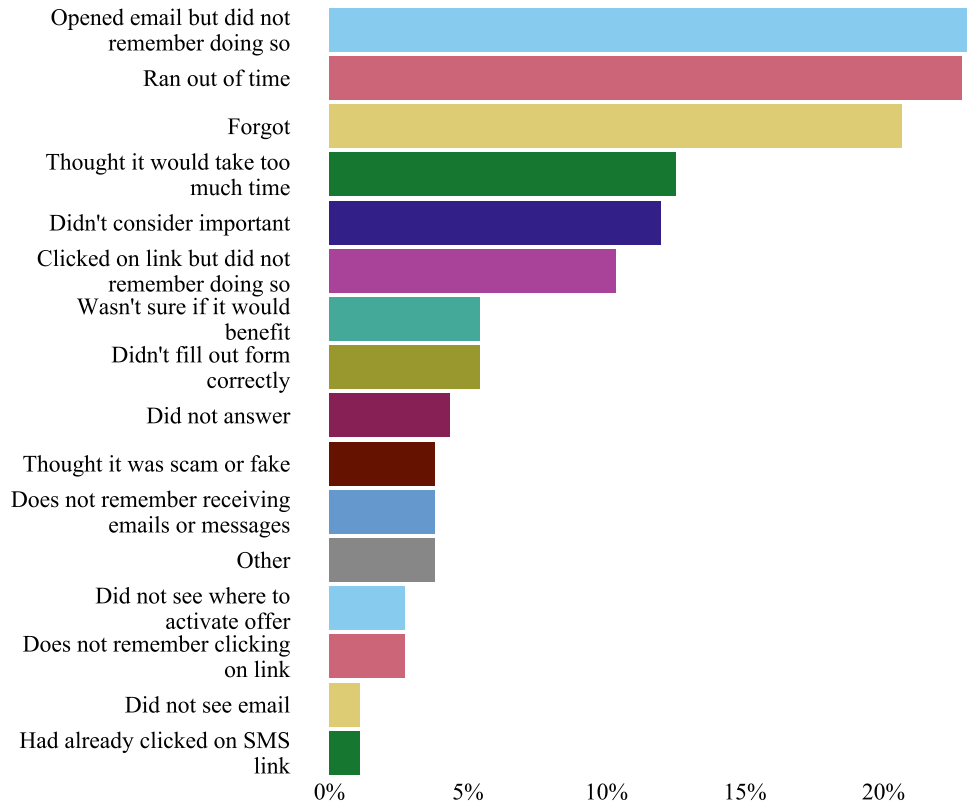
Figure C.2: Self-Reported Time Cost of Accepting the Offer



The figure shows how long firms expected it would take them to fill out the form to accept the offer (left panel) and how long it actually took them (right panel). The left panel shows responses to the survey question: “How long did you expect completing the form to activate the lower fee would take you?” This question was asked to users who recall receiving the first email or SMS ( $N = 289$ ). The right panel shows responses to the survey question: “How long did it take you to fill out the offer?” This question was asked to respondents who recalled receiving the first email or SMS and recalled accepting the offer or clicking on the link in the email to accept the offer ( $N = 186$ ).

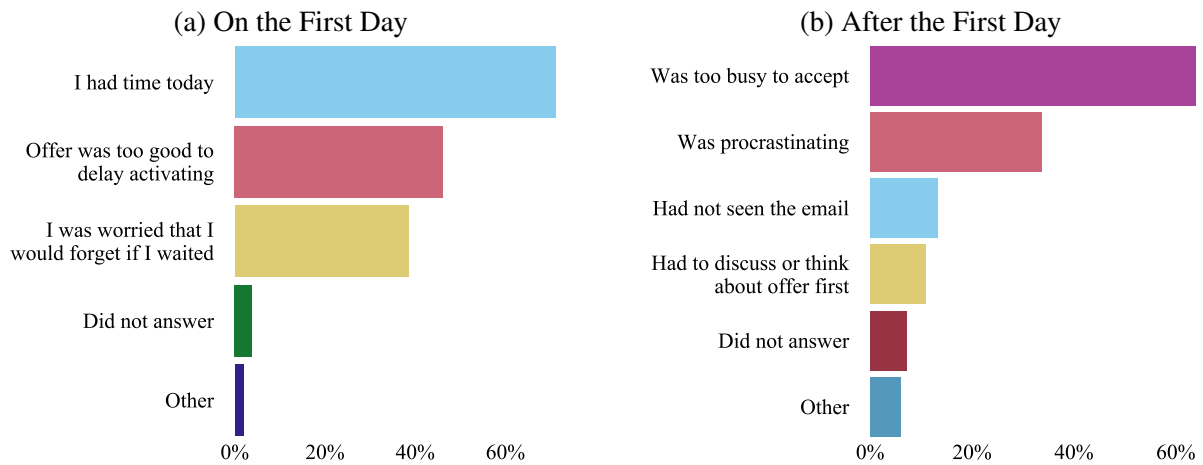


Figure C.3: Self-Reported Reasons Why Firms Did Not Adopt



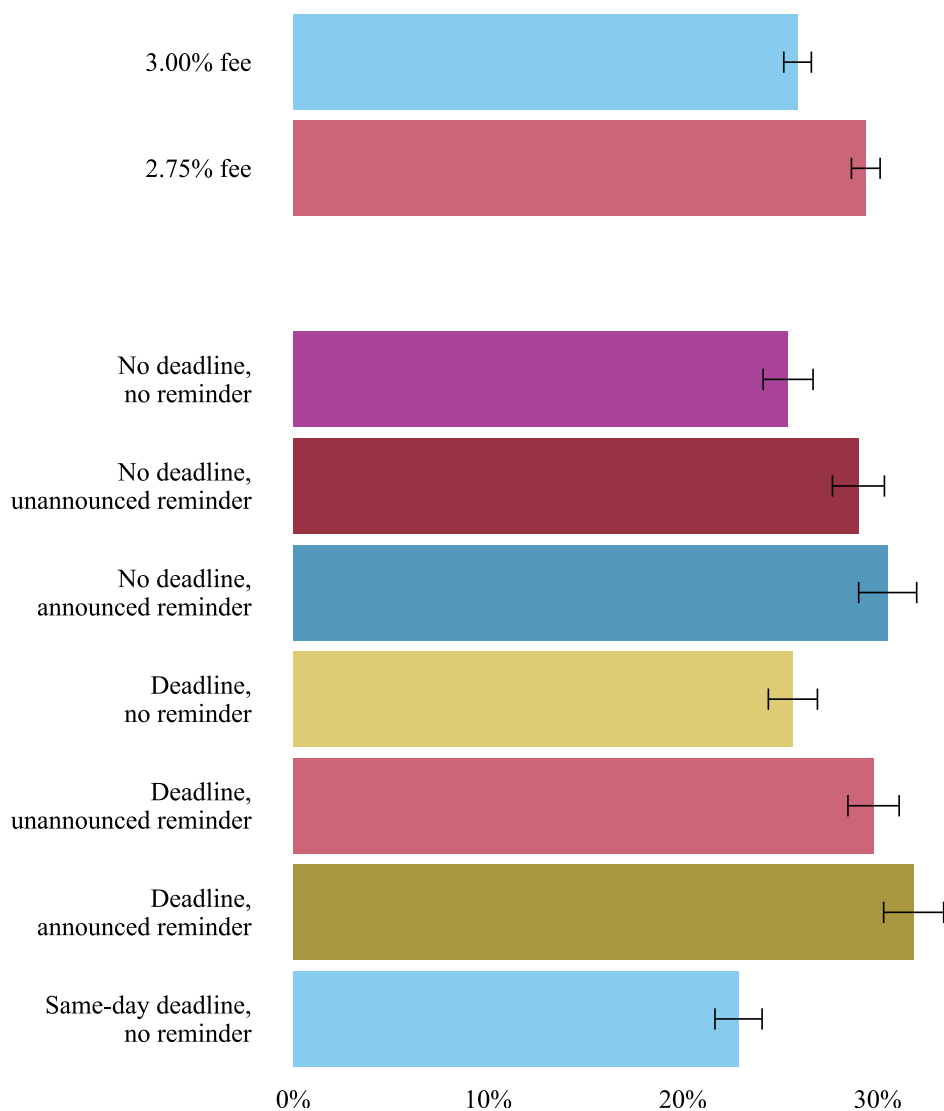
The figure shows a barplot with the reasons given by firms for not adopting. The survey was fielded on a random subsample of  $N = 471$  firms, and this figure includes 169 firms that did not adopt. The figure combines responses from the following survey questions: If our administrative data show that the respondent did not open the email, we ask “Our records show that you did not open the email. Why not?” If our administrative data show that the respondent opened the email but did not click the link, we ask “What prevented you from clicking the link and filling out the form?” If our administrative data show that the respondent clicked the link but did not complete the form, we ask “We observe that you did not complete the form after clicking the offer link. Why did you not complete the form?”

Figure C.4: Why Firms Accepted the Offer On or After the First Day



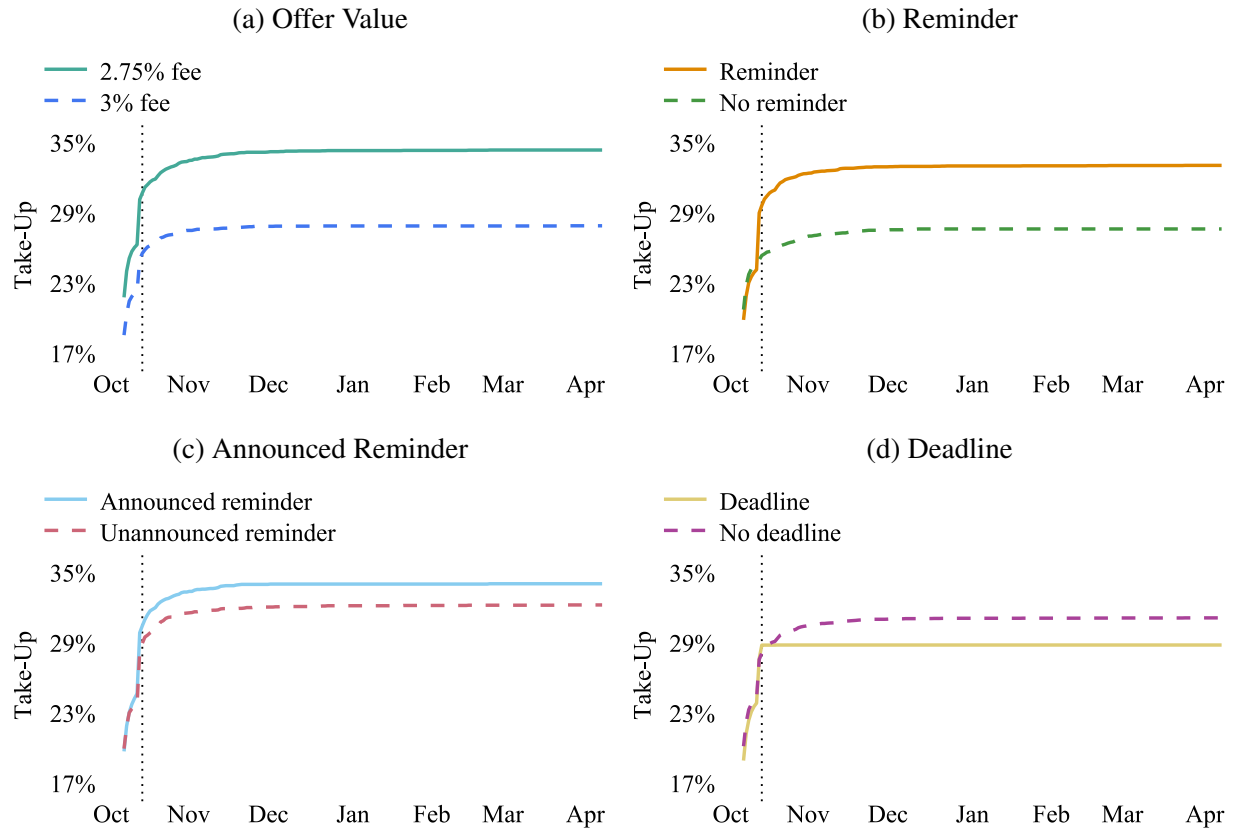
The left panel shows responses to the survey question: “Our records show that you activated the offer on September 29, even though your deadline to activate the offer was not until October 6. Why did you activate the offer on September 29?” This question was asked to firms that recall receiving the first email or SMS, received a deadline, and accepted the offer on day 1 ( $N = 52$ ). The right panel shows responses to the survey question: “We sent you the emails and SMS to let you know about this offer on September 29, but we see that you filled the form on {activation date}. Why did you wait until {days to accept} day(s) later?” This question was asked to firms that recalled receiving the first email or SMS, had a deadline, accepted the offer after day 1, and recalled accepting the offer or clicking the link to accept the offer ( $N = 83$ ). Respondents could provide more than one response, so totals add up to more than 100%.

Figure C.5: Take-up by Treatment Arm



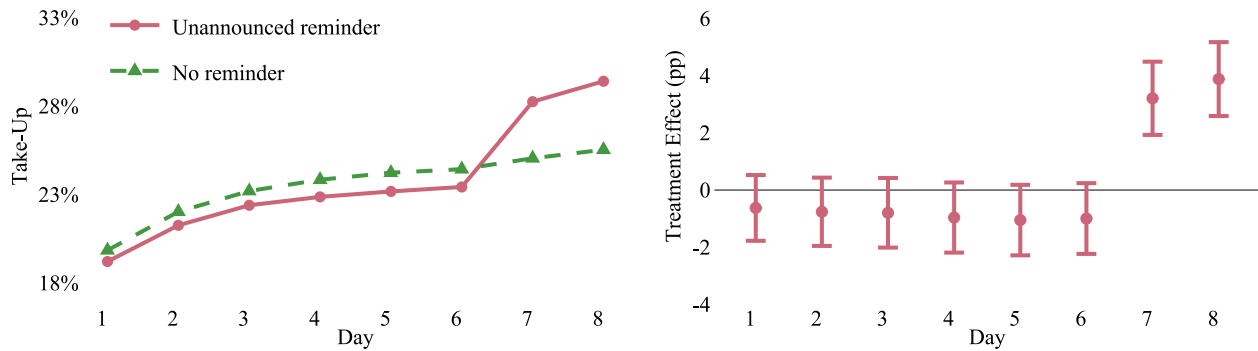
This figure shows cumulative take-up by day 8 (October 6), which is the day of the deadline for firms that had a deadline, separately for each treatment arm. The two fee arms in the upper panel of the figure are cross-randomized with the deadline and reminder arms in the lower panel. The coefficients and 95% confidence intervals are estimated using a version of model (5) where we regress cumulative take-up by day 8 on strata fixed effects and, in the upper panel, dummies for the 2.75% and 3% fee arms or, in the lower panel, dummies for the various reminder and deadline treatment arms along with the same-day deadline arm. (In both panels, the control group that did not receive an offer is the omitted dummy.) The figure uses data from the full sample of firms ( $N = 33,978$ ).

Figure C.6: Long-Term Take-Up by Treatment Arm



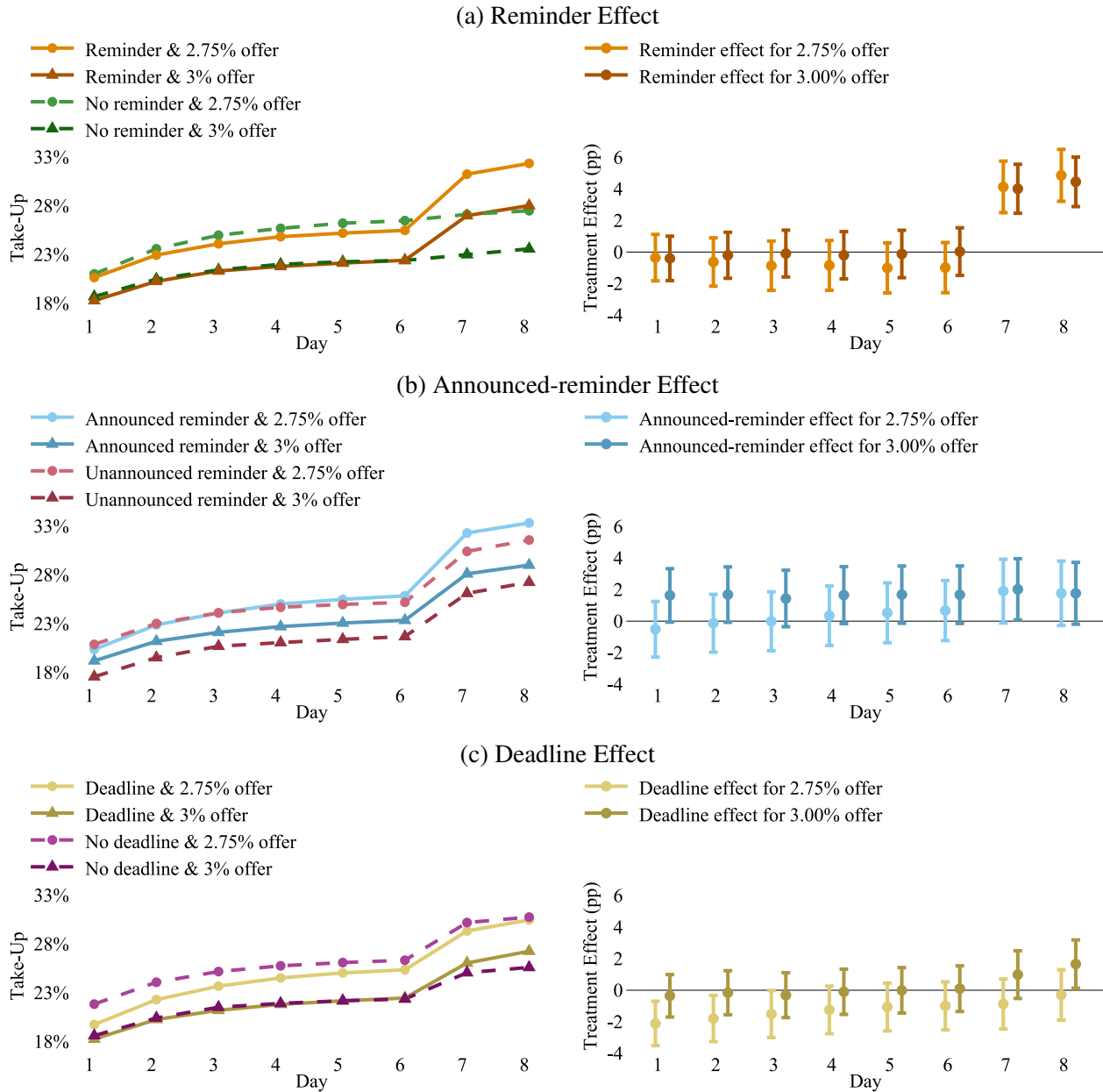
The four panels show long-term take-up by reminder, and announced reminder, deadline, and 2.75% fee groups. The unit of observation is a firm. The dotted vertical line indicates the day of the deadline. Data include take-up from September 29, 2020, to March 31, 2021. Panel (a) includes 17,220 firms with 2.75% and 3.00% offers, restricted to firms without a deadline. Panel (b) includes 12,579 firms with and without reminders, restricted to firms without a deadline. Panel (c) includes 8,124 firms with announced or unannounced reminders, restricted to firms without a deadline. Panel (d) includes 25,327 firms with and without deadlines. The sample in all panels does not include the same-day deadline and pure control groups.

Figure C.7: Treatment Effect of Unannounced vs. No Reminder



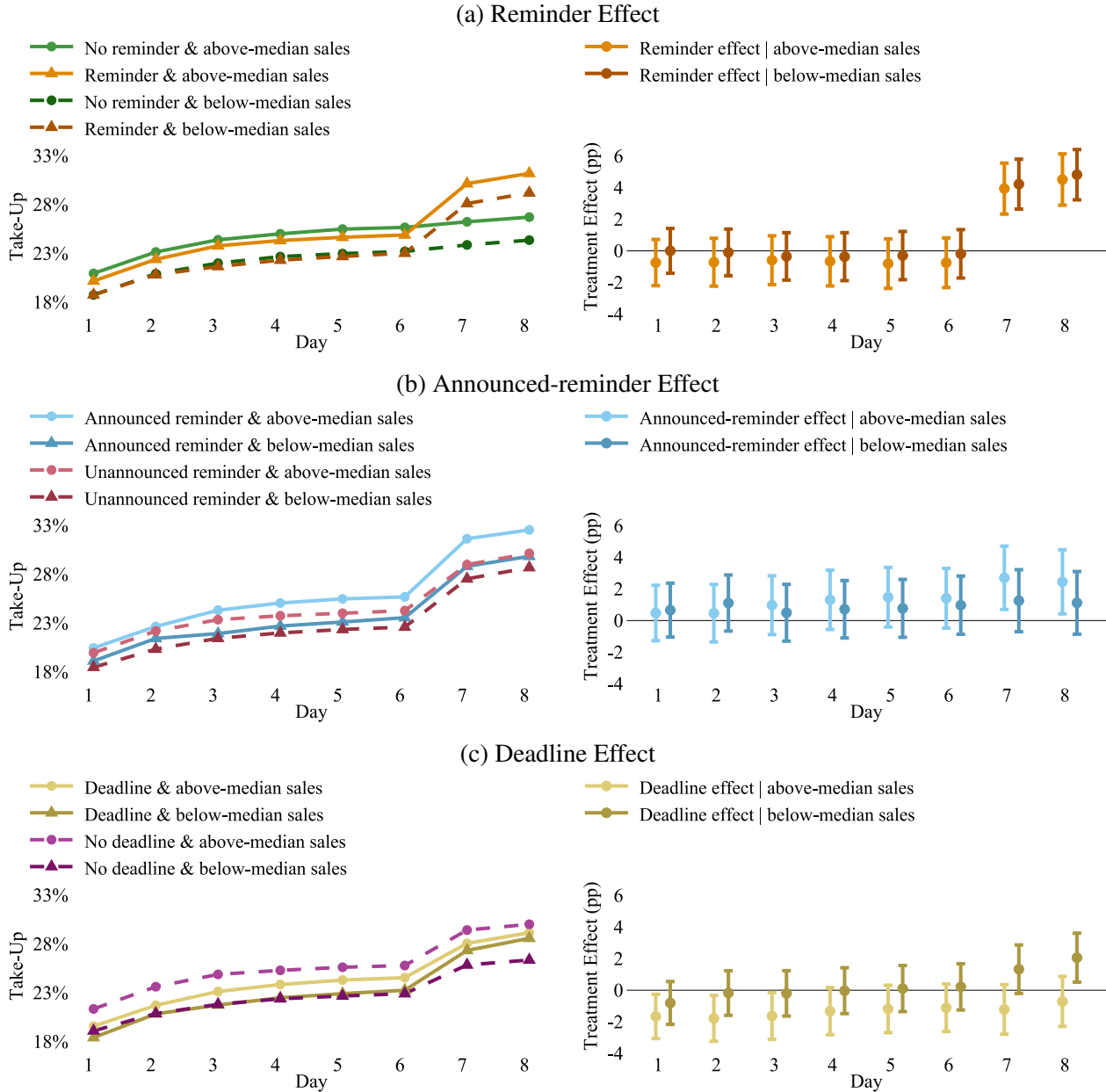
The figure shows average take-up rates (line graph on the left) and treatment effects (coefficient graph on the right) for the unannounced-reminder versus no-reminder groups, separately for each day of the experiment. The unit of observation is a firm. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on the unannounced-reminder treatment, controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). The figure includes 18,155 firms with unannounced reminders and no reminders, excluding the announced reminder, same-day deadline and pure control groups.

Figure C.8: Heterogeneous Treatment Effects by Offer Value



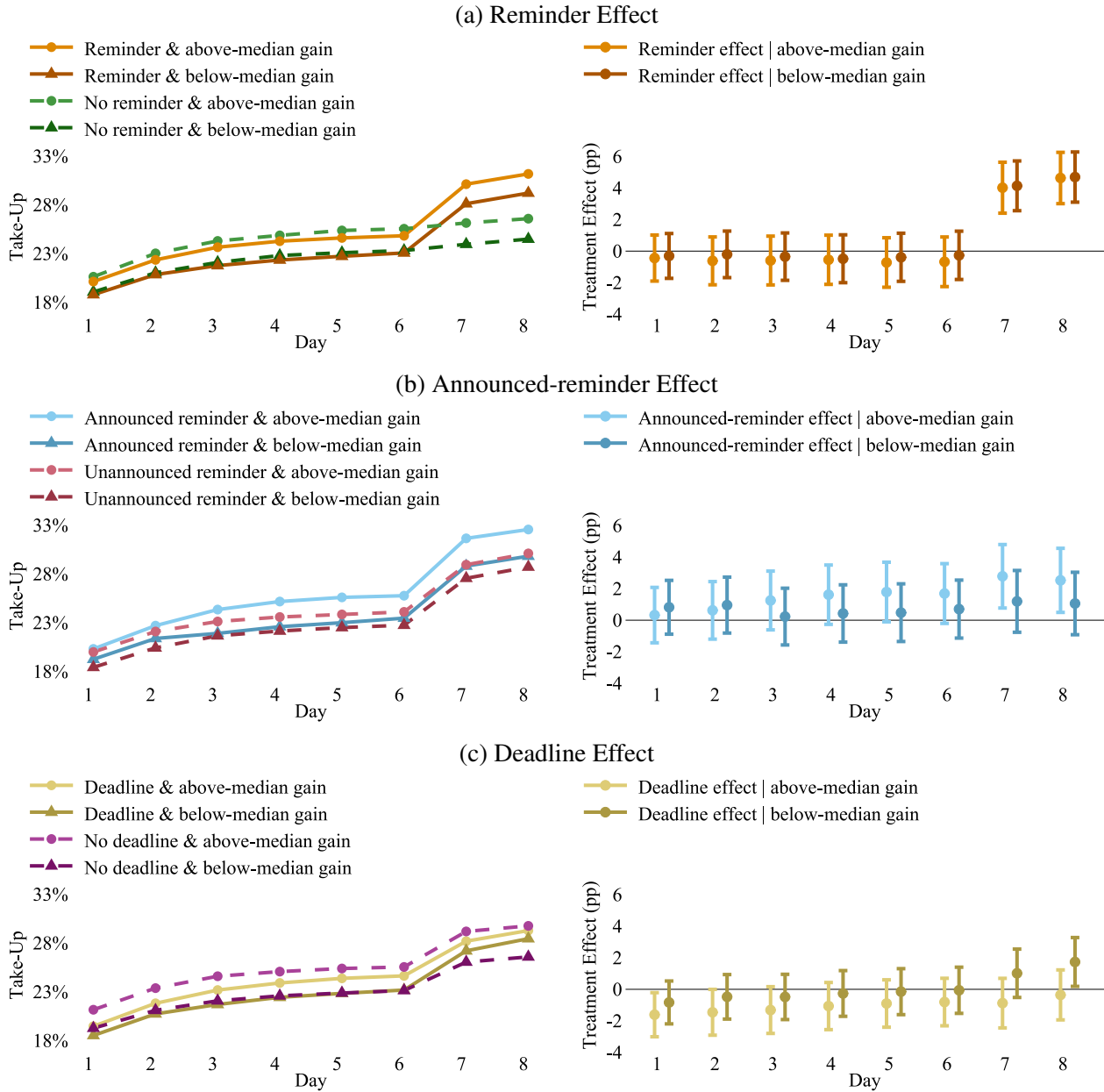
This figure shows take-up and heterogeneous treatment effects of the reminder, announced reminder, and deadline, separately for the subsamples with a 2.75% and 3% offer. The unit of observation is a firm. Line graphs show average take-up rates by treatment and offer value. Coefficient graphs show the corresponding coefficient estimates for the differential take-up of the groups, separately for each day of the experiment. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on treatment, a dummy indicating a 2.75% fee, and the interaction between treatment and the 2.75% fee dummy, controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). Panel (a) includes 25,327 firms with and without reminders, excluding the same-day deadline and pure control groups. Panel (b) includes 16,254 firms with announced and unannounced reminders, excluding firms without reminders. Panel (c) includes 25,327 firms with and without deadlines, excluding the same-day deadline and pure control groups.

Figure C.9: Heterogeneous Treatment Effects by Baseline Sales



This figure shows take-up and heterogeneous treatment effects of the reminder, announced reminder, and deadline, separately for the subsamples with above- and below-median baseline sales. The unit of observation is a firm. Line graphs show average take-up rates by treatment and baseline-sales group. Coefficient graphs show the corresponding coefficient estimates for the differential take-up of the groups, separately for each day of the experiment. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on treatment, a dummy indicating above-median baseline sales, and the interaction between treatment and the above-median baseline sales dummy, controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). Panel (a) includes 25,327 firms with reminders and no reminders, excluding the same-day deadline and pure control groups. Panel (b) includes 16,254 firms with announced and unannounced reminders, excluding firms without reminders. Panel (c) includes 25,327 firms with and without deadlines, excluding the same-day deadline and pure control groups.

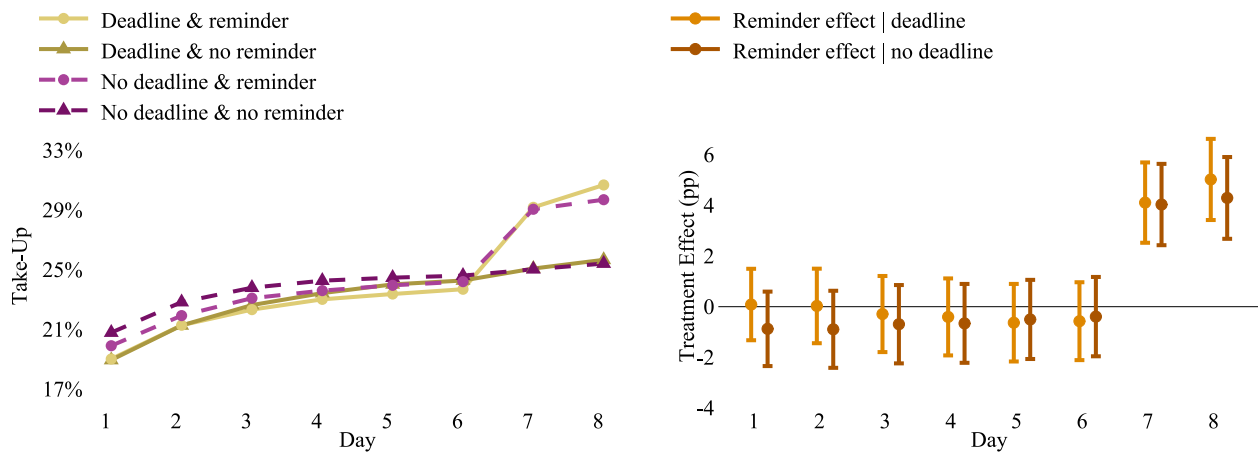
Figure C.10: Heterogeneous Treatment Effects by Expected Gain from Take-Up



The figure shows average take-up rates (line graphs on the left) and heterogeneous treatment effects (coefficient graphs on the right) by expected gain from take-up and, respectively, by reminder, announced reminder and deadline groups, and separately for each day of the experiment. The unit of observation is a firm. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on treatment, a dummy indicating above-median expected gain, and the interaction between treatment and the above-median expected gain dummy, controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). Panel (a) includes 25,327 firms with reminders and no reminders, excluding the same-day deadline and pure control groups. Panel (b) includes 16,254 firms with announced and unannounced reminders, excluding firms without reminders. Panel (c) includes 25,327 firms with and without deadlines, excluding the same-day deadline and pure control groups.

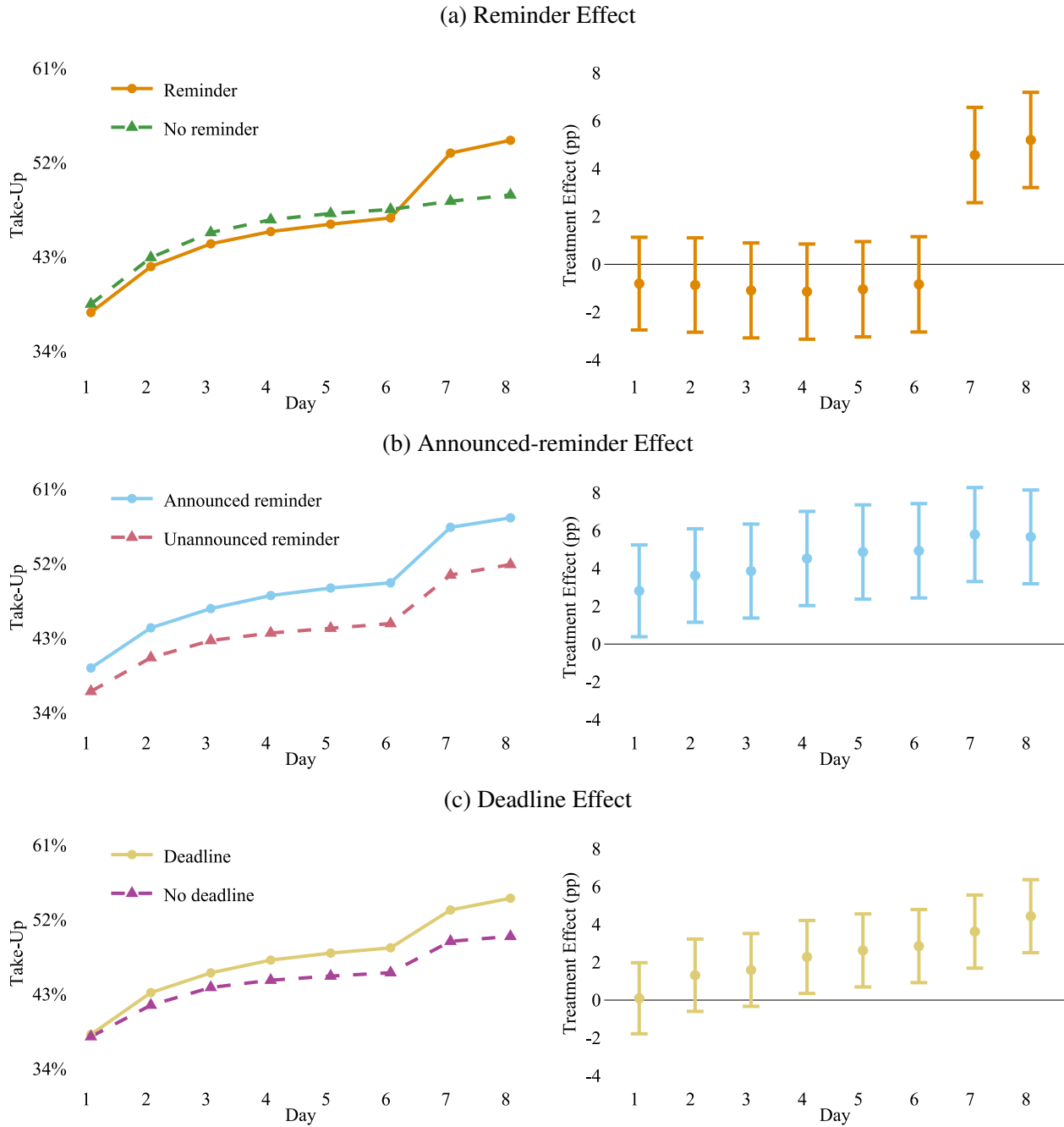


Figure C.11: Interaction Effects of Deadline and Reminder



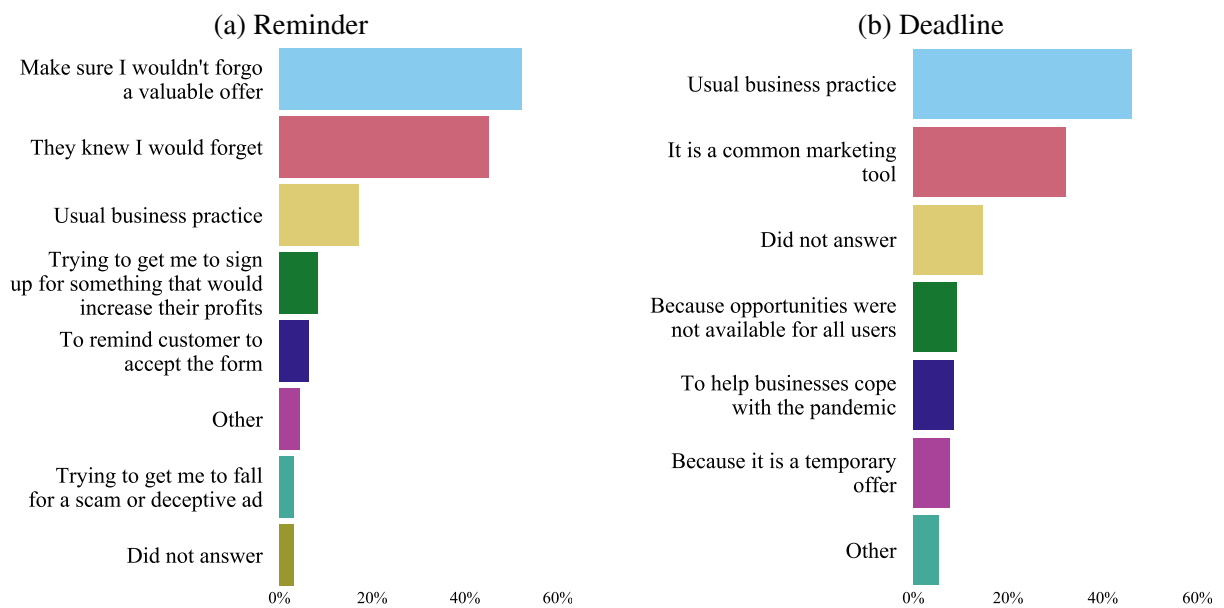
The figure shows average take-up rates (line graph on the left) and heterogeneous treatment effects (coefficient graph on the right) by reminder (includes both announced and unannounced) and deadline groups, separately for each day of the experiment. The unit of observation is a firm. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on treatment, a deadline dummy, a reminder dummy and the interaction between the deadline and reminder dummies, controlling for strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). The figure includes 25,327 firms with reminders and no reminders, excluding the same-day deadline and pure control groups.

Figure C.12: Effect of Reminder on Take-Up Conditional on Opening Email



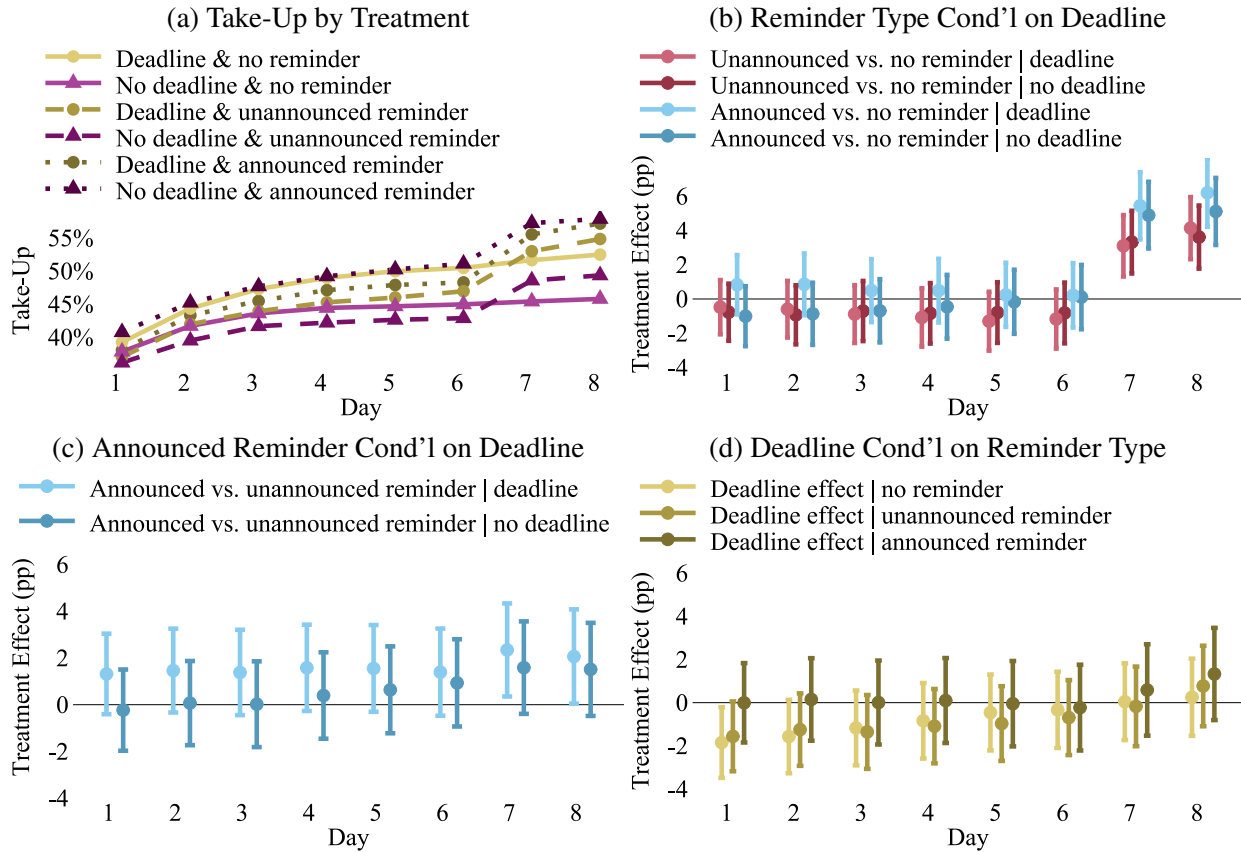
This figure shows take-up rates and treatment effects of the offer by reminder, announced reminder and deadline groups, separately for each day of the experiment, for firms that opened the email before the day of the reminder. The unit of observation is a firm. Line graphs show average take-up rates by treatment group. Coefficient graphs show the corresponding coefficient estimates for the differential take-up of the groups, separately for each day of the experiment. Coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on treatment. Regressions include strata fixed effects. Data include take-up from September 29 to October 6 (the day of the deadline). Panel (a) includes 10,246 firms with reminders and no reminders, excluding the same-day deadline and pure control groups. Panel (b) includes 6,396 firms with announced and unannounced reminders, excluding firms without reminders. Panel (c) includes 10,246 firms with and without deadlines, excluding the same-day deadline and pure control groups.

Figure C.13: Why Firms Thought the Offer Had a Deadline and Reminder



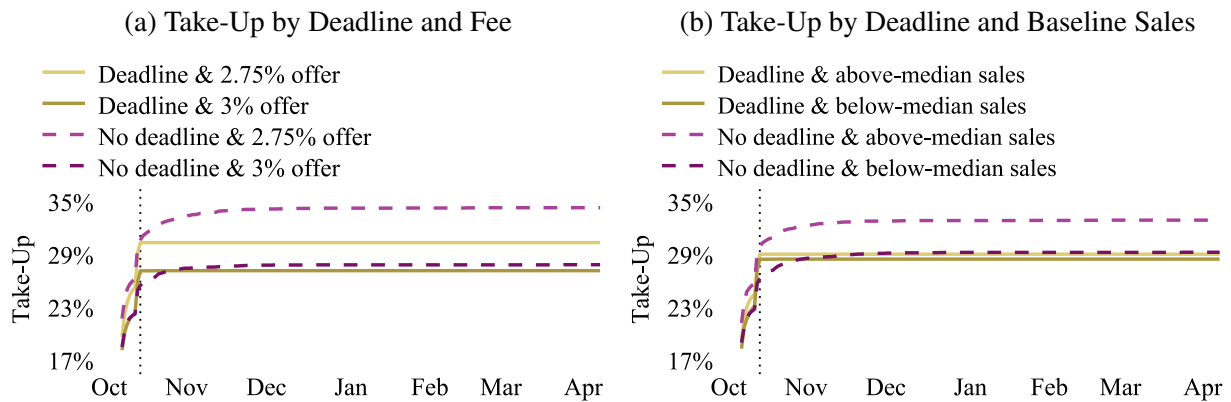
The left panel shows responses to the question: “Why do you think we sent you a reminder?” This question was asked to firms that recalled receiving the first email or SMS, were assigned to receive a reminder, did not accept the offer prior to the reminder, and recalled receiving the reminder ( $N = 157$ ). The right panel shows responses to the survey question: “Why do you think the offer had a deadline?” This question was asked to firms that recalled receiving the first email or SMS, had a deadline, and recalled that the offer had a deadline ( $N = 130$ ). Respondents could provide more than one response, so totals add up to more than 100%.

Figure C.14: Interaction Effects of Deadline and (Un-)Announced Reminder



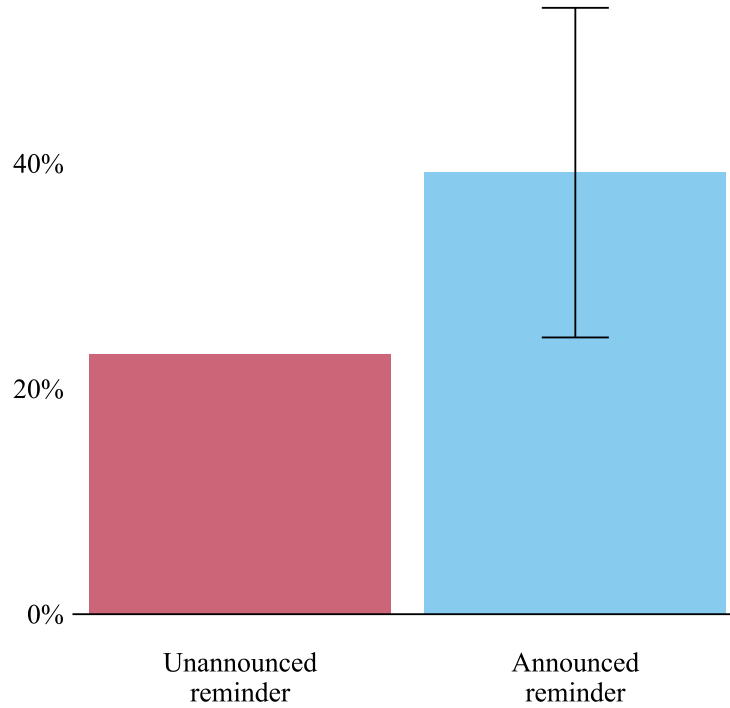
The figure shows average take-up rates (line graph in panel (a)) and treatment effects (coefficient graphs in panels (b), (c), and (d)) for various combinations of reminder type (or no reminder) and deadline (or no deadline). The unit of observation is a firm. In panel (b), the light and dark red coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on a deadline dummy, unannounced reminder dummy, and the interaction between the deadline and unannounced reminder dummies, including 18,155 firms with unannounced and no reminders, excluding the announced reminder, same-day deadline and pure control groups; the light and dark blue lines come from a regression of take-up on a deadline dummy, announced reminder dummy, and the interaction between the deadline and announced reminder dummies, including 16,245 firms with announced and no reminders, excluding the unannounced reminder, same-day deadline and pure control groups. In panel (c), coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on a deadline dummy, announced reminder dummy and the interaction of the deadline and announced reminder dummies, including 16,254 firms with announced and unannounced reminders, excluding firms without reminders. In panel (d), coefficient estimates and 95% confidence intervals come from daily regressions of cumulative take-up on a deadline dummy, announced reminder dummy, unannounced reminder dummy, and the interaction of the deadline, announced reminder, and unannounced reminder dummies, including 25,327 firms with and without deadlines, excluding the same-day deadline and pure control groups. All regressions control for strata fixed effects. Cond'l = conditional.

Figure C.15: Long-Term Take-up by Deadline, Offer Value, and Firm Size



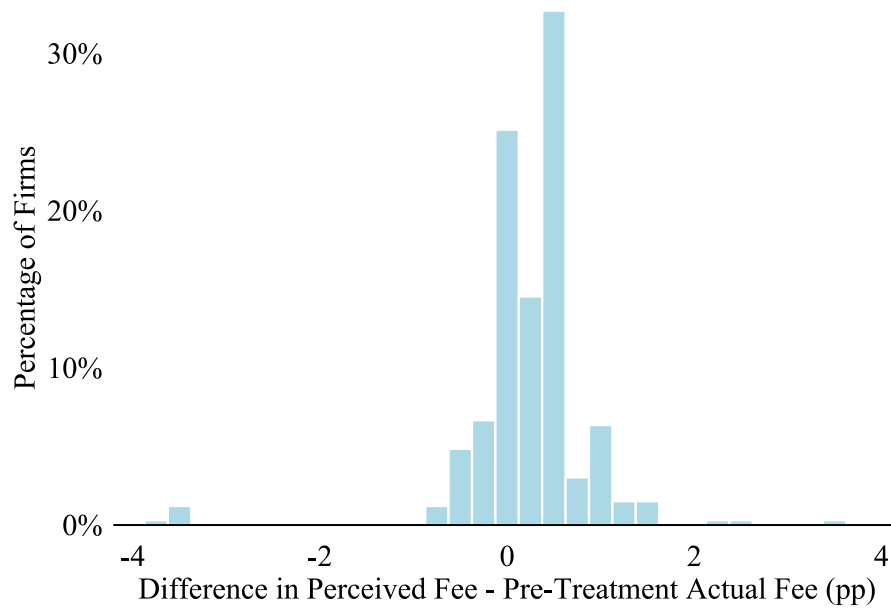
The figure shows long-term take-up by deadline, offer value and firm size. The unit of observation is a firm. The dotted vertical line indicates the day of the deadline. Data include take-up from September 29 to March 31 by 25,327 firms with and without deadlines, excluding the same-day deadline and pure control groups.

Figure C.16: Effect of Announced Reminder on Perceived Offer Value



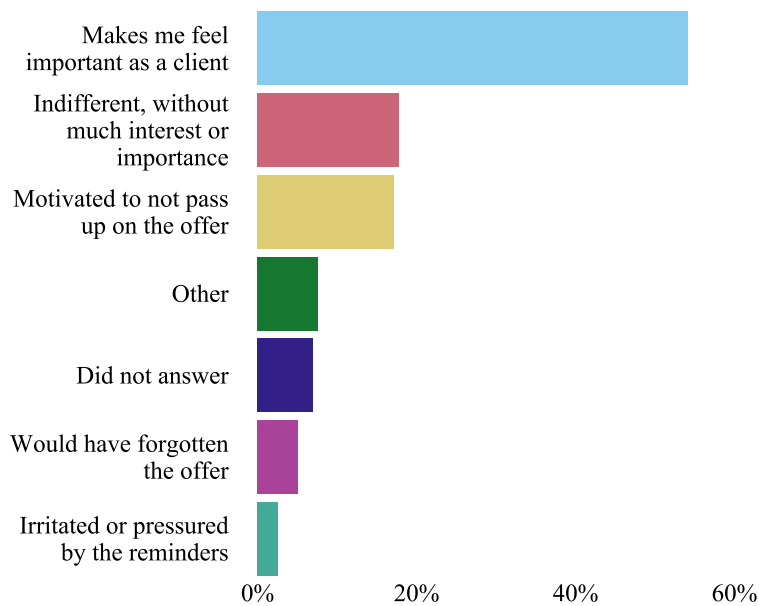
This figure contains a barplot with the percentage of firms that answered yes to the survey question “Did the reminder change your perception of the offer’s value?”, separately by reminder type. Data comes from the survey conducted on a random sample of 471 firms in the experiment, with this question asked to the 157 firms that recall receiving the first email or SMS, recall receiving a reminder, received an offer with a reminder, and accepted the offer after receiving the reminder or did not accept the offer. 5 firms that did not know the answer to the question were excluded from the sample ( $N = 152$ ). Coefficients and 95% confidence intervals come from a regression of a dummy of responding “yes” on a constant and an indicator for announced reminder, with heteroskedasticity-robust standard errors. The constant from this regression gives take-up in the unannounced-reminder arm (left bar) while the constant plus the coefficient on announced reminder gives take-up in the announced-reminder arm (right bar). We use the standard error on the announced-reminder coefficient to generate the 95% confidence interval on that bar, such that we can see that the difference in adoption between the two groups is statistically significant given that the unannounced-reminder arm’s take-up is not in the confidence interval.

Figure C.17: Difference in Pre-Treatment Actual Fee and Perceived Fee



This figure shows a histogram of the differences in firms' actual and perceived pre-treatment fees. The "difference in perceived fee – pre-treatment actual fee" is measured as the difference between a manager's survey response to the question "What was your fee with (provider) the week before you received the offer?"  $N = 330$  firms from the survey, excluding 118 firms that answered that they did not know their fee, and 23 firms that did not answer the question.

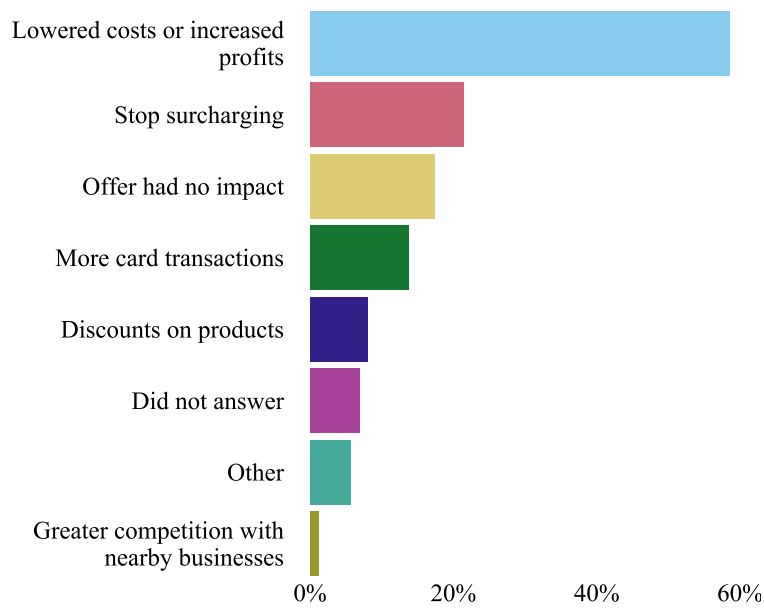
Figure C.18: How Firms Felt About Receiving Reminder



The figure contains a barplot showing how firms felt when receiving the reminder. Data comes from survey conducted on a random sample of firms in the experiment ( $N = 471$ ), with 157 firms asked this question. This question was asked to firm owners assigned to the announced or unannounced reminder arms who recalled receiving the first email or SMS, did not adopt prior to the reminder, and recalled receiving a reminder, received an offer with a reminder, and accepted the offer after receiving the reminder or did not accept the offer. Survey question: “How did receiving the reminder make you feel?” Respondents could provide more than one response, so totals add up to more than 100%.



Figure C.19: Impact of Take-Up on Business Outcomes



The figure contains a barplot with the impact the lower fee offer had on firms. Data comes from survey conducted on a random sample of firms in the experiment ( $N = 471$ ) with 248 firms asked this question. This question was asked to users who accepted the survey, and recalled accepting or clicking on the offer. Survey question: “Is this offer working for your business? What impact has it had?” Respondents could provide more than one reason for not accepting the offer on the first day, so totals add up to more than 100%.

## D Heterogeneous Effect of Announced Reminder by Length of Business Relation

In this appendix, we detail our methodology for testing whether the treatment effect of the announced reminder relative to the unannounced reminder is decreasing in the length of a firm’s business relationship with the FinTech company. We find evidence that the effect of announced reminders is decreasing in the length of the business relationship.

### D.1 Framework for Estimating Heterogeneous Announced Reminder Effect

We restrict the sample to managers who were randomized to receive either an announced reminder or an unannounced reminder. We express expected take-up as a conditional expectation, conditional on covariates and “treatment,” i.e., whether the manager received an announced reminder. We assume this conditional expectation is a sum of two polynomials: (i) a function  $f$  of the length of the business relation and the announced reminder dummy; and (ii) a function  $g$  of covariates and the announced reminder dummy:

$$\mathbb{E}[Y | X, T] = \mathbb{E}[Y | X_1, X_{-1}, T] = f(X_1, T) + g(X_{-1}, T), \quad (30)$$

where  $Y \in \{0, 1\}$  represents take-up of the profitable opportunity;  $T \in \{0, 1\}$  represents whether the reminder was unannounced or announced; and  $X \in \mathbb{R}^p$  represents a  $p$ -dimensional vector of all the observable characteristics. We divide  $X$  into  $X_1 \in \mathbb{R}$ , which represents the length of business relationship, and  $X_{-1} \in \mathbb{R}^{p-1}$ , which represents a  $(p - 1)$ -dimensional vector of observable characteristics, excluding the length of business relationship.

Following Imbens and Rubin (2015), we write the conditional average treatment effect (CATE) as

$$\tau(x) \equiv \mathbb{E}[Y(T = 1) - Y(T = 0) | X = x]. \quad (31)$$

Taking the stable unit treatment values assumption (SUTVA) and the overlap assumption (i.e.,  $0 < Pr(T = 1|X) < 1$ ) as given and noting that the announced reminder ( $T$ ) is randomly assigned, we can rewrite equation (31) as

$$\tau(x) = \mathbb{E}[Y | X_1, X_{-1}, T = 1] - \mathbb{E}[Y | X_1, X_{-1}, T = 0]. \quad (32)$$

Plugging equation (30) into equation (32) and taking predictions,

$$\hat{\tau}(x) = [\hat{f}(X_1, T = 1) + \hat{g}(X_{-1}, T = 1)] - [\hat{f}(X_1, T = 0) + \hat{g}(X_{-1}, T = 0)]. \quad (33)$$

### D.2 Model Specification

We use a penalized regression model, ridge regression, to estimate heterogeneous treatment effects of the announced reminder. We use 10-fold cross-validation (CV) to select the penalty parameter that minimizes the mean squared error. The reason we use ridge regression rather than a more black-box machine learning algorithm is that we need to extract coefficient estimates from our model to test whether the treatment effect is decreasing in the length of business relationship.

We consider various potential model specifications, where we vary the following: (i) the degree of the polynomials  $f$  and  $g$ , which can each vary from degree-one to degree-four; (ii) winsorization of the top 5% of the continuous covariates; and (iii) standardization of the covariates.<sup>26</sup> Henceforth, a notation like  $(2, 1)$  represents a model configuration where the polynomial degree of  $f$  is 2, and that of  $g$  is 1. We evaluate which of these potential models performs best using the out-of-sample root mean squared error (OOS RMSE).

Specifically, we calculate the OOS RMSE from 64 distinct ridge regression models (where  $64 = 4$  potential degrees for  $f \times 4$  potential degrees for  $g \times 2$  winsorized or not  $\times 2$  standardized or not). For each model, this involves a four-step process:

1. Divide our dataset into a training set and a hold-out test set using a 9:1 ratio.
2. Train our model on the training set and perform 10-fold CV within the training set.
3. Use the trained model to predict the outcome (take-up) on the test set.
4. Compute the OOS RMSE by comparing predicted take-up to observed take-up for each observation in the test set.

To compare the 64 models and account for randomness in the process of splitting the data, we repeat the above process using 300 distinct seeds, which determine distinct random splits into training and testing sets.<sup>27</sup> We then compare the percentage of instances (across the 300 seeds) in which each model outperforms the benchmark non-winsorized and non-standardized  $(1, 1)$  model, measured by OOS RMSE. Figure D.1 shows the results of this comparison. Models incorporating higher-degree polynomials, such as the  $(\cdot, 4)$  models, generally exhibit poorer performance (and to an even greater extent for the standardized models). This is likely due to overfitting. Among the 64 models, the  $(2, 2)$  model with no winsorization or standardization yields the best fit on average; thus, we select this model for the remainder of the analysis.

### D.3 Results

Using the  $(2, 2)$  ridge regression model without winsorization or standardization, we estimate equation (33). We test whether the treatment effect of the announced reminder relative to the unannounced reminder is decreasing in the length of business relationship. In other words, we test whether  $\hat{\tau}(x)$  is decreasing in  $X_1$ . Specifically, we test whether

$$\frac{\partial \hat{\tau}(x)}{\partial X_1} = \frac{\partial \hat{f}(X_1, T = 1)}{\partial X_1} - \frac{\partial \hat{f}(X_1, T = 0)}{\partial X_1} \quad (34)$$

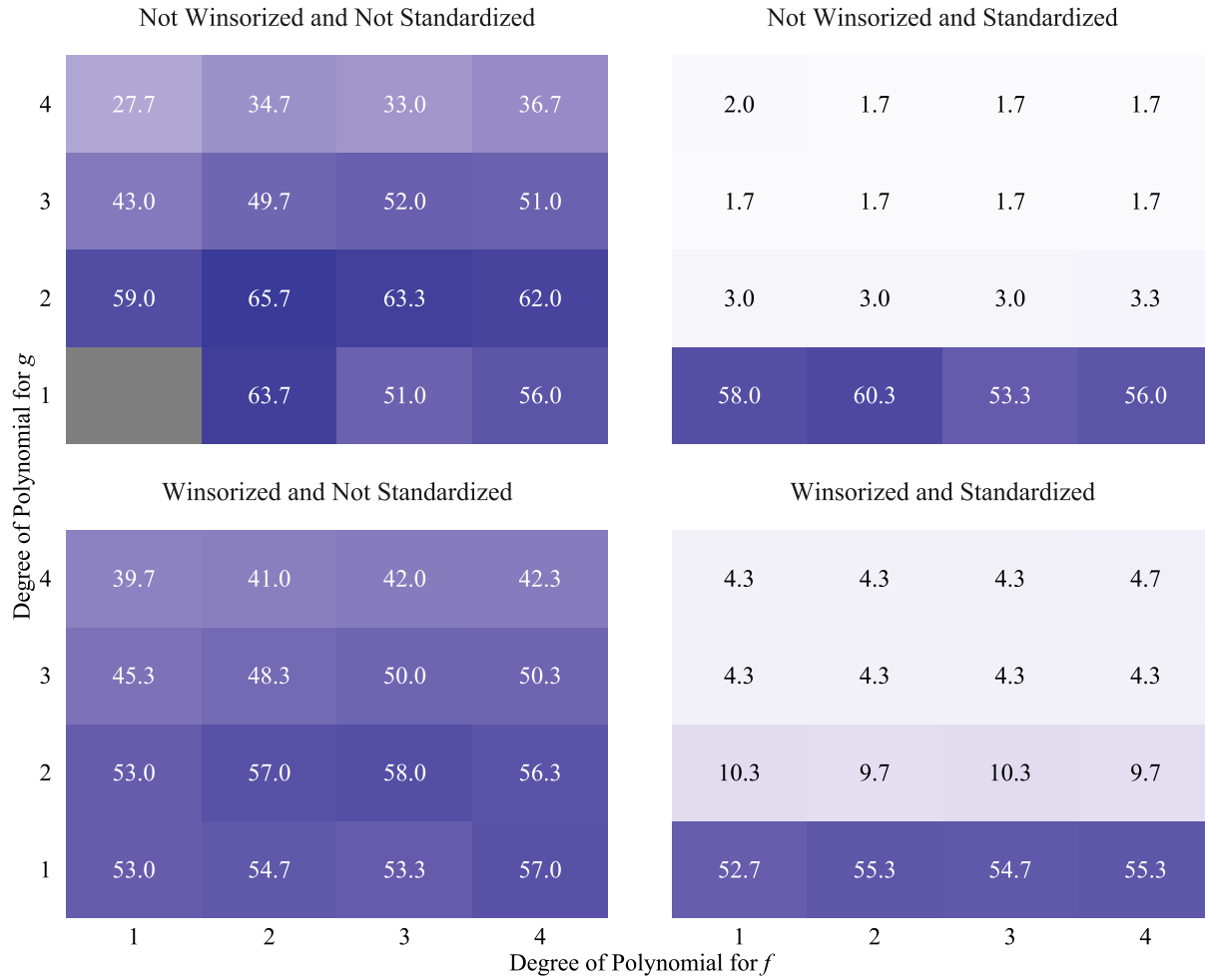
is negative at various values of  $X_1$ , including the 25th percentile (11.8 months), median (20 months), mean (24.2 months), and 75th percentile (33.1 months).

---

<sup>26</sup> We do not consider polynomials higher than fourth-order because the fourth-degree polynomials already have substantially worse out-of-sample fit than lower-order polynomials due to overfitting. We do not winsorize the bottom 5% of the continuous covariates as they are all bounded below by 0. Standardization converts each covariate to a Z-score by subtracting the sample mean and then dividing the demeaned variable by the standard deviation.

<sup>27</sup>The additional splits of the training data done during 10-fold CV are also affected by the seed.

Figure D.1: Percent of OOS RMSEs lower than benchmark model OOS RMSE



This figure shows the percentage of out-of-sample root mean squared errors (OOS RMSEs) for the given models that are lower than those of the benchmark non-winsorized and non-standardized (1,1) model. Using 300 distinct randomization seeds, which determine 300 different training and test set splits, we compute the corresponding 300 OOS RMSEs for each model. We then compare the 300 OOS RMSEs of a given model with the 300 OOS RMSEs from the benchmark model and calculate the percent of times, out of 300, that the OOS RMSE of the given model is smaller than that of the benchmark model. For example, the 57.0% for the winsorized and not standardized (4,1) model indicates that this model has a lower OOS RMSE than the benchmark model for 171 distinct randomization seeds out of the 300 total seeds.

To formally test whether (34) is negative, we employ a bootstrapping method with 1,000 iterations, and document the percentage of bootstrap iterations in which (34) is negative. Algorithm 1 describes our exact bootstrapping procedure.

---

**Algorithm 1** Bootstrapping Algorithm

---

**Require:** Original data set  $\mathcal{D}$  of size  $|\mathcal{D}|$  and number of resamples  $B = 1000$

- 1: **for**  $b = 1$  to  $B$  **do**
  - 2:   Sample with replacement from  $\mathcal{D}$  to get  $\mathcal{D}_b$  of size  $|\mathcal{D}_b| = |\mathcal{D}|$ .
  - 3:   Estimate  $\hat{f}(X_1, T = 1)$  and  $\hat{f}(X_1, T = 0)$  in equation (33) from  $\mathcal{D}_b$ .
  - 4:   Compute  $\partial \hat{\tau}(x) / \partial X_1$  using equation (34).
  - 5:   Evaluate  $\partial \hat{\tau}(x) / \partial X_1$  at the 25th percentile, median, mean, and 75th percentile of  $X_1$  in  $\mathcal{D}$ .
  - 6: **end for**
  - 7: Compute the percentage of bootstrap iterations out of  $B = 1000$  total bootstrap iterations in which  $\partial \hat{\tau}(x) / \partial X_1 < 0$ , evaluated at the 25th percentile, median, mean, and 75th percentile of  $X_1$  in  $\mathcal{D}$ .
- 

We find that the derivative is negative for 93.3%, 92.5%, 90.5%, and 81.3% of the iterations at the 25th percentile, median, mean, and 75th percentile of  $X_1$ , respectively. We conclude that the effect of the announced reminder on take-up of the profitable opportunity is higher for managers who have been using the FinTech payments technology for less time and thus likely have less trust in the FinTech company.

**Appendix References**

Imbens, Guido and Donald B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York, NY: Cambridge University Press.