Buy Baits and Consumer Sophistication: Theory and Field Evidence from Large-Scale Rebate Promotions

Matthias Rodemeier*

This version: July 2022
First version: November 2020

Abstract

Are consumers in the marketplace aware of their behavioral biases? Which biases can be profitably exploited by firms? I answer these questions in the context of a widely regulated form of price discrimination: rebates that require active redemption. I show theoretically how to recover consumers' subjective beliefs about their biases—that prevent them to redeem the rebate—from aggregate demand responses to rebates, redemption reminders, and a simple discount that does not require redemption. In a large-scale field experiment with a major online retailer, I find that consumers correctly increase demand when the firm offers a redemption reminder, but they fail to reduce demand when the firm increases the hassle required to redeem. Structural estimates reveal that, while consumers are almost fully sophisticated about the probability of *forgetting* to redeem the rebate, they vastly underestimate the *hassle* of redeeming it by 20 EUR per consumer. Exploiting this misperception increases the profitability of rebates by 150%.

JEL Codes: D18, D61, D83, D49, D91, L21

Keywords: consumer sophistication, consumer protection, price discrimination, behavioral public economics, field experiments

---

Acknowledgments: I thank Alec Brandon, Leonardo Bursztyn, Luigi Butera, Felix Chopra, Markus Dertwinkel-Kalt, Nicola Gennaioli, Andreas Gerster, Sam Hirshman, Justin Holz, Kilian Huber, Chris Knittel, Botond Kőszegi, Thibaut Lamadon, John List, Andreas Löschel, Rob Metcalfe, Maximilian Mueller, Nadine Riedel, Carlo Schwarz, Frederik Schwerter, Nicolas Serrano-Velarde, Dmitry Taubinsky, Jurre Thiel and seminar participants at Aachen, Amsterdam, CBS, CEU, CERGE-EI, Chicago, Frankfurt School, Innsbruck, Louvain, TSE, Mannheim, and ZEW for helpful discussions that improved this paper. The pilot studies have been pre-registered at the AEA RCT registry under trial numbers AEARCTR-0003150 and AEARCTR-0004475. The main field experiment including the theoretical model has been pre-registered under trial number AEARCTR-0005830. All remaining errors are mine.
*Bocconi University, Becker Friedman Institute, IGIER, and IZA, rodemeier.matthias@gmail.com.

"These companies are only offering the illusion of a rebate to the many people like me who never get around to claiming it. Because of such thick sludge, redemption rates for rebates tend to be low, yet the lure of the rebate still can stimulate sales—call it 'buy bait.'"

- Richard H. Thaler, 2018

1 Introduction

Growing evidence in psychology and economics shows that behavioral anomalies affect consumption choices. Consumers might ignore shrouded shipping costs when purchasing products online, they might over-borrow once they have a credit card, or they might fail to cancel expensive subscriptions when the hassle of cancellation seems large. Are consumers able to anticipate these behavioral shortcomings? Answering this question is of fundamental importance because it determines whether consumers can avoid environments that try to exploit them. A consumer who is aware of her inattention might avoid online shops that shroud shipping costs. She might decide not to get a credit card if she is aware of her time inconsistency. And she might avoid subscriptions if she anticipates the hassle required to get out of the contract. Consumer sophistication about their behavioral tendencies also determines whether firms can profitably exploit behavioral biases, whether biases actually lead to systematic mistakes, and whether consumer protection regulation is justified.

While there is some laboratory evidence that measures consumers' beliefs about their behavioral biases, there is little evidence as to whether consumers in the marketplace are sophisticated or not. I provide the first revealed-reference test of consumer sophistication from a "natural field experiment" (Harrison and List 2004): real market participants make choices in their natural environment, not knowing they are being observed by a researcher. The experiments is conducted with one of Europe's largest retailers for furniture and homeware, and gathers choices from over 600,000 consumers. The study exploits an ideal setting that is often criticized for exploiting behavioral biases and, for that reason, is widely regulated by consumer protection laws: online rebates that require active redemption. In these promotions, consumers are offered a price discount, but the discount is only applied to the purchase value if consumers enter the promotion code into a field on the checkout page of the web shop.[1]

This setting is interesting because consumers may fail to claim the rebate for two reasons. First, they may forget, because reminders to redeem are typically not provided. It is often argued that this friction, which I refer to as "inattention," may be a source of profits for firms. Second, even if consumers remember rebate redemption, they need to go through the effort of redeeming it, namely

[1]Digital rebates are the main promotion tool in online shopping and quantitatively important. Industry experts estimate that the total value of digital rebates that have been redeemed was 47 billion USD in the United States in 2017 (Juniper Research 2021). This number is projected to increase to 91 billion USD in 2022. Digital rebates account for around 80% of all rebates in terms of redemption value, superseding traditional promotions such as mail-in rebates.

finding and entering the promotion code. The consumer may then fail to redeem also because of such "hassle costs." Ex ante underestimation of hassle costs may be another source of profits for the firm. Distinguishing between these two frictions is also important in light of their starkly different implications for consumer protection policies: mandating firms to provide reminders to redeem the rebate would not prevent consumers from making mistakes if they are sophisticated about their inattention but underestimate hassle costs. My field setting is ideal to study the importance of this issue. The retailer I partner with routinely offers these rebates, so I can assess whether consumers in the marketplace, who have an opportunity to learn, are nevertheless unsophisticated with respect to their inattention, to hassle costs, or both.

Assessing whether consumers are sophisticated is not straightforward and in particular it is not enough to document that redemption rates are low. As long as the consumer is aware of her low redemption probability, she will respond less to the rebate than to a simple price reduction and the firm may not be able to profit from a low redemption rate. To evaluate sophistication we must measure consumers' *perceived* redemption probability and compare it to the true redemption rate.

In the first part of the paper, I show theoretically that consumers' perceived and true inattention and hassle costs can be recovered from observed choices in a small set of treatments: a standard rebate ("rebate" henceforth), a rebate in which a salient reminder to redeem is announced at the outset ("reminder" henceforth), and a simple price reduction that does not require redemption ("discount" henceforth).

Consider first how these treatments allow us to estimate the *true* extent of inattention and of hassle costs. I show that the true probability to remember redemption is identified by the ratio of redemption rates with and without the reminder (i.e. "rebate" vs. "reminder"). Using this method, I find that consumers remember to redeem 78% of the time. I then show that the true hassle cost felt by the consumers can be estimated using the minimum increase in rebate value at which all consumers redeem the rebate when they are reminded to do so. Using exogenous variation in rebate value, I estimate that hassle costs are around 20 EUR. This is a substantial amount, given that consumers only need to find and enter a promotion code.

I then proceed to show how to estimate consumers' sophistication. Consumers' subjective belief about inattention and hassle costs can be estimated by observing how the buying probability, which I refer to as "demand," responds to the treatment variation. Consumers' beliefs about inattention can be identified by comparing the demand response to "rebate" to the demand response to "reminder." Using this identification strategy, I structurally estimate that the perceived probability of remembering rebate redemption is around 73%. This is close to the true probability to remember, which, as previously discussed, is equal to 78%. Since perceived and true values are roughly equal, I conclude that consumers almost fully anticipate their inattention when deciding whether to buy. I show that this finding replicates at higher stakes as I exogenously vary the rebate value.

Importantly, the picture is very different for hassle costs. Perceived hassle costs can be estimated by comparing the demand response to "reminder" to the demand response to "discount." I find that there is no significant difference in the demand responses to these treatments, even though the actual redemption probability is 20 percentage points lower for the rebate with the reminder than for the automatically-applied discount. This implies that, despite hassle costs being large, there is full naiveté about them. The magnitude is economically large: I find that the optimal demand response to rebates of fully sophisticated consumers would only be 30% of the observed treatment effect.

There is little heterogeneity in treatment effects based on observables. This is important because rebates could be a neoclassical form of price discrimination if more price elastic consumers have higher redemption probabilities. However, I find that consumers visiting from regions with different income levels have fairly homogenous price elasticities, that are independent from inattention and hassle costs. These additional results indicate that rebates are unlikely to constitute a traditional form of price discrimination, but are rather used to exploit naiveté about hassle costs.

The extent of naiveté about hassle costs has an important implication: firms can introduce hassle costs, which strongly reduce redemption rates, without sacrificing sales. Using detailed data on markups, I show that the typical rebate is 150% more profitable than an automatically applied discount of equal value. This effect seems to be entirely driven by consumers not anticipating hassle costs. Exploiting inattention, by not offering a reminder, has no significant positive effect on profits. The reason is that consumers anticipate their inattention and reduce demand if the firm does not offer a reminder.

These results inform consumer protection policies in the US and across the globe. Policymakers have substantially regulated the features of rebates due to the suspicion that consumers are not sophisticated.[2] This paper provides the empirical foundation for these widely used regulatory policies and substantiates the underlying motivation. An important takeaway is that mandatory reminders may not be sufficient to protect consumers, because naiveté is concentrated around hassle costs rather than inattention. More invasive regulations that restrict the use of rebates or impose the burden of redemption on the side of the firm may be called for.[3]

The rest of this paper is structured as follows. Section 2 discusses the contributions of this paper to the literature. In Section 3, I present a simple model of consumer responses to rebates

[2]In the US, the Federal Trade Commission warns consumers to not be "'baited' by rebates that never arrive" (Federal Trade Commission 2020). The regulation of these promotions lies in the responsibility of each state. The Connecticut "Unfair Trade Practices Act" allows firms to advertise a price net of the rebate only if the rebate is immediately applied (Connecticut Unfair Trade Practices Act Regulations 2004). The "Unfair Trade Practice and Consumer Protection Act" in Rhodes Island explicitly states that the burden of rebate redemption should be on the side of the retailer if prices are advertised net of the rebate (Rhode Island General Laws 2016). California, Maryland, and New Jersey have similar regulations that illustrate the regulators' concerns that consumers may be naive about their probability to redeem.

[3]An additional policy that this paper does not speak to is mandatory information provision about hassle costs. Firms could be mandated to warn consumers that redemption is more cumbersome than many consumers assume ex-ante. It is unclear, however, whether these warnings would be able to fully de-bias consumer beliefs about hassle costs and how these warnings could be implemented in practice. Note also that firms have a lot of control over the shopping environment and could, therefore, frequently change the details of the redemption procedure in a way that is difficult to inform consumers about concisely.

and derive empirically testable predictions of sophisticated behavior. The model guides the field experimental design discussed in Section 4. Reduced-form results are presented in Section 5. In Section 6, I estimate the structural parameters of sophistication and discuss policy implications. Section 7 discusses additional mechanisms that can explain the data. Section 8 concludes.

2 Contributions to the Literature

This paper makes four contributions. First, it provides the first test of consumer sophistication from a natural field experiment. The experiment leverages a popular platform in which rebates are frequently offered and I observe how actual market participants respond to them. Two prior studies have provided clean tests of sophistication about redemption behavior but relied on laboratory experiments. Ericson (2011) lets students choose between two future payments, where one payment is certain and the other payment is only received if the subject remembers to send an email to the author 5 months later. A multiple price list elicits a lower bound for subject's perceived probability of sending the email. While this lower bound corresponds to 76%, only 53% of subjects end up sending the email. Tasoff and Letzler (2014) run a related experiment, in which students need to file a mail-in form rather than send an email. Subjects' lower bound is 79% but only 30% of subjects eventually mail in the form.[4,5]

The field experiment conducted in my study circumvents the potential criticism that subjects' choices do not generalize to actual market settings. For instance, one may argue that subjects in the prior studies are overconfident because they have no experience with the particular redemption procedure. Prior work also documents that market experience can eliminate certain behavioral biases (List 2003, List 2011, Allcott et al. 2020). The rebate used in my study is the standard promotion used by the firm and its competitors, and therefore exists in a long-run general equilibrium. One important takeaway from this study is, therefore, that also consumers in the marketplace are partially naive.[6]

The second contribution is the insight that while both inattention and hassle costs are substantial frictions, consumers fully anticipate their inattention while they severely underestimate hassle costs. A particular novel result is that seemingly minor hassle costs cause substantial, but fully

[4] An information treatment and an email reminding subjects to file the form neither increase redemption, nor do they alter beliefs.

[5] Prior work in marketing indicates that subjects overestimate their redemption probability, but the elicitation methods rely on non-incentivized survey questions in which subjects are asked how often they forget to redeem rebates (Jolson, Wiener, and Rosecky 1987, Silk 2004). Self-reported and non-incentivized predictions may suffer from several reporting biases that have been documented in the literature (see, e.g., Cummings, Harrison, and Rutström 1995, List and Gallet 2001). In fact, Ericson (2011) shows in his experiment that unincentivized statements on subjective redemption probabilities differ substantially from lower bounds elicited in the multiple price list.

[6] Generalizability may, of course, always be attenuated by the fact that firms can design the shopping environment in different ways, such that the presented results do not translate to competitor shops. However, the redemption procedure of the rebate in this study is the standard practice of both the firm and of most of its competitors.

unanticipated, disutility to consumers. The results also advance our general understanding of which behavioral biases consumers are aware of. My structural estimates on perceived and true inattention could be used to quantify parameters of leading models on (rational) inattention (Sims 2003, Bordalo, Gennaioli, and Shleifer 2013, Kőszegi and Szeidl 2013, Matějka and McKay 2015 Caplin, Dean, and Leahy 2017) and memory (Bordalo, Gennaioli, and Shleifer 2020).[7] The small number of prior studies that identify structural parameters of agents' sophistication about their behavioral biases mostly do so in the context of present bias (Augenblick and Rabin 2019, Allcott et al. 2020, Bai et al. 2017, Carrera et al. 2019, Chaloupka IV, Levy, and White 2019). In many contexts, insufficient demand for commitment devices is interpreted as evidence of naiveté about time-inconsistent behavior.[8] A recent paper by Miller, Sahni, and Strulov-Shlain (2022) investigate sophistication regarding future inertia in canceling newspaper subscriptions and finds that subscription demand is larger, but not large enough, when an automatic cancellation feature is offered. Different to my study, they do not distinguish between various sources of naiveté. The source of naiveté turns out to be important in my setting as consumers are well aware of their inattention but fully ignorant of their hassle costs, resulting in different welfare and policy implications. Another paper by Bronchetti et al. (2020) elicits subjects' valuation for a reminder using a multiple price list. While average willigness to pay is positive, valuations are still below the true returns of being attentive. This provides evidence that subjects in the study do not allocate attention fully rationally.

The third contribution is to provide evidence that firms have an incentive to exploit the lack of sophistication. I find that profits more than double relative to automatically applied discounts. The fact that these rebate promotions exist in equilibrium implies that firms offering rebates are not driven out of the market due to competition. This is consistent with the theoretical prediction that competition does not eliminate exploitation when at least some consumers are naive (Gabaix and Laibson 2006). In this context, the paper also provides the first empirical evidence from a natural field experiment for the theoretical literature on firm practices of exploiting consumer naiveté (DellaVigna and Malmendier 2004, Grubb 2009, Heidhues and Kőszegi 2010, Kőszegi 2014, Heidhues and Kőszegi 2017). Most theory papers on naiveté-based exploitation are motivated by descriptive facts of real-world contracts or common anecdotes, but hard empirical evidence is scarce. The first empirical study comes from DellaVigna and Malmendier (2006) and uses observational data on gym membership contracts. They show that many gym members lose money by choosing flat-rate contracts, likely because they overestimated their gym attendance ex ante. Grubb and Osborne (2015)

[7]This paper also joins a rapidly growing literature that attempts to identify deep primitives of economic models motivated by insights from psychology (see DellaVigna 2018 for an overview).

[8]Löschel, Rodemeier, and Werthschulte (2020) estimate willingness to pay for "soft" commitments such as self-set goals in the context of time inconsistency.

estimate a structural model of cellular plan choice with observational data and find that consumers choose too risky plans because they underestimate the variance of future calling behavior.[9]

To the best of my knowledge, my study provides the first evidence from a natural field experiment on the economic implications of behaviorally-motivated price discrimination for consumer welfare and firm profits.[10,11]

As a fourth and final contribution, the paper provides a novel identification strategy, which point-identifies consumers' subjective behavioral biases based on observed demand responses. Instead of relying on willigness to pay elicitations with multilpe price lists, this study identifies consumer beliefs from observed demand responses to rebates. This allows me to explicitly test whether the demand response to rebates is excessively high. The identification strategy extends Chetty, Looney, and Kroft (2009)'s seminal sufficient statistics approach of identifying behavioral biases. The novelty is to show that, by observing two choice margins (demand and rebate redemption), one may identify not only the behavioral bias but also consumers' beliefs about this bias.[12,13] The approach may be applied more generally to any setting in which two or more choices require some consistency.

3 Theoretical Framework

3.1 A Simple Model of Consumer Responses to Rebates

In this section, I develop a simple model of consumer behavior and derive an empirical test of consumer sophistication. The model has been pre-registered at the AEA RCT Registry under trial ID AEARCTR-0005830 and directly produced the experimental design.[14]

[9] Gottlieb and Smetters (2012) argue—based on stylized facts and a theoretical model—that demand for life insurance plans with high cancellation fees is best explained by consumers who do not correctly anticipate future liquidity needs. Ausubel (1991) hypothesizes that a likely explanation for high interest rates in credit card markers is that consumers underestimate their probability of borrowing.

[10] Rebate promotions are distinct from well-studied firm practices of shrouding fixed product surcharges, as empirically analyzed in the context of shipping surcharges (Hossain and Morgan 2006, Brown, Hossain, and Morgan 2010), sales taxes (Chetty, Looney, and Kroft 2009, Taubinsky and Rees-Jones 2017), and hidden fees of cinema tickets (Dertwinkel-Kalt, Köster, and Sutter 2020). Shrouding surcharges simply makes a part of the total price less salient, but the actual price for a given product remains the same for all consumer types. By contrast, rebate promotions vary the price of the goods as a function of the behavioral bias. Consumers do not simply make mistakes because costs are not salient, but because they do not correctly predict their own behavior (i.e., rebate redemption).

[11] Gilpatric (2009) develops a theoretical model to study the implications of time inconsistency for rebate redemption rates and firm profits.

[12] This extension has important implications for the welfare effects of behavioral frictions and the design of policies aimed at correcting these frictions. Because sophisticated consumers do not make mistakes in expectation, corrective policies may introduce substantial inefficiencies if policymakers ignore the degree of sophistication, thereby implicitly assuming full naiveté. See Bernheim and Taubinsky (2019) for an overview of policies aimed at correcting behavioral biases.

[13] Prior studies identify lower bounds of the subjective redemption probabilities, while the sufficient statistics approach used in this paper provides point estimates.

[14] The notation in the the pre-registration model is less concise. The model predictions are the same.

In the most concise version of the model, the consumer faces two choices.[15] First, she chooses where to buy, $b \in \{0, 1\}$, where $b = 1$ means she buys at a location I call "the store," and $b = 0$ means she chooses the outside option. The outside option may represent one or multiple competitors of the store, but it may also mean the consumer does not buy anywhere. The store offers a rebate of value s that is only redeemed if the consumer actively claims it.

Second, she decides whether to claim the rebate conditional on buying at the store. Let her rebate redemption choice be represented by $r \in \{0, 1\}$, where $r = 1$ means she redeems and $r = 0$ means she does not redeem.

Claiming the rebate causes hassle costs c. In the field experiment, the consumer has to search for a discount code and enter it into a respective promotion code field during the checkout page. Claiming the rebate may also require effort because some consumers with less experience in online shopping need to understand the details of how to redeem.

Another redemption friction is inattention: the consumer may eventually forget to claim the rebate even if redemption would have been worth the hassle. Her probability of remembering rebate redemption is given by $\theta \in [0, 1]$.

I allow for the possibility that the consumer is not perfectly aware about these frictions when deciding whether to buy at the store. Let $\hat{c}$ and $\hat{\theta}$ be the consumer's *perceived* hassle costs and probability of remembering to redeem the rebate, respectively. The difference between the perceived and true values of inattention and of hassle costs measures the degree of sophistication. A consumer is said to be *sophisticated* about her redemption frictions if and only if $\hat{\theta} = \theta$ and $\hat{c} = c$. She is *naive* if $\theta \neq \hat{\theta} = 1$ and $c \neq \hat{c} = 0$.

We can characterize consumer behavior as a two-stage decision process and solve it backwards. Given that the consumer buys at the store and remembers rebate redemption, she chooses $r = 1$, if and only if

$$s - c \geq \kappa, \tag{1}$$

where κ is an idiosyncratic taste parameter affecting the redemption decision. In the field experiment, κ can represent various unobserved factors. For example, some subjects may have rare technical issues with their browser that interfere with rebate redemption. As another example, κ may represent the value of an alternative gift card the consumers has received as a birthday present. In the online shop only one promotion code can be used during a purchase, such that the consumer must either use the rebate or the gift card. The value of the gift card then becomes the opportunity cost of redeeming the rebate. These idiosyncrasies do not affect the internal validity of the experiment since treatments are independent of κ by randomization.

[15] In Appendix A.4, I extend the model by a third stage in which the consumer also decides which goods to buy. This extension does not affect the predictions of the model. In Appendix C, I also extend the model by introducing heterogeneity in inattention and hassle costs, and discuss implications for the empirical identification.

Denote $r^* = r(s, c, \kappa)$ as the redemption choice the consumers will actually make conditional on being attentive. Analogously, $\hat{r} = r(s, \hat{c}, \kappa)$ denotes the expected redemption decision conditional on being attentive.

When the consumer chooses whether to buy at the store, she takes into account expected redemption frictions. Denote the outside option utility she gets from not buying at the store by ϵ. Then, she chooses $b = 1$ if and only if the perceived expected utility from buying at the store exceeds the utility from the outside option:

$$\hat{\theta}\left[\hat{r}(s - \hat{c}) + (1 - \hat{r})\kappa\right] + (1 - \hat{\theta})\kappa \geq \epsilon. \tag{2}$$

The left-hand side of equation 2 consists of three parts. First, if she remembers rebate redemption and decides to redeem, she expects to receive $s - \hat{c}$. Second, if she remembers the rebate, but decides not to redeem it, she receives κ. Third, if she forgets about the rebate, she receives κ, as well.[16] Weighting each of these states of the world by its respective perceived probability gives expected utility of buying at the store. The consumer buys whenever this exceeds ϵ.

The model does not make any assumption regarding the formation of consumers' beliefs about their redemption frictions. Consumers may be Bayesian agents that form their beliefs about inattention and hassle costs rationally. They may also have systematically distorted beliefs due to some of the biases in belief formation as documented in the literature.[17]

Expected utility is a function of the true redemption frictions and given by

$$U(b) = b\left[\theta\left(r(s - c) + (1 - r)\kappa\right) + (1 - \theta)\kappa\right] + (1 - b)\epsilon, \tag{3}$$

whereas the actual buying decision, denoted $\hat{b}$, is determined by the perceived redemption frictions:

$$\hat{b} = \arg\max_b\{\hat{U}(b)\} = \arg\max_b\{b\left[\hat{\theta}\left(\hat{r}(s - \hat{c}) + (1 - \hat{r})\kappa\right) + (1 - \hat{\theta})\kappa\right] + (1 - b)\epsilon\}. \tag{4}$$

Let $\Psi = U(\hat{b}) - \hat{U}(\hat{b})$ be the difference between the utility the consumer obtains from the buying decision and the utility she believes she obtains from that decision:

$$\Psi = b(\hat{c}, \hat{\theta})\left[\theta r^*\left(s - c - \kappa\right) - \hat{\theta}\hat{r}(s - \hat{c} - \kappa)\right]. \tag{5}$$

The difference Ψ is commonly referred to as an "internality" or "behavioral wedge" in the

[16] Alternatively, one can also assume that she does not get κ if she forgets to redeem the rebate. The same predictions in Proposition 1 can be derived under this alternative assumption. However, the requirement of the approximation that $1 - \theta$ is relatively small becomes more binding.

[17] See, e.g., Falk and Zimmermann (2018) and Enke and Zimmermann (2019).

behavioral economics literature. It is a price metric that quantifies the welfare loss the consumer bears due to making a mistake. Importantly, this difference is only nonzero in my model if there is a lack of sophistication. The existence of redemption frictions, that is, $\theta \neq 1$ and $c \neq 0$, is not sufficient to conclude that consumers make systematic mistakes. The degree of sophistication is what governs whether redemption frictions create mistakes, not the behavioral friction itself.

Equation 5 also has general implications for the literature on inattention. Most models of inattention do not explicitly allow for sophistication, thereby implicitly assuming full naiveté. This approach ignores that consumers may anticipate their inattention and make appropriate adjustments on an extensive margin, which in my case is whether to buy or not. If consumers are fully aware of their probability of being attentive, inattention does not generate systematic mistakes in choices but is just a simple form of transaction cost.[18] Optimal policy, therefore, crucially depends on the degree of sophistication.

3.2 Aggregating Consumer Choices

To aggregate individual behavior, let the idiosyncratic taste parameters follow an atomless joint distribution $G(\kappa, \epsilon)$. The marginal distributions are denoted by $H(\kappa)$ and $F(\epsilon)$. The true probability of redeeming the rebate (unconditional on buying), denoted $R(s, \theta, c)$, is given by

$$R(s, \theta, c) = \theta \int^{s-c} dH(\kappa). \tag{6}$$

The probability of buying at the store, denoted $B(s, \hat{\theta}, \hat{c})$, can be written as

$$B(s, \hat{\theta}, \hat{c}) = \underbrace{\int^{s-\hat{c}} \int^{\hat{\theta}(s-\hat{c})+(1-\hat{\theta})\kappa} dF(\epsilon|\kappa) dH(\kappa)}_{\text{share of buyers expecting to redeem}} + \underbrace{\int_{s-\hat{c}} \int^{\kappa} dF(\epsilon|\kappa) dH(\kappa)}_{\text{share of buyers expecting not to redeem}}, \tag{7}$$

which consists of the share of buyers who expect to redeem and the share of buyers who do not expect to redeem.

Both the buying and the redemption probability depend on the size of the rebate. However, the redemption probability is determined by the true redemption frictions, whereas the buying probability is a function of the consumers' beliefs about these frictions.

[18] Informational interventions that directly eliminate inattention may still yield welfare gains to rational consumers, by reducing the *transaction costs* of paying attention. See Farhi and Gabaix (2020) and Rodemeier and Löschel (2020) for theoretical frameworks and empirical evidence of the efficiency effects of informational interventions.

3.3 Testable Predictions of Consumer Sophistication

In order to measure whether consumers correctly anticipate redemption frictions, we need to identify perceived and true values for inattention and hassle costs: $\hat{\theta}, \theta, \hat{c}$ and c.[19] Proposition 1 shows that we can empirically identify these parameters through three main treatments that, respectively, create variation in i) the rebate value, ii) the probability of remembering rebate redemption, and iii) the hassle required to redeem. These are "sufficient statistics:" reduced-form treatment effects that approximate structural parameters of interest (Chetty 2009).

To establish this result, I introduce some additional notation and let $\Delta_c R$ denote the effect of eliminating hassle costs on the redemption probability. Analogously, I let $\Delta_{\hat{c}} B$ denote the change in the buying probability in response to an elimination of perceived hassle costs.

Proposition 1. *Perceived inattention and perceived hassle costs can be approximated by reduced-form treatment effects on demand:*

$$\hat{\theta} \approx \frac{\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial s} B(s, 1, \hat{c})} \tag{8}$$

$$\hat{c} \approx \frac{\Delta_{\hat{c}} B(s, 1, \hat{c})}{\frac{\partial}{\partial s} B(s, 1, \hat{c})}. \tag{9}$$

True inattention and true hassle costs can be approximated by reduced-form treatment effects on the redemption rate:

$$\theta = \frac{R(s, \theta, c)}{R(s, 1, c)} \tag{10}$$

$$c \approx \frac{\Delta_c R(s, 1, c)}{\frac{\partial}{\partial s} R(s, 1, c)}. \tag{11}$$

Consumers are sophisticated if and only if

$$\frac{\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial s} B(s, 1, \hat{c})} \approx \frac{R(s, \theta, c)}{R(s, 1, c)} \tag{12}$$

and

$$\frac{\Delta_{\hat{c}} B(s, 1, \hat{c})}{\frac{\partial}{\partial s} B(s, 1, \hat{c})} \approx \frac{\Delta_c R(s, 1, c)}{\frac{\partial}{\partial s} R(s, 1, c)}. \tag{13}$$

[19] One might think that the subjective redemption probability, denoted $R(s, \hat{\theta}, \hat{c})$, could be simply identified by comparing the demand response to a rebate with the demand response to a simple price reduction. Intuitively, if a 2 USD rebate increases demand by the same amount as a 1 USD price reduction, then the subjective redemption probability might be approximately 0.5. However, as I show in Appendix A.2, this strategy, while intuitive, only identifies $R(s, \hat{\theta}, \hat{c})$ under an implicit, and potentially strong, assumption on the distribution of marginal consumers. Proposition 1 does not require this assumption and is more general.

The approximations in equations 8 and 9 each require that $F(\epsilon|\kappa)$ is approximately linear on the interval $[\hat{\theta}(s-\hat{c})+(1-\hat{\theta})\kappa, s-\hat{c}]$ for all κ. This approximation is more accurate when perceived inattention, $1-\hat{\theta}$, is relatively small, or when κ is close to $s-\hat{c}$, or when both conditions apply. Obviously, it is also accurate when the demand function for a given κ, which is $1-F(\epsilon|\kappa)$, is approximately linear over the respective interval, for each κ. Approximate linearity in $F(\epsilon|\kappa)$ assures that the density of consumers that are at the margin is roughly equal across treatment conditions, which is a common assumption in the sufficient statistics literature.[20] Equation 9 and 11 are also not exact because they are first-order approximations of perceived and true hassle costs, respectively.

Equation 8 says that perceived inattention is approximated by the ratio of two treatment effects on the buying probability caused by a small change in the rebate size. The numerator is the treatment effect under perceived inattention, $\hat{\theta}$, whereas the denominator is the effect under perceived full attention. The treatment effect in the numerator can be empirically identified by observing how the buying probability responds to a small change in the value of a rebate. For the treatment effect in the denominator, we need to observe the same margin but to a rebate that consumers know to remember with certainty. In the field experiment, the store offers a reminder at the checkout telling consumers not to forget rebate redemption. Importantly, consumers receive an information treatment at the start of their website visit, telling them that they will receive a redemption reminder at the end.

Equation 8 has a very intuitive interpretation. It tells us that, if a rebate with a reminder increases demand by twice as much as a rebate without a reminder, then consumers' perceived probability of remembering rebate redemption in the absence of the reminder is $\hat{\theta} \approx 0.5$.

Perceived hassle costs are identified in a similar way. The numerator in equation 9 is the treatment effect on the buying probability of fully eliminating the hassle of rebate redemption, given that the store already offers a reminder. In one treatment of the experiment, the store offers a rebate that is automatically redeemed and, therefore, requires no effort. Comparing demand in this condition to demand under the rebate that requires active redemption and includes a reminder identifies $\Delta B(s, 1, \hat{c})$. Since $\hat{c}$ has a money metric, this treatment effect on the buying probability in the numerator needs to be scaled by the demand derivative with respect to the rebate value.

The intuition behind equation 9 is simple, as well. Money-metric hassle costs equal the necessary change in the rebate value that would cause the same increase in the buying probability as a complete elimination of hassle costs would generate. For example, if the buying probability increases by one percentage point to a 10 EUR increase in rebate value, but by two percentage points in response to the elimination of hassle costs, then perceived hassle costs equal approximately 20 EUR.

The true redemption frictions are identified through behavioral responses in the redemption rate

[20] If $1-\hat{\theta}$ is large, or $F(\epsilon|\kappa)$ is highly nonlinear, then structural parameter estimates can be biased by differences in the share of consumers at the margin. More generally, the sufficient statistics approach typically requires that the density of marginal consumers is approximately the same across conditions. For example, the widely used assumption that the demand function for a good is locally linear implies that the CDF of valuations for that good is locally linear.

caused by the treatments. The true probability of remembering rebate redemption is simply the ratio of the redemption probability with and without a reminder. This is straightforward from equation 6.

True hassle costs are identified analogously to perceived hassle costs. They are equal to the minimum increase in rebate value necessary to generate the same increase in the redemption rate that an elimination of hassle costs would generate. This is a simple measure of compensating variation.

The identification strategies of the true redemption frictions parallel Chetty, Looney, and Kroft (2009), who identify inattention to sales taxes by comparing demand elasticities with and without a treatment increasing the salience of taxes. However, different to their work, my approach allows for an identification of both the behavioral bias and the degree of consumer sophistication about this bias. This is possible because I observe two, instead of one, choice margin, and the choices need to be internally consistent.

Consumers are sophisticated when perceived and true values for each redemption friction are equal, which yields the empirically testable predictions in equations 12 and 13. These predictions are straightforward and only rely on reduced-form treatment effects. Proposition 1 therefore provides us with a recipe for an experimental design. The ideal experiment observes how both the buying and the redemption probability respond to a redemption reminder, a treatment that eliminates the hassle of redemption, and a treatment that varies the rebate value.

Finally, note that many treatment effects in Proposition 1 are derivatives. Empirically, however, we can only observe non-marginal changes. Therefore, we require demand and redemption to be locally linear in s, which is a common assumption when estimating demand functions.[21] The identification of true inattention using equation 10 is more general as it neither relies on an approximation, nor does it involve derivatives.

4 Field Experiments

4.1 Cooperation with Online Retailer

I test the theoretical predictions in a natural field experiment in cooperation with a major online retailer for furniture and homeware in the European Union. The retailer operates online stores in the majority of European countries and sells a large variety of products. The main experiment presented in this paper is implemented among consumers in Germany in 2020. Two pilot studies with smaller samples and fewer treatments precede this main experiment.

[21] In fact, the previously mentioned requirement that $F(\epsilon|\kappa)$ is locally linear for each κ implies that demand is locally linear.

4.2 Pilot Studies

The first pilot was implemented in the United Kingdom in 2018 with a sample size of 19,811 consumers. The second pilot was implemented with 52,302 consumers in Germany in 2019. The pilot studies involve fewer treatments and more specific target groups. A description of the pilot studies and the results are presented in Appendix B.

All pilot studies have been pre-registered at the AEA RCT Registry.

4.3 Ethics

All interventions designed for the purpose of this study intend to make consumers better off relative to the standard rebate policy of the firm. Treatments either increase attention toward rebate redemption or intend to eliminate all redemption frictions simultaneously. The firm typically offers rebates that need to be actively redeemed during the checkout in the web shop. They also run regular pricing experiments in which a control group that does not receive a discount typically exists.

It is important to highlight that the cooperating firm does not use any uncommon promotion practices, but rather follows the standard pricing policies in the industry. Claimable rebates are the predominant form of price promotions in online shopping and can be found in virtually every major web shop.

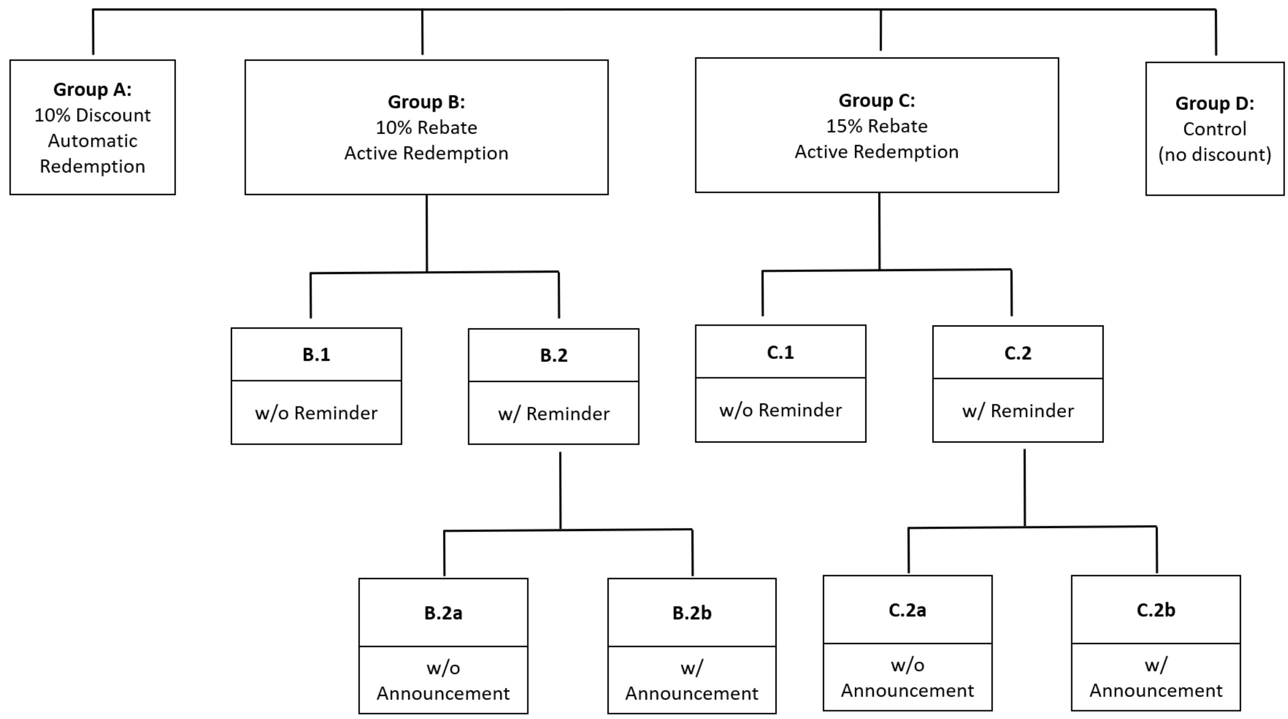
4.4 Design

Figure 1 illustrates the experimental design. Upon visiting the website of the shop, the subject is randomized into one of eight experimental cells with equal probability. Subjects are individually identified and remain in the same experimental condition on follow-up visits.[22] The experimental cells can be categorized into four main groups: A, B, C, and D. Subjects in group A receive an automatically applied discount of 10% on all products. In group B, subjects receive the same price reduction but only if they actively claim the rebate during checkout. Group C receives a larger rebate value of 15% but is otherwise identical to group B. Groups B and C involve several subgroups in which I randomize a reminder and an information treatment about the reminder, as described below. Group D serves as the control group. Subjects in this group are not offered a promotion.

Group A (10% off, automatic redemption): In group A, the promotion banner in Figure 2a is displayed at the top of the browser, saying *"Only for a short period of time: 10% off everything."* An information icon next to this text informs visitors that the discount will be automatically applied at the checkout. Subjects can then browse through the online store and add products to their shopping basket. Once they click on the shopping basket, they see the checkout page presented in Figure 4. On

[22]The firm uses the HTTP cookie to individually identify subjects.

Figure 1: Experimental Design



Note: This figure illustrates the experimental design. Subjects are randomized into one of eight groups with equal probability upon visiting the website.

the checkout page, another banner is displayed telling subjects the discount has been automatically applied. At the center left of the page is a field into which the promotion code has already been entered. An additional small pop-up box above the promotion-code field tells subjects that the code has already been applied. The actual promotion code and some product descriptions have been censored with black bars to protect the anonymity of the company.

Note that even though the discount is automatically applied, it can still be invalidated by actively deleting the promotion code from the field. This may happen on purpose or by accident. For example, subjects that have an alternative promotion, such as a gift card, may delete the discount code in order to enter the code of the gift card. The reason is that only one price promotion can be applied per purchase, such that the consumer needs to decide which code to use. There may also be technical issues related to the device or browser of the subject. The redemption rate will therefore not equal 100%, but will be around 90%. However, because subjects are randomly assigned to the treatments, idiosyncrasies that reduce the redemption rate below 100% are also present in the other experimental conditions, and therefore, do not threaten the internal validity of the experiment.

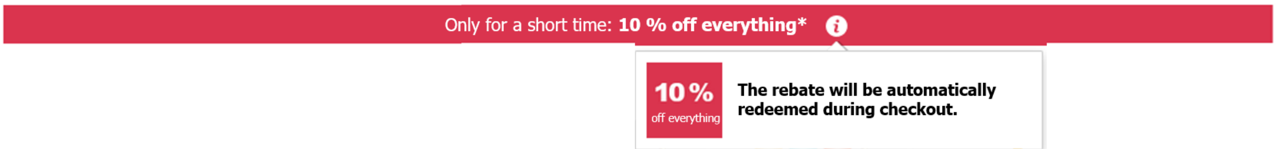
Group B (10% off, active redemption): In this group, subjects are also offered 10% off all products but need to actively claim the rebate. The banner displayed at the top of the browser is shown in Figure 2b and includes the same text as the banner in group A. However, instead of the information icon, a hyperlink says, *"Go to rebate"*. If subjects click on the hyperlink, they are forwarded to another subpage of the website shown in Figure 3. The text on the subpage informs consumers that they need to copy the rebate code shown on the page and enter it into the respective field during the checkout. Subjects only get 10% off their purchase value if they read the text to find the rebate code, copy it onto their clipboard, and remember to paste it into the rebate redemption field on the checkout page

Subjects in group B can be divided into the following three experimental subgroups in which I vary a reminder and an information treatment during checkout.

Group B.1 (without reminder): This treatment represents the standard promotion used in online shopping. Subjects are not reminded to redeem the rebate during the checkout process. An example of the checkout page is shown in Figure 5a and does not involve any reference to the rebate. Subjects see the products they added to the shopping basket, the quantity of each product, and the total purchase price. To redeem the rebate, subjects must paste the rebate code into the respective field that is located below the list of products and on the left side of the page. Once they paste in the code, they need to click on the button saying "apply" right next to the field in order to redeem it. The total purchase price is then reduced by the size of the rebate.

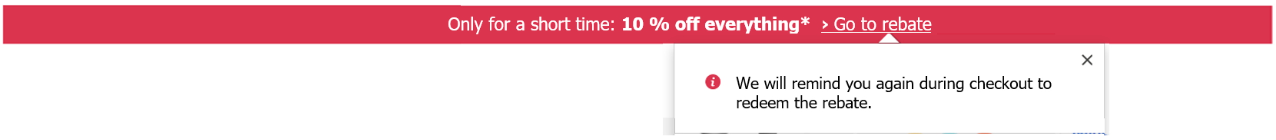
Group B.2a (with reminder): Subjects receive the same treatment as in group B.1 but are

Figure 2: Examples of Promotion Banners



(a) Banner in Group A: 10% Discount, Automatic Redemption

(b) Banner in Groups B.1 and B.2a: 10% Rebate, Active Redemption



(c) Banner in Groups B.2b: 10% Rebate, Active Redemption with Announcement of Reminder

Note: This figure shows an English translation of the banners displayed in the experimental groups A, B.1, B2.a, and B2.b.

Figure 3: Subpage with Rebate Code in Group B.1, B.2a and B.2b

Note: This figure shows an English translation of the subpage showing the rebate code in experimental groups B.1, B2.a, and B2.b.

offered an additional reminder during the checkout process. Figure 5b shows an additional banner, that is displayed on the checkout page and that tells subjects to not forget rebate redemption. In addition, a pop-up box just above the promotion code field highlights the field and tells subjects to enter the rebate code here. These two reminders were designed to be very salient in order to capture subjects' attention as much as possible.

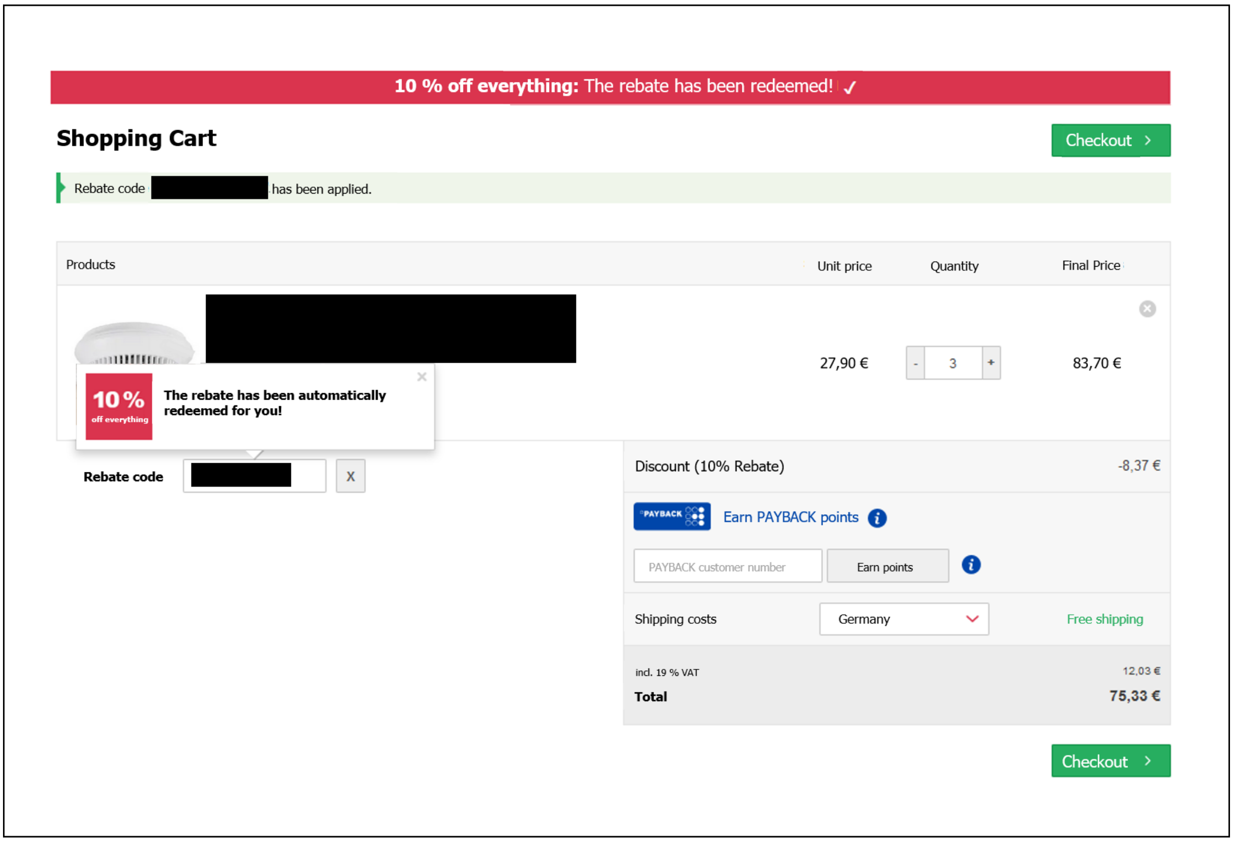
Group B.2b (with reminder and announcement of reminder): This group features an additional information treatment that explicitly announces the reminder at the start of the website visit. The announcement is shown in Figure 2c. The promotion banner is identical to the one in group B.2a but is now accompanied by a pop-up box telling subjects that they will be reminded to redeem during the checkout process.

While both group B.2a and B.2b receive the same reminder during checkout, only group B.2b is informed about this reminder at the very moment they start browsing on the website. Treatment B.2b is directly motivated by the model and intended to identify demand under fully anticipated attention, i.e. $B(s, 1, \hat{c})$. Treatment B.2a, on the other hand, is not produced by the theoretical model. It simply serves as an additional variation to understand the effect of announcing the reminder.

Group C (15% off, active redemption): This group is structured in the same way as group B, but subjects receive a 15% instead of a 10% rebate. Groups C.1, C.2a and C.2b are analogous to groups B.1, B.2a, and B.2b, respectively. The figures showing the treatments can be found in Appendix H.

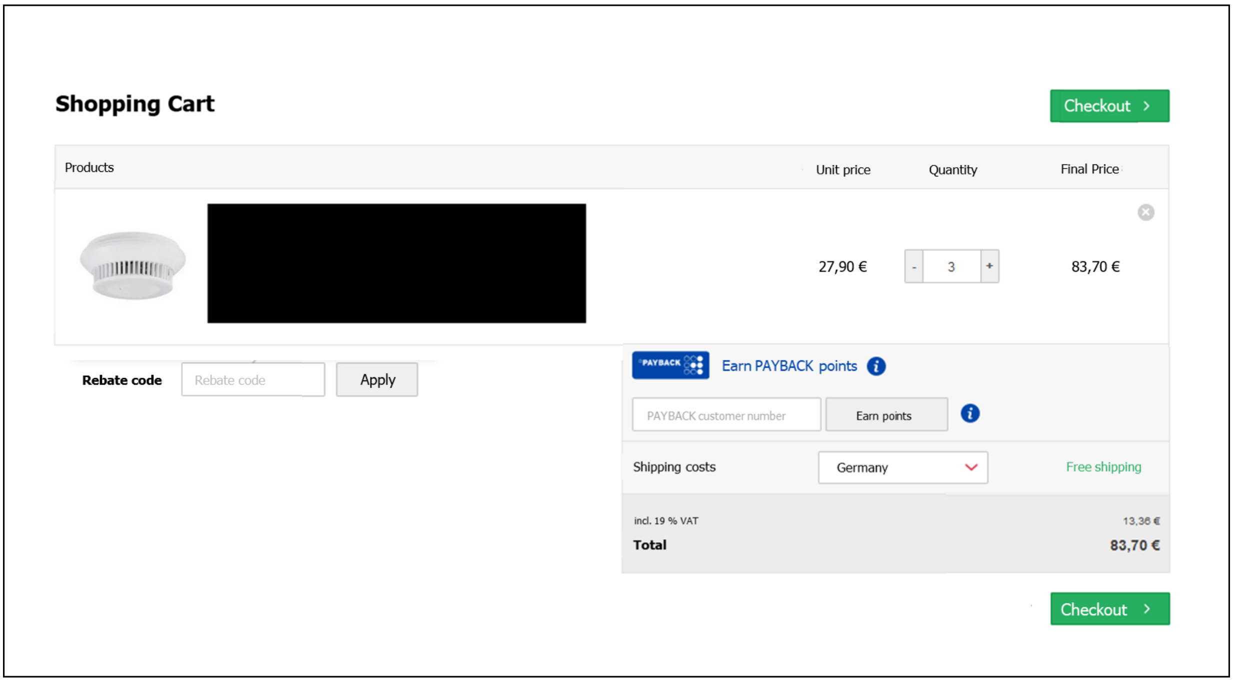
Group D (no promotion): No banner is displayed at the top of the browser and subjects are not offered a rebate. They simply see the status quo of the website without any price promotions.

Figure 4: Checkout Page in Group A: Automatic Redemption

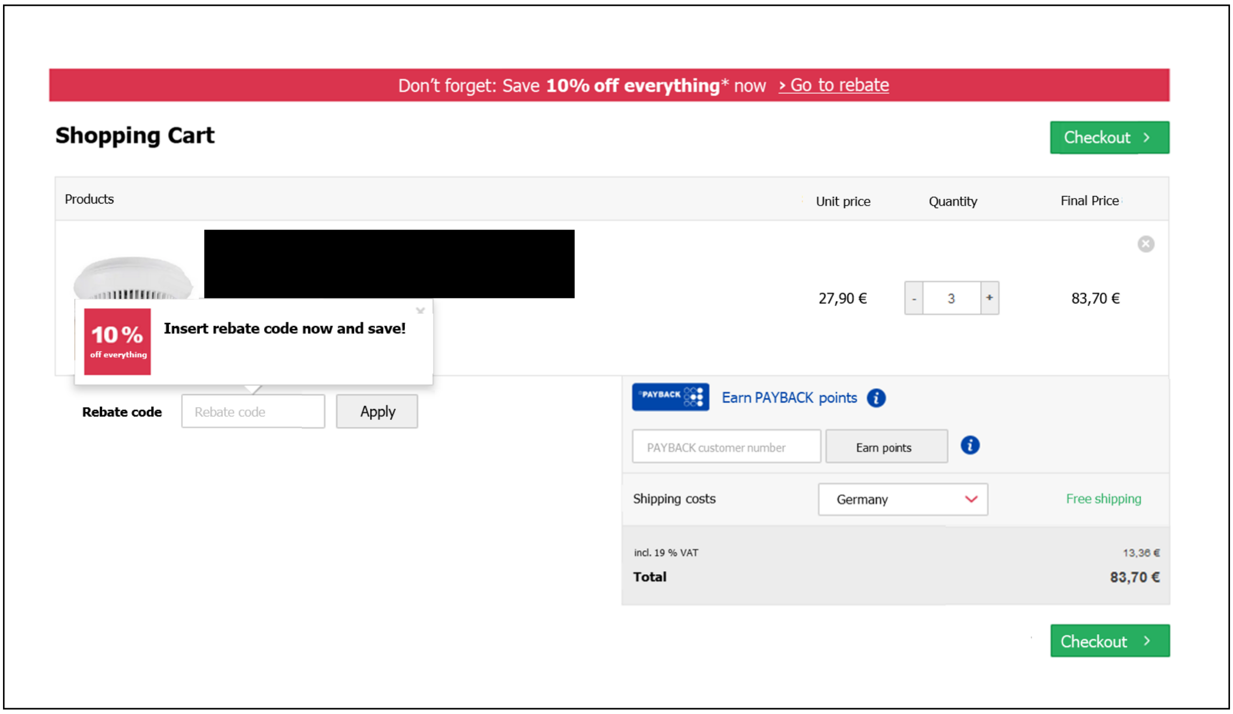


Note: This figure shows an English translation of the checkout page in experimental group A.

Figure 5: Checkout Pages in B-Groups



(a) Checkout Page in Group B.1: No Reminder



(b) Checkout Page in Group B.2a and B.2b: With Reminder

Note: This figure shows an English translation of the checkout pages in experimental groups B.1, B.2a and B.2b.

4.5 Identification of Consumer Sophistication

The experimental design allows for both a qualitative test of consumer sophistication and a structural quantification of the degree of sophistication. The qualitative test has the advantage that it does not rely on the structure imposed by the theoretical model in Section 3 and is more general. The downside is that it only allows us to conclude whether consumers exhibit *some* degree of sophistication, but not how pronounced it is. To quantify welfare implications and guide policy-making, we need to identify the deep primitives underlying consumer choices. In the following, I connect the experimental design to the qualitative and structural tests of sophistication.

4.5.1 A Reduced-Form Test of Sophistication

A simple comparison of differences in the buying probability between rebates with and without redemption frictions is sufficient to conclude whether consumers are fully or (at least) partially naive. Full naiveté implies consumers' buying probability is the same for an automatically applied discount as for a rebate that only gets redeemed with a probability below 1. By contrast, *any* decrease in the buying probability as a response to an exogenous reduction in the redemption rate is evidence of some degree of consumer awareness about redemption frictions.

If consumers anticipate that they might *forget* about rebate redemption, the buying probability must be higher in group B.2b than in group B.1. For the same reason, it must be higher in C.2b than in C.1. Similarly, if they perceive that hassle costs reduce the redemption rate, the buying probability must be higher in group A than in group B.2b. We can even test whether sophistication increases with stakes by comparing whether the buying probability responds more strongly to the reminder treatment with a 15% than to the treatment with a 10% rebate.

Moreover, we are able to quantify which redemption friction consumers perceive themselves as being more affected by. Imagine, for instance, that the buying probability in group B.2b is only slightly smaller than the buying probability in group A, but it is much smaller in group B.1 than in group B.2b. This observation implies consumers perceive their inattention to be a more severe redemption friction than hassle costs.

4.5.2 Identification of Structural Parameters

Before discussing how the experiment identifies the model parameters, first note a nuanced difference between the type of rebate that is offered in the experiment and the rebate modeled in the theory section. The theoretical model involves a lump-sum rebate, whereas in the experiment, the company offers an ad valorem rebate whose absolute monetary value depends on the purchase value. I use a lump-sum rebate in the main part of the paper because it is less involved to model. However, I show in Appendix A.4 that an extended model with an ad valorem rebate yields the exact same proposition as a model with a lump sum rebate. All structural parameters can be identified by aggregate demand responses to an ad valorem rebate, as well, but s from the theory section translates to $s = t \times y$ in the experiment, where $t \in \{10\%, 15\%\}$ is the ad valorem rebate and y is the average purchase value in the treatment group that receives the automatically applied discount (i.e., group A).[23]

Perceived inattention, $\hat{\theta}$, can be identified by comparing the difference in the buying probability between the 10%- and 15%-rebate groups with and without inattention. Specifically, a linear approximation of the numerator in equation 8 is identified by comparing group B.1 with C.1, whereas for the denominator, we need to compare group B.2b with C.2b. To identify the buying probability under fully anticipated attention, that is, under $\hat{\theta} = 1$, we need to use the empirical moments from the treatment groups in which the reminder is explicitly announced at the outset. Otherwise, it would be unclear whether the buying probability fully incorporates the anticipation of the reminder.

We can identify perceived hassle costs by analyzing how the buying probability changes as we move from an automatically applied (hassle-free) discount to a rebate that needs to be actively claimed but that consumers remember during checkout. This change is the difference in the buying probability between group A and group B.2b. Scaling this treatment effect by the effect of a 1 EUR increase in the rebate value on the buying probability yields an approximation of $\hat{c}$, as shown in equation 9.

Identifying true inattention and true hassle costs is more difficult. The empirical challenge stems from the fact that I only observe the redemption probability for individuals who have self-selected into the pool of buyers. This results in a classical sample selection problem in which an outcome of interest (here, rebate redemption) is only observed conditional on another outcome (here, buying). Differences in redemption rates across treatments may then not have a causal interpretation if unobservables that affect the buying decision correlate with unobservables affecting redemption. In Section 5, I address this selection issue empirically by using a sample selection model with an arguably credible exclusion restriction (i.e., an instrument).

For expositional purposes, assume for a moment that the treatments do not cause systematic selection into the subsample of buyers—an assumption I actually provide evidence for later on. In this case, differences in the redemption probability have a causal interpretation and also reflect average differences for the sample of website visitors. True inattention, θ, can be identified by the redemption probability with and without a reminder. Since I observe these moments at two different rebate values, the design overidentifies inattention. Inattention could be identified by taking the ratio of the empirical intensive-margin moments in groups B.1 and B.2b, or by the moments in groups C.1 and C.2b. As I show in the appendix, a third way of identifying inattention is through a comparison of redemption elasticities with respect to the rebate value. Since the system is, therefore, overidentified,

[23] For the estimation, I use the median instead of the mean to adjust for outliers in the revenue distribution. The median purchase value in group A is 96 EUR.

I approximate the efficient weighting matrix by a two-step GMM estimator as described in Section 6.

To identify true hassle costs, c, we first estimate the average treatment effect of reducing hassle costs to zero on the redemption probability, that is, the difference between group A and group B.2b. We then scale this effect by the treatment effect of increasing the rebate value by 1 EUR under full attention, as can be linearly approximated from a comparison between group B.2b and group C.2b.

4.6 Sample

I observe a total of 816,662 website visits by 601,471 individually identified subjects. Table 1 reports summary statistics for each of the eight experimental groups. Each group consists of around 75,000 subjects who visit the website using one of three possible devices.[24] Approximately 35% visit the website using a desktop, 56% use a mobile phone, and 9% use a tablet. These fractions are balanced across all experimental groups and provide confidence of successful randomization.

The four variables at the bottom of the table do not need to balance, because they are potentially endogenous to the treatment variation. The average subject in the control group makes around 1.35 website visits. As the discount size increases, the number of website visits tends to slightly increase. The average buying probability in the control group equals 1.8% and is substantially larger in any of the treatment groups. The redemption probability is zero in the control group by construction and close to 90% in group A, in which the rebate is redeemed automatically. Redemption probabilities are dramatically lower once the rebate needs to be actively redeemed, as they fall to between 53% and 72% depending on the experimental condition.

In the next section, I estimate and discuss average treatment effects on a number of outcomes capturing consumer behavior and on firm profits.

[24] Adding up all observations in Table 1 results in 601,805 instead of 601,471 observations, because there are 334 follow-up purchases; i.e., some consumers buy more than once.

Table 1: Summary Table

Variable	A 10%, automatic	B.1 10%, w/o reminder	B.2a 10%, w/ reminder	B.2b 10%, w/ reminder + announcement	C.1 15%, w/o reminder	C.2a 15%, w/ reminder	C.2b 15%, w/ reminder + announcement	D Control
Desktop user (Yes=1)	0.351	0.351	0.352	0.352	0.350	0.352	0.349	0.353
	(0.477)	(0.477)	(0.477)	(0.478)	(0.477)	(0.478)	(0.477)	(0.478)
Mobile phone user (Yes=1)	0.563	0.563	0.562	0.563	0.564	0.562	0.566	0.562
	(0.496)	(0.496)	(0.496)	(0.496)	(0.496)	(0.496)	(0.496)	(0.496)
Tablet user (Yes=1)	0.087	0.086	0.087	0.085	0.085	0.086	0.085	0.085
	(0.281)	(0.280)	(0.281)	(0.279)	(0.279)	(0.280)	(0.278)	(0.278)
Number of sessions	1.354	1.354	1.353	1.354	1.372	1.361	1.359	1.350
	(1.140)	(1.143)	(1.126)	(1.101)	(1.194)	(1.123)	(1.117)	(1.098)
Made purchase (Yes=1)	0.022	0.021	0.022	0.021	0.023	0.024	0.024	0.018
	(0.147)	(0.144)	(0.148)	(0.145)	(0.149)	(0.154)	(0.153)	(0.133)
Number of purchases	1.041	1.024	1.059	1.032	1.027	1.028	1.026	1.025
	(0.288)	(0.157)	(0.461)	(0.213)	(0.182)	(0.177)	(0.178)	(0.179)
Redeemed rebate (Yes=1)	0.877	0.530	0.630	0.681	0.569	0.691	0.723	0.000
	(0.328)	(0.499)	(0.483)	(0.466)	(0.495)	(0.462)	(0.448)	(0.000)
N	75,256	75,428	74,731	75,368	74,911	75,835	75,403	74,873

Note: This table presents the mean of observable variables in different treatment conditions. Standard deviations are reported in parentheses.

5 Reduced-Form Estimates

5.1 Consumer Behavior

Figure 6 plots the redemption and buying probability for each treatment condition. For expositional purposes, I first pool the two reminder groups with and without announcement of the reminder into one group for each rebate size.[25] The right ordinate depicts the buying probability, while the left ordinate shows the redemption probability conditional on buying.

Table 2 complements Figure 6 by showing estimated average differences in the buying and redemption probability across treatments. All coefficients are multiplied by 100 to ease readability. The empirical specification that produces the coefficients in column 1 for the buying probability is a linear probability model of the following form:

$$Buy_i = \phi_1 + \alpha A + \beta_1' \mathbf{B}_i + \zeta_1' \mathbf{C}_i + \xi_i, \tag{14}$$

where Buy_i is an indicator equal to 1 if subject i made at least one purchase, and 0 otherwise. The buying probability in the control group equals ϕ_1. The indicator A equals one if subject i was offered the automatically-applied discount, and zero otherwise. The column vectors $\mathbf{B}_i$ and $\mathbf{C}_i$ indicate whether subjects received a 10%- or 15%-rebate, respectively, and each include interaction terms with the reminder and the announcement of the reminder. Average treatment effects are given by α, β_1 and ζ_1. ξ_i is the residual.

Estimated differences in the redemption probability are shown in column 2. The empirical specification is the following linear probability model:

$$Redeem_i = \phi_2 + \beta_2' \mathbf{B}_i + \zeta_2' \mathbf{C}_i + \nu_i, \tag{15}$$

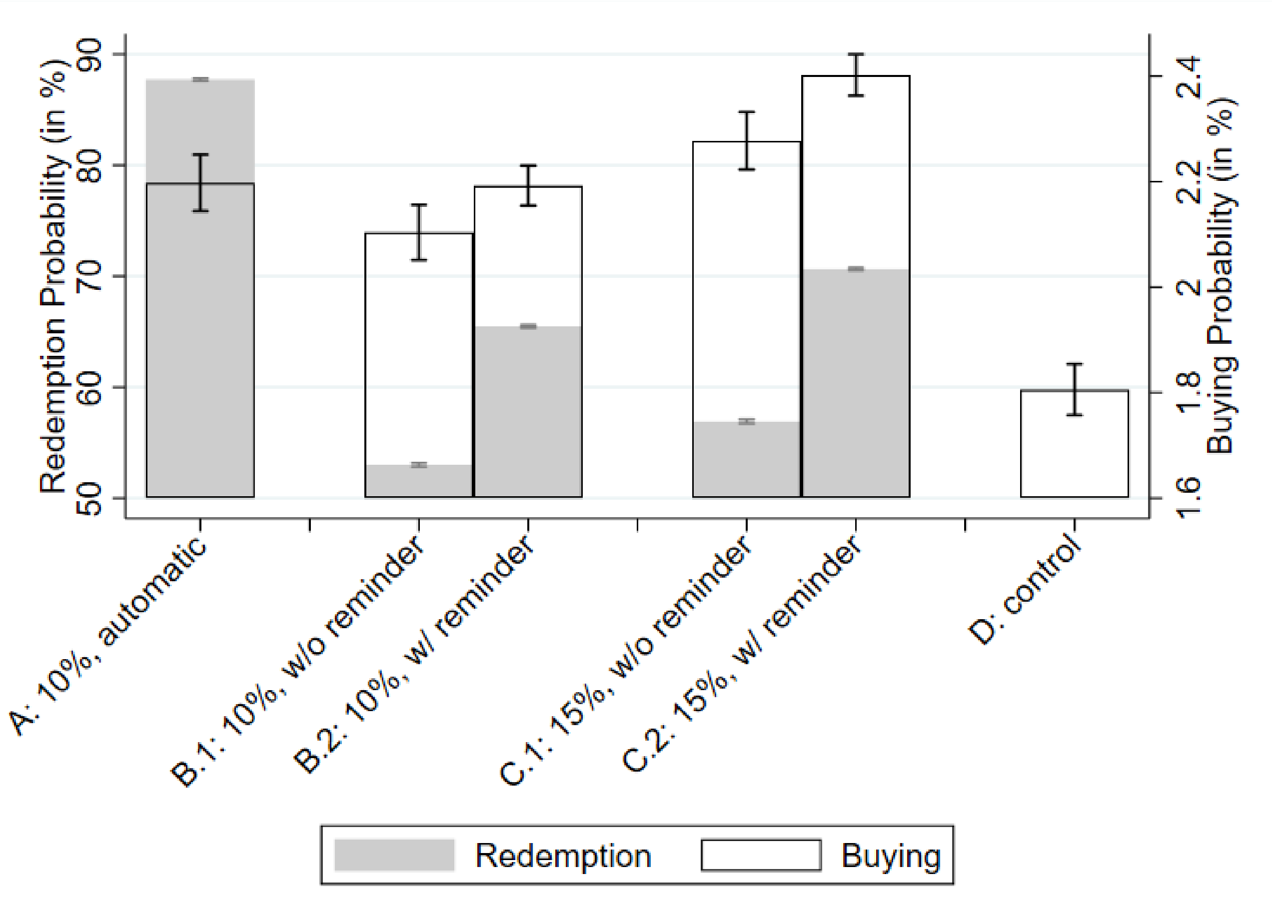
where $Reedem_i$ equals 1 if subject i redeemed the rebate, and 0 otherwise. I exclude control group subjects from the estimation because they could not redeem the rebate by construction of the experimental design. The regression constant, ϕ_2, is the redemption probability in group A, in which the discount is automatically applied. The coefficients in β_2 and ζ_2 measure average differences in the redemption probability across rebate treatments. ν_i denotes the unobserved residual of subject i. Since the rebate is only redeemed if the subject decides to buy, only the subsample of observations who ended up making a purchase is included in the estimation.

The control group has a buying probability of 1.8%. The treatment group that receives an automatically applied 10% discount has a substantially larger buying probability of 2.2%. The difference constitutes an increase of 22% relative to control and is highly statistically significant ($p < 0.01$). The redemption probability is close to 90% and implies that idiosyncratic noise reduces the redemption probability by around 10 percentage points. As previously discussed, the

[25]The disaggregated effects of the reminder treatments can be found in Figure H4 in the Appendix.

reason the redemption probability does not exactly equal 1 may be attributable to the fact that same consumers use alternative gift cards, or to device- and browser-specific technical issues. Because these unobservables should be balanced by random treatment assignment, they only affect the *levels* of but not the *differences* between the redemption probabilities across experimental conditions.

Figure 6: Buying and Redemption Probabilities



Note: This figure shows the buying probability for the entire sample and the redemption probability conditional on buying. The error bars represent standard errors. Reminder treatments with and without announcement are pooled for each discount size.

Table 2: Differences in Buying and Redemption Probabilities

	(1) Buying Probability (in %)	(2) Redemption Probability (in %)
A: 10%, discount	0.392***	87.727***
	(0.075)	(0.844)
B: 10%, rebate	0.298***	-34.734***
	(0.073)	(1.519)
× reminder	0.133*	9.985***
	(0.078)	(1.769)
× announcement	-0.089	5.088***
	(0.079)	(1.706)
C: 15%, rebate	0.472***	-30.810***
	(0.075)	(1.478)
× reminder	0.142*	12.184***
	(0.080)	(1.635)
× announcement	-0.034	3.162**
	(0.081)	(1.533)
D: Control	1.806***	
	(0.050)	
Joint effect of reminder	0.138**	11.131***
	(0.056)	(1.202)
Joint effect of announcement	-0.061	4.078***
	(0.056)	(1.142)
Regression constant	D	A
N	601,805	11,872

Note: The table reports average treatment effects from the OLS regressions specified in equations 14 and 15. The joint effects of the reminder and the announcement are estimated from an alternative regression in which the respective treatments at the 10%- and 15%-rebate are pooled into one variable. Robust standard errors are in parentheses and clustered on the subject level. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

Inattention

Requiring consumers to actively claim the rebate is associated with a sharp and statistically highly significant ($p < 0.01$) drop in the redemption probability of 35 percentage points for the 10%-rebate (group B.1). If the firm offers a reminder at checkout (B.2), the redemption probability rises again, by 10 percentage points ($p < 0.01$). This result is consistent with the notion that limited memory accounts for a large share of redemption frictions: simply reminding consumers of redemption during checkout is associated with a substantial and highly significant increase in redemption rates.

When the rebate value is 15% the reminder is associated with an even larger increase in redemption rates by 12 percentage points ($p < 0.01$).[26]

A comparison of the buying probability with and without a reminder in Figure 6 reveals a remarkable degree of consumer anticipation about their inattention. The movement in the buying probabilities in groups B.1 and B.2 mirrors the movement in the redemption probabilities. Adding a reminder to the rebate causes an additional increase in the buying probability of 0.13 percentage points (B.1 vs. B.2)—a large incremental effect of 45% relative to B.1. The same behavioral pattern is replicated for the treatment groups that receive a 15% discount. The reminder causes an additional increase in the buying probability of 0.14 percentage points (C.1 vs. C.2)—a relative treatment effect of 30%. The treatment effects are direct evidence of consumer sophistication about inattention: consumers are aware they might forget to redeem the rebate and reduce their buying probability relative to a scenario in which they will be attentive due to a reminder.

Regression results in Table 2 tell us that the reminders are individually significant at the 10%-level. Their joint average effect reported at the bottom of the table, equals 0.14 percentage points, and is significant at the 5%-level.

Overall, the results provide us with reduced-form evidence that consumers anticipate their inattention when responding to a price reduction in the form of a rebate.

Hassle Costs

The effect of hassle costs on the redemption rate is identified by the difference in redemption between group A and group B.2. The active redemption requirement causes a large reduction in the redemption rate by approximately 25 percentage points ($p < 0.01$) in comparison to the automatically applied discount—a relative decrease of 28%. This provides reduced-form evidence that, by revealed preference, hassle costs are a substantial redemption friction.

Because we see no difference in the buying probability between A and B.2, there is no indiciation for differential sorting of consumers. The difference in redemption rates should, therefore, be the causal treatment effect of hassle costs on redemption, i.e. $\Delta_c R(s, 1, c)$.

The fact that we do not observe an effect of hassle costs on the buying probability implies that consumers are fully naive about hassle costs. If consumers anticipated hassle costs, we would see a substantial drop in the buying probability as we move from group A to group B.2. However, consulting Figure 6, we see that there is no significant difference in demand between A and B.2. The buying probability is only 0.05 percentage points smaller for B.2 than for A, and this difference is statistically indistinguishable from zero. The important takeaway from these results is that the

[26]Recall again that these difference do not necessarily constitute the causal treatment effect, because the reminder may also affect the type of consumers sorting into the pool of buyers. I address the concern of sample selection in more detail in Section 5.3.

demand response to rebates is too large because consumers do not anticipate the hassle of rebate redemption.[27]

The reduction in the redemption probability caused by hassle costs is more than twice as large as the reduction associated with removing the reminder. Differences in the buying probability instead suggests that consumers perceive inattention as a more important friction to rebate redemption than hassle costs. These results are suggestive evidence of a perverse correlation between perceived and true redemption frictions: consumers are less sophisticated about behavioral biases that cause larger losses to their welfare. This result is also found in the two pilot studies.

Announcement of the Reminder

As documented in column 1 of Table 2, the announcement has no statistically significant effect on demand, neither for the 10%- nor for the 15%-rebate. The joint effect, as reported at the bottom of the table, is not significant either.[28] One reason why there is no incremental effect of the announcement is that even subjects that do not get the announcement can still easily realize that the shop offers a reminder. The reason is that the typical subject browses to the shopping basket frequently during their visit, and thereby sees the reminder. For these subjects the announcement may, therefore, not provide additional information and does not change behavior.[29]

An interesting observation is that the announcement increases redemption rates by 5.1 percentage points ($p < 0.01$) and 3.2 percentage points ($p < 0.05$) for the 10%- and 15%-rebate, respectively. This may suggest that the announcement increased attention to the reminder, which in turn increased redemption rates. In appendix F, I analyze browsing behavior and show that announcing the reminder causes consumers to visit the rebate page more often and view the reminder more frequently.

The takeaway from this is that redemption rates in B.2b and C.2b are more likely to reflect behavior under full attention than the ones in B.2a and C.2a, respectively. I will therefore use the former empirical moments for the structural estimation.

[27] One might argue that consumers may simply not care about a price reduction and, therefore, do not respond to the lower redemption rate caused by hassle costs. However, this interpretation is inconsistent with the observation that consumers buying probability increases as the stakes increase from a 10%- to a 15%-rebate. Thus, consumers clearly respond to price reductions and should, therefore, also reduce their buying probability when the redemption probability exogenously falls.

[28] Note that both coefficients, while not significant, are negative. To understand why the announcement may have a (small) negative effect, note that the effect should be governed by the consumers' prior beliefs regarding the type of rebate the firm is offering. If consumers' prior belief is that the rebate is automatically applied, the announcement implicitly tells them that the rebate needs to be actively redeemed. In this case, we expect a negative effect on demand because some consumers learn that obtaining the promotion requires effort. The negative coefficients may suggests that, without the announcement, some consumers confuse a rebate with an automatically-applied discount, i.e. expect no hassle costs.

[29] As specified in the pre-analysis plan, the relevant treatment to identify true and perceived inattention is the reminder with announcement.

5.2 Heterogeneity

I analyze heterogeneity in treatment effects to separate the role of neoclassical price discrimination from naiveté-based exploitation. For neoclassical price discrimination, we would expect to observe correlated heterogeneity in redemption probabilities and price elaticities, such that the firm could charge higher prices from consumers with lower price elasticities. This would require that less price-elastic consumers face larger redemption frictions. Conversely, rebates cannot price discriminate between preference types if price elasticities are homogeneous because then elasticities are independent of redemption frictions. Understanding heterogeneity in treatment effects is, therefore, crucial in separating the role of neoclassical price discrimination from naiveté-based exploitation.

As documented in the pre-analysis plan, I analyze heterogeneous treatment effects for different income groups. I consider income as a reasonable proxy to separate consumer types, because price elasticities are partially determined by marginal utility of income, and because prior research has documented heterogeneity in behavioral biases across income groups (e.g., Allcott, Lockwood, and Taubinsky 2019).[30]

I use two different datasets on income that vary in their level of aggregation. First, I obtain zip-code level income data, and merge it to each subject based on the approximate location the subject is visiting from. This information is provided by the firm.[31] Since zip-code level income data is not available for every zip code, and some subjects use browser settings that hide their origin, income data can be merged to a sub-sample of 420,857 subjects.[32] Second, I use more aggregated, state-level income data and merge it to each subject based on the state from which they are visiting the website. This reduces the loss of observations considerably, and results in a sample of 582,629 subjects. In the main part of the paper, I discuss results for the larger sample, and provide results from the zip-code level dataset in Appendix C. Both analyses show quantitatively similar results and yield the same conclusion.[33]

I construct a dummy variable equal to one if a subject visits from a region with an income equal or above the sample median, and zero otherwise.[34] I then extend the regressions in equation 14 and 15 by adding the dummy variable and interacting it with the treatment variables. Results are documented in Table 3.

The first important observation is that there is no statistically significant heterogeneity in the demand response to an automatically-applied discount. While low-income consumers are 0.065

[30]In unreported regressions, I also analyze heterogeneity based on voting shares for the main political parties in Germany. I find little robust evidence for significant heterogeneity in price elasticities.

[31]Income data comes from the Institute of Economics and Social Sciences ("Wirtschafts- und Sozialwissenschaftliche Institut").

[32]As pre-registered, I already reduce the loss of observation by merging income data based on the first 4 digits of the 5-digit zip code.

[33]The use of zip-code level data has been pre-registered, while the use of state-level income data has not. I added the state-level income data to increase the sample size.

[34]Following the pre-registration, I calculate the median based on the income distribution of the *sample*, not based on the income distribution of the German population.

percentage points more likely to respond to the price reduction than high-income consumers, this difference is both economically and statistically small. The redemption rates are 3.5 percentage points lower for high-income consumers, suggesting that idiosyncrasies affecting redemption are slightly more relevant for high income consumers.

None of the coefficients that interact the income dummy with the treatment variables is statistically significant. This is true for interaction effects on both demand and redemption behavior. Unreported regressions seperating the sample based on income quartiles rather than the median also show no heterogeneous treatment effects.[35]

Based on these results, there is no clear rationale to use rebates to price-discriminate between income groups, as price elasticities are roughly homogeneous. Instead, rebates are more likely to be a profitable promotion because consumers are naive about hassle costs.

It is important to highlight that there may be other idiosyncratic correlations between preferences and redemption probabilities that this analysis does not capture. To fully understand the role of heterogeneity for rebate promotions I would require knowledge of the joint distribution of price elasticities and redemption probabilities. Since this distribution obviously remains unobserved, I have to resort to pre-registered observable characteristics of consumers.

[35] I pre-registered to analyze heterogeneity based on income quartiles, as well.

Table 3: Heterogeneity by Income

	(1) Buying Probability (in %)	(2) Redemption Probability (in %)
A: 10%, discount	0.429***	89.568***
	(0.106)	(1.063)
B: 10%, rebate	0.283***	-35.956***
	(0.103)	(2.076)
× reminder	0.055	12.223***
	(0.093)	(2.153)
C: 15%, rebate	0.497***	-33.209***
	(0.107)	(2.025)
× reminder	0.019	14.842***
	(0.098)	(2.043)
D: Control	1.802***	
	(0.071)	
Above median income	0.039	-3.489**
	(0.102)	(1.705)
Interaction effects:		
A: 10%, discount × above median income	-0.065	
	(0.153)	
B: 10%, rebate × above median income	0.042	2.564
	(0.149)	(3.069)
× reminder	0.076	0.601
	(0.135)	(3.064)
C: 15%, rebate × above median income	-0.001	4.463
	(0.153)	(2.979)
× reminder	0.173	-2.242
	(0.141)	(2.897)
N	582,629	11,613

Note: The table reports treatment effects for consumers from regions with an income below and above the sample median. Robust standard errors are in parentheses and clustered on the subject level. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

5.3 Effect on Firm Profits

Are rebates more profitable for the firm than automatically-applied discounts? This is not a trivial question if consumers are sophisticated, because then low redemption rates also imply lower demand responses to rebates than to discounts. The firm provides me with rich data on markups for each product offered in the online store. Based on the data, I calculate profits for each website visitor in the experiment. In addition, I calculate the total value of the shopping basket before any promotion has been applied, for each subject. This variable, therefore, equals revenue plus the value of any promotion that has been applied.

Table 4 reports the percentage change in shopping basket value and profits for each treatment relative to control group profits. Column 1 shows that there is no difference in the shopping basket value between the automatically applied discount (A) and the equivalent rebate (B). Both promotions increase the value of the shopping basket by around 20% ($p < 0.01$). Thus, the average subject chooses the same value of goods with a rebate than with a discount even though the discount is only redeemed half of the time. This is is why in column 3, the increase in profits is more than twice as large for the rebate (+9.3%) than for the discount (+3.7%).[36] While the coefficient for the discount is not significant, the coefficient for the rebate is marginally significant ($p < 0.07$). Overall, effects on the shopping basket value can be measured more precisely since effects are (obviously) larger, while standard errors are roughly the same. The difference in magnitudes is large: simply changing the promotion details by requiring consumers to redeem the rebate may increase the profitability of the promotion by 150%.

This effect seems to be entirely driven by the fact that consumers underestimate hassle costs, not inattention. The reminder (column 2 and 4) has no significant—and in particular no negative—effect on profits. Put differently, exploiting inattention (by not reminding consumers to redeem) offers no significant benefit for the firm.[37] The reason that not reminding consumers does not increase profits is that, even though this reduces redemption, it also reduces demand because consumers anticipate their inattention. Thus, it is consumers' naiveté about hassle costs that makes rebates overall more profitable than discounts.

In conclusion, simple and costless changes in the promotion feature have dramatic effects on firm returns because of consumers' partial naiveté. These results rationalize firm practices referred to as buy baits or "sludges" (Thaler 2018) more generally, and provide empirical evidence for the theoretical literature on naiveté-based exploitation (e.g., Heidhues and Kőszegi 2017).

[36] One might be tempted to interpret the positive coefficient of A on profits as suggestive evidence that the firm has not set prices optimally. However, note that we only observe effects on profits within the experimental time frame. The effects of a permanent price reduction on long-term profits are not observed.

[37] Since the coefficient is even positive, this may suggest that offering a reminder increases profits.

Table 4: Shopping Basket Value and Profits

	Shopping Basket Value (in % to control)		Profits (in % to control)	
	(1)	(2)	(3)	(4)
A: 10%, discount	19.814***	19.814***	3.707	3.707
	(6.822)	(6.822)	(6.121)	(6.121)
B: 10%, rebate	20.364***	18.040***	9.272*	8.175
	(5.238)	(6.795)	(5.055)	(6.459)
× reminder		3.491		1.648
		(6.260)		(5.722)
C: 15%, rebate	32.403***	28.278***	11.853**	11.225*
	(5.295)	(6.584)	(4.999)	(6.098)
× reminder		6.170		0.939
		(6.198)		(5.300)
N	601,805	601,805	601,805	601,805

Note: The table reports average treatment effects on the value of the shopping basket and on profits. Both outcome variables are in percent relative to control group values. The shopping basket value is gross of the potential price reduction. Robust standard errors are in parentheses and clustered on the subject level. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

5.3.1 Effects on Consumer Loyalty

An important caveat of the previous analysis should be made explicit. The data only allows us to draw conclusions for short-term effects of the treatments on profits. This limitation is a potential concern because exploiting consumer biases may have negative long-term effects on the probability of returning to the shop. Even consumers who are naive in the short term might realize they are being exploited in the long term and, as a result, decide not to return to the store. In Appendix E, I investigate dynamic effects of redemption frictions on the probability of returning to the store. I find no significant negative effects on customer loyalty. However, the time period in which I observe the sample is relatively short and, as a result, follow-up purchases tend to be rare in the data. But note that even if long-term effects on customer loyalty could be negative, they do not seem sufficiently large to drive firms out of the market. As previously mentioned, both the cooperating firm and the majority of large online retailers frequently offer these type of rebates, indicating the use of buy baits persists in a long-run equilibrium.

5.4 Addressing Sample Selection

Before moving to the estimation of the structural parameters, I address a potential concern related to systematic sorting of subjects into the subsample of buyers. Because the treatments may affect the type of subjects that select into the pool of buyers, differences in the redemption probability may not have a causal interpretation. Selection is not a concern for the identification of the treatment effect of hassle costs on redemption rates, because there is no evidence that hassle costs affect the buying probability. Thus, the experiment identifies the causal effect of hassle costs on redemption without additional assumptions. The conclusion that hassle costs are large and fully unanticipated remains valid. However, the effect of the reminder on redemption rates may not be identified because the buying probability responds to the reminder. This could indicate systematic sorting of consumers.

A large literature in econometrics has developed techniques to address bias resulting from sample selection building on the seminal work in Heckman (1976) and Heckman (1979). A consensus in the literature is that convincing identification in these models requires a credible exclusion restriction: a variable that does not directly affect the outcome of interest but affects whether subjects select into the sample.

Internet Outages as Exclusion Restriction. I address the potential bias resulting from sample selection by estimating a fully parametric selection model with an arguably credible exclusion restriction: regional and temporal variation in sudden internet outages. The exclusion restriction requires that internet outages affect the redemption probability only indirectly through its effect on the buying probability. I view this assumption as a plausible one: an internet outage reduces the probability to redeem because it simply cuts people off from access to the online shop—so they cannot buy—but not because of other channels.

Using this exclusion restriction, I estimate a selection model with normally distributed residuals and a binary dependent variable for both the selection and outcome equation, as first formulated by Van de Ven and Van Praag (1981). Monte Carlo simulations show that when these distributional assumptions are violated, the model still performs well in many cases as long as a valid exclusion restriction exists (Cook and Siddiqui 2020). A challenge with sample selection models are high levels of collinearity between treatment regressors and the correction term. To reduce the level of collinearity I allow for heterogeneous effects of outages on demand by interacting the instrument with income. I provide a detailed discussion about this model in Appendix G, where I elaborate on the dataset on internet outages, the construction of the exclusion restriction, and the estimation of the selection model.

Before estimating the selection model, I first analyze whether internet outages have a significant effect on the buying probability. Table 5 provides results from a linear probability model of the buying decision on internet outages and the treatments. Column 1 only includes the instrument, whereas column 2 adds the experimental treatments. Major internet outages cause an economically large

Table 5: Effect of Internet Outages on Buying Probability

	Buying Probability (in %)	
	(1)	(2)
Internet Outage	-0.180***	-0.180***
	(0.052)	(0.052)
A: 10%, discount		0.393***
		(0.075)
B: 10%, rebate		0.298***
		(0.073)
× reminder		0.134*
		(0.078)
× announcement		-0.089
		(0.079)
C: 15%, rebate		0.472***
		(0.075)
× reminder		0.142*
		(0.080)
× announcement		-0.034
		(0.081)
Constant	2.227***	1.836***
	(0.021)	(0.051)
N	601,805	601,805

Note: The table reports average treatment effects from a linear probability model of internet outages and treatment indicators on the buying probability. Column 1 only includes internet outages as a regressor, and column 2 adds the experimental treatments. Robust standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

and highly statistically significant decrease in the buying probability by 8.3%, or 0.18 percentage points. The addition of experimental treatments in column 2 does not affect the coefficient of the instrument. This finding is reassuring because it indicates the exclusion restriction is not correlated with the experimental treatments—a result we expect due to random treatment assignment.

Treatment Effects on Redemption Corrected for Selection. Table 6 reports the main estimation results. The log-likelihood value is –61,309 and corresponds to the global maximum as I show in the appendix. The correlation between residuals is 0.42, which would imply that unobservables that increase the buying probability also increase the redemption probability. Two independent linear probability models would then overestimate the redemption probability, because subjects with a systematically larger likelihood of redeeming have selected into the subsample of buyers.

However, there is no indication for significant sample selection bias: the coefficient showing Fisher's Z-transformation of the correlation between residuals is not statistically significantly different from zero. This implies that, given joint normality of residuals, the simple OLS regression in equation 15 identifies the causal treatment effects on redemption.

The treatment coefficients are also similar to the ones of the two OLS models with independent residuals. The effect of hassle costs, as identified by B.2b, equals a reduction in the redemption probability by 25 percentage points, and is therefore the same to the OLS model. The (announced) reminder increases the redemption rate by 4 percentage points, i.e. less than in the model with independent residuals. Overall, the treatment coefficients are fairly similar to the treatment effects estimated in the main part of the paper, indicating the degree of selection bias is small if the model is correctly specified.[38]

Table 6: Estimation Results from Sample Selection Model

	Redemption Probability (in %)
B.1: 10%, w/o reminder	-36.696*
	(20.252)
B.2a: 10%, w/ reminder	-28.982**
	(12.967)
B.2b: 10%, w/ reminder+announce	-25.277***
	(9.748)
C.1: 15%, w/o reminder	-34.017*
	(17.881)
C.2a: 15%, w/ reminder	-23.661**
	(9.348)
C.2b: 15%, w/ reminder+announce	-21.068***
	(7.630)
ρ	0.416
	(0.441)
Fisher's Z-transformation	0.443
	(0.533)
Log likelihood	-61308.62
N	510,207

Note: This table reports estimation results from the sample selection model in equation 48. Control group subjects are excluded from the estimation because they cannot redeem by construction. The correlation between intensive and extensive margin residuals is denoted by ρ. Fisher's Z transformation is the inverse hyperbolic tangent of ρ and asymptotically normally distributed. Standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

[38] Standard errors are substantially larger in the selection model due to collinearity with the correction term—a well-known issue in sample selection models. See Appendix G for a more detailed discussion.

6 Structural Parameter Estimates

The reduced-form results tell us that consumers exhibit some degree of sophistication regarding their inattention, but are fully naive about their hassle costs. To quantify the *degree* of perceived and true inattention and hassle costs, respectively, we require structural estimates of the underlying choice parameters. For this purpose, I estimate the sufficient statistics derived in Proposition 1 using a two-step GMM estimator. Appendix A.5 provides the derivation of the moment conditions.

Table 7 reports the estimation results. Differences in the buying probability across treatments imply consumers' subjective probability of remembering rebate redemption is 73%. Their true probability of remembering is 78% and is, therefore, remarkably close to consumers' expectations. Recall that Proposition 1 only requires these point estimates to be *approximately* equal. We can therefore conclude that consumers may well be sophisticated about their inattention, because their buying and redemption choices are approximately consistent.

By contrast, consumers vastly underestimate the hassle costs of redeeming the rebate. Since the buying probability hardly responds to an introduction of hassle costs, the perceived hassle only equals 1 EUR. However, the strong drop in the redemption probability implies hassle costs are a significant redemption friction equal to approximately 21.16 EUR. Although this number may seem large, note that the hassle of redeeming the rebate may come in various forms. For instance, the hassle of finding out where to find the rebate code and how to exactly redeem it may be substantial for certain consumers. Especially people with a low level of experience with online shopping may struggle to understand how to redeem the rebate and give up after some time spent engaging with the feature. Hassle costs therefore represent not only the process of copying and pasting the rebate code into the respective field, but also the time and effort required to understand how redemption is done and where to find the necessary rebate code. This process may be more challenging for subjects with lower levels of digital literacy, such as more senior citizens. In addition, people with larger opportunity costs of time might be more willing to forgo discounts when required to complete the technical redemption process themselves.

There may, however, be a potential upward bias because the estimated parameter is only a *linear* approximation of hassle costs. If demand is highly convex in rebate value as we move from a 10%- to a 15%-rebate, then the presented estimate is too large and represents an upper bound of the true hassle costs. While true hassle costs are smaller than estimated in this case, they are still far above perceived hassle costs, which are statistically indistinguishable from zero.

Overall, the large difference between perceived and actual hassle costs yields the conclusion that the average consumer does not fully anticipate the struggle of redemption when deciding to make a purchase, and underestimates the hassle of rebate redemption by up to 20.16 EUR.

Table 7: Structural Estimates

	Extensive Margin	Intensive Margin
Inattention and Hassle Costs:		
$\hat{\theta}$	0.730 *	
	(0.390)	
$\hat{c}$ (in EUR)	1.002	
	(1.380)	
θ		0.784 ***
		(0.014)
c (in EUR)		21.156 ***
		(5.289)
Other Parameters:		
β_A	0.004 ***	
	(0.001)	
β_{B1}	0.003 ***	
	(0.001)	
β_{B2}	0.003 ***	
	(0.001)	
β_D	0.018 ***	
	(0.001)	
τ_A		0.877 ***
		(0.011)
$\frac{dR(s,1,c)}{ds}$		0.009 ***
		(0.003)
N	451,239	8,365

Note: The table reports estimation results from the GMM estimations specified in equations 35 and 36. Standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

How large is counterfactual demand if consumers were fully sophisticated?[39] Since $c = 21$€ and $\hat{c} = 1$€, consumers respond to the 10 EUR rebate the same way sophisticated consumers would respond to a rebate with a value of 31 EUR. The sophisticated demand response to a rebate with reminder is, therefore, only around 1/3 of the observed treatment effect: $\frac{\beta_{B.2b}\times 10}{31} = 0.11$. The sophisticated demand response to a 10 EUR rebate *without* a reminder is lower because inattention becomes an additional redemption friction. The optimal demand response of sophisticated consumers to such a rebate can be found by multiplying the sophisticated demand response to a rebate with reminder by the true probability of being attentive: $\frac{\beta_{B.2b}\times 10}{31} \times 0.78 = 0.086$. Thus, the observed treatment effect of the standard rebate in B.1 on demand, $\beta_{B.1} = 0.298$, is $\frac{0.298}{0.089} - 1 = 235\%$ too large. Conversely, the optimal demand response of fully sophisticated consumers is only around 30% of the observed treatment effect.

In sum, both reduced-form and structural results point to a substantial misperception of the effort related to rebate redemption. This misperception lures consumers into making a purchase even though the hassle of obtaining the discount turns out to be too large to be worthwhile. The presented results, therefore, substantiate the motive for consumer protection laws that limit the use of claimable rebates.

7 Additional Mechanisms

I discuss a number of alternative mechanisms that may affect the interpretation of the data, as well as the identification of the structural parameters.

Risk preferences and loss aversion. Demand may be more elastic to price reductions than to rebates with the same expected value because consumers are averse to risk and losses. First, consumers' utility of income may be concave, which implies classical risk-averse preferences. Second, consumers may be loss-averse, meaning the disutility of losing a monetary amount significantly exceeds the utility of a monetary gain of equal size.

Loss aversion is arguably a more relevant factor in my empirical setting than risk aversion, because the rebate value is relatively small. For reasonable degrees of risk aversion, even risk-averse consumers should behave risk neutral over small gambles (Rabin 2000).

Both risk and loss aversion would have the same directional effect on the empirical estimates. In particular, introducing risk and loss aversion provides two additional reasons for why consumers should respond *even less* to the rebate than to the automatically-applied discount. The previously presented estimates imply that the demand response to a rebate is already excessively high under the

[39] It might seem intuitive to simply substitute $\hat{c} = c = 21.16$ €into $\Delta_{\hat{c}} B(s, 1, \hat{c}) \approx \hat{c}\frac{\partial B(s,1,\hat{c})}{\partial s}$ to find the optimal demand response under sophistication, i.e. to find $\Delta_c B(s, 1, c)$. However, this approach does not yield the correct result because the demand *derivative* itself depends on c (see proof of Proposition 1 in the Appendix to verify).

assumption of risk neutrality and no loss aversion. If consumers are risk- and loss-averse, then the demand response of sophisticated consumers should be even lower than previously discussed. Thus, extending the model to capture an aversion to risk and losses would further strengthen the qualitative conclusion that the demand response to rebates is too large because consumers overestimate their redemption probability. As a consequence, the implication for consumer protection regulation would remain unchanged.

An important limitation is that the structural estimates would change. The presented estimates in the previous section would overstate the degree of consumer sophistication, because the model calculates a counterfactual demand response under full sophistication that is too large for risk- and loss-averse consumers.

Social preferences. Existing evidence in literature on social preferences indicates that subjects exhibit altruistic and reciprocal preferences (e.g., Fehr and Gächter 1998). Even sophisticated consumers who would benefit from rebate redemption might decide not to buy at the store, because they consider rebates an unfair marketing practice. Consumers may also receive direct disutility from a firm's attempt to exploit their own or other consumers' behavioral tendencies.

Introducing a distaste for exploitation to the model would have the same directional effect as the introduction of risk and loss aversion: it provides another reason for why consumers should respond less to a rebate than to an automatically-applied discount. This model extension would, therefore, not affect the qualitative conclusion that demand overreacts to rebates. However, it would increase the magnitude of that overreaction.

8 Conclusion

This paper studies consumers' sophistication about their own behavioral tendencies in the context of large-scale rebate promotions. I develop novel theoretical predictions of sophisticated behavior that can be empirically tested by observing a small set of aggregate demand elasticities. I then take this model to a natural field experiment and estimate consumer sophistication and its economic implications, using choices from hundreds of thousands of consumers.

I find that consumers exhibit a remarkable degree of sophistication regarding their inattention but are almost fully naive with respect to hassle costs. As a result, claimable rebates cause an excessive increase in the buying probability because consumers do not fully anticipate the hassle of redemption. Exploiting consumer naiveté is an impressive lever of profit and increases rebate returns by 150%.

Results have important implications for the policy debates around the regulatory framework that limits the use of rebate promotions in many countries across the world. I provide the first evidence from a natural field experiment that tests whether rebates harm consumers by causing them to make systematically distorted buying decisions. The evidence indicates that the regulators' qualitative intuition is correct and that consumer protection laws may have large positive effects on consumer welfare.

A limitation of the study is that it cannot draw conclusions regarding supply-side responses to regulatory interventions. For instance, it is unclear whether a ban on claimable rebates would result in a long-run equilibrium in which firms do not offer price promotions at all instead of offering frictionless discounts. Future research can make important contributions by studying firm responses to various policy counterfactuals in order to obtain a complete picture of the economic implications of consumer protection regulation.

References

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman. 2020. “Are high-interest loans predatory? Theory and evidence from payday lending.” Working Paper.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky. 2019. “Regressive sin taxes, with an application to the optimal soda tax.” *The Quarterly Journal of Economics* 134 (3): 1557–1626.

Augenblick, Ned, and Matthew Rabin. 2019. “An experiment on time preference and misprediction in unpleasant tasks.” *Review of Economic Studies* 86 (3): 941–975.

Ausubel, Lawrence M. 1991. “The failure of competition in the credit card market.” *The American Economic Review*, pp. 50–81.

Bai, Liang, Benjamin Handel, Edward Miguel, and Gautam Rao. 2017. “Self-control and demand for preventive health: Evidence from hypertension in India.” NBER Working Paper No. 23727.

Bernheim, B. Douglas, and Dmitry Taubinsky. 2019. “Behavioral Public Economics.” In *Handbook of Behavioral Economics*, edited by B. Douglas Bernheim, Stefano Della Vigna, and David I. Laibson, Volume 1. Elsevier.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. “Salience and Consumer Choice.” *Journal of Political Economy* 121 (5): 803–843.

———. 2020. “Memory, Attention, and Choice.” *The Quarterly Journal of Economics* 135 (3): 1399–1442.

Bronchetti, Erin T, Judd B Kessler, Ellen B Magenheim, Dmitry Taubinsky, and Eric Zwick. 2020. “Is Attention Produced Rationally?” NBER Working Paper No. 27443.

Brown, Jennifer, Tanjim Hossain, and John Morgan. 2010. “Shrouded attributes and information suppression: Evidence from the field.” *The Quarterly Journal of Economics* 125 (2): 859–876.

Caplin, Andrew, Mark Dean, and John Leahy. 2017. “Rationally inattentive behavior: Characterizing and generalizing Shannon entropy.” NBER Working Paper No. 23652.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky. 2019. “How are preferences for commitment revealed?” NBER Working Paper No. 26161.

Chaloupka IV, Frank J, Matthew R Levy, and Justin S White. 2019. “Estimating Biases in Smoking Cessation: Evidence from a Field Experiment.” NBER Working Paper No. 26522.

Chetty, Raj. 2009. “Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods.” *Annual Review of Economics* 1 (1): 451–488.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. “Salience and Taxation: Theory and Evidence.” *The American Economic Review* 99 (4): 1145–1177.

Connecticut Unfair Trade Practices Act Regulations. 2004. Connecticut Unfair Trade Practices Act Regulations, Sec. 42-110b-19.

Cook, Jonathan A, Noah Newberger, and Joon-Suk Lee. 2020. "On identification and estimation of Heckman models." Working paper.

Cook, Jonathan A, and Saad Siddiqui. 2020. "Random forests and selected samples." *Bulletin of Economic Research* 72 (3): 272–287.

Cummings, Ronald G, Glenn W Harrison, and E Elisabet Rutström. 1995. "Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive-compatible?" *The American Economic Review* 85 (1): 260–266.

DellaVigna, Stefano. 2018. "Structural behavioral economics." In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 1, 613–723. Elsevier.

DellaVigna, Stefano, and Ulrike Malmendier. 2004. "Contract design and self-control: Theory and evidence." *The Quarterly Journal of Economics* 119 (2): 353–402.

———. 2006. "Paying Not to Go to the Gym." *American Economic Review* 96 (3): 694–719 (June).

Dertwinkel-Kalt, Markus, Mats Köster, and Matthias Sutter. 2020. "To buy or not to buy? Shrouding and partitioning of prices in an online shopping field experiment." *European Economic Review*, vol. 130. Artciel 103593.

Enke, Benjamin, and Florian Zimmermann. 2019. "Correlation Neglect in Belief Formation." *The Review of Economic Studies* 86 (1): 313–332.

Ericson, Keith M Marzilli. 2011. "Forgetting we forget: Overconfidence and memory." *Journal of the European Economic Association* 9 (1): 43–60.

Falk, Armin, and Florian Zimmermann. 2018. "Information Processing and Commitment." *The Economic Journal* 128 (613): 1983–2002.

Farhi, Emmanuel, and Xavier Gabaix. 2020. "Optimal Taxation with Behavioral Agents." *American Economic Review* 110 (1): 298–336.

Federal Trade Commission. 2020. Consumer Information: Rebates. URL: https://www.consumer.ftc.gov/articles/0096-rebates (Retrieved: 09/20/2020).

Fehr, Ernst, and Simon Gächter. 1998. "Reciprocity and economics: The economic implications of homo reciprocans." *European Economic Review* 42 (3-5): 845–859.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded attributes, consumer myopia, and information suppression in competitive markets." *The Quarterly Journal of Economics* 121 (2): 505–540.

Gilpatric, Scott M. 2009. "Slippage in rebate programs and present-biased preferences." *Marketing Science* 28 (2): 229–238.

Gottlieb, Daniel, and Kent Smetters. 2012. "Narrow framing and life insurance." NBER Working Paper No. 18601.

Grubb, Michael D. 2009. "Selling to overconfident consumers." *American Economic Review* 99 (5): 1770–1807.

Grubb, Michael D, and Matthew Osborne. 2015. "Cellular service demand: Biased beliefs, learning, and bill shock." *American Economic Review* 105 (1): 234–71.

Harrison, Glenn W., and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42 (4): 1009–1055.

Heckman, James J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." In *Annals of economic and social measurement, volume 5, number 4*, 475–492. NBER.

———. 1979. "Sample selection bias as a specification error." *Econometrica* 47 (1): 153–161.

Heidhues, Paul, and Botond Kőszegi. 2010. "Exploiting naivete about self-control in the credit market." *American Economic Review* 100 (5): 2279–2303.

———. 2017. "Naivete-based discrimination." *The Quarterly Journal of Economics* 132 (2): 1019–1054.

Hossain, Tanjim, and John Morgan. 2006. "... plus shipping and handling: Revenue (non) equivalence in field experiments on ebay." *The BE Journal of Economic Analysis & Policy* 6, no. 2.

Jolson, Marvin A, Joshua L Wiener, and Richard B Rosecky. 1987. "Correlates of rebate proneness." *Journal of Advertising Research* 27 (1): 33–43.

Juniper Research. 2021. "Mobile Online Coupons: Leading Vendors, Technologies Market Forecasts 2017-2022." Technical Report.

Kőszegi, Botond. 2014. "Behavioral Contract Theory." *Journal of Economic Literature* 52 (4): 1075–1118.

Kőszegi, Botond, and Adam Szeidl. 2013. "A model of focusing in economic choice." *The Quarterly journal of economics* 128 (1): 53–104.

List, John A. 2003. "Does market experience eliminate market anomalies?" *The Quarterly Journal of Economics* 118 (1): 41–71.

———. 2011. "Does market experience eliminate market anomalies? The case of exogenous market experience." *American Economic Review* 101 (3): 313–17.

List, John A, and Craig A Gallet. 2001. "What experimental protocol influence disparities between actual and hypothetical stated values?" *Environmental and Resource Economics* 20 (3): 241–254.

Löschel, Andreas, Matthias Rodemeier, and Madeline Werthschulte. 2020. "When Nudges Fail to Deliver: Field Experimental Evidence from Goal Setting on Smartphones." CESifo Working Paper No. 8485.

Matějka, Filip, and Alisdair McKay. 2015. "Rational inattention to discrete choices: A new foundation for the multinomial logit model." *American Economic Review* 105 (1): 272–98.

Miller, Klaus M, Navdeep S Sahni, and Avner Strulov-Shlain. 2022. "Sophisticated Consumers with Inertia: Long-Term Implications from a Large-Scale Field Experiment." Working Paper.

Müller, Karsten, and Carlo Schwarz. 2020. "Fanning the flames of hate: Social media and hate crime." *Journal of the European Economic Association*.

Rabin, Matthew. 2000. "Risk aversion and expected-utility theory: A calibration theorem." *Econometrica* 68 (5): 1281–1292.

Rhode Island General Laws. 2016. Commercial Law – General Regulatory Provisions. Chapter 6-13.1. Deceptive Trade Practices. Section 6-13.1-1-.

Rodemeier, Matthias, and Andreas Löschel. 2020. "Information Nudges, Subsidies, and Welfare: Evidence from a Natural Field Experiment." CESifo Working Paper No. 8259.

Silk, Tim. 2004. Examining Purchase and Non-redemption of Mail-in Rebates: The Impact of Offer Variables on Consumers' Subjective and Objective Probability of Redeeming. PhD dissertation, University of Florida 2004.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of Monetary Economics* 50 (3): 665–690.

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.

Taubinsky, Dmitry, and Alex Rees-Jones. 2017. "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment." *The Review of Economic Studies* 85 (4): 2462–2496.

Thaler, Richard H. 2018. "Nudge, not sludge." *Science* 361 (6401): 431.

Van de Ven, Wynand PMM, and Bernard MS Van Praag. 1981. "The demand for deductibles in private health insurance: A probit model with sample selection." *Journal of Econometrics* 17 (2): 229–252.

Online Appendix: Not for Publication

A Mathematical Appendix

A.1 Proof of Proposition 1

The probability of buying at the store can be written as

$$B(s, \hat{\theta}, \hat{c}) = \int^{s-\hat{c}} \int^{\hat{\theta}(s-\hat{c})+(1-\hat{\theta})\kappa} f(\epsilon|\kappa) d\epsilon h(\kappa) d\kappa + \int_{s-\hat{c}} \int^{\kappa} f(\epsilon|\kappa) d\epsilon h(\kappa) d\kappa. \tag{16}$$

For convenience, let $Q(s, \hat{\theta}, \hat{c}, \kappa) = \int^{\hat{\theta}(s-\hat{c})+(1-\hat{\theta})\kappa} f(\epsilon|\kappa) d\epsilon h(\kappa)$ and $M(\kappa) = \int^{\kappa} f(\epsilon|\kappa) d\epsilon h(\kappa)$. Then,

$$\begin{aligned}
\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c}) &= Q(s, \hat{\theta}, \hat{c}, s-\hat{c}) + \int^{s-\hat{c}} \frac{\partial Q(s, \hat{\theta}, \hat{c}, \kappa)}{\partial s} d\kappa - M(s-\hat{c}) \\
&= \underbrace{Q(s, \hat{\theta}, \hat{c}, s-\hat{c}) - M(s-\hat{c})}_{=0} + \int^{s-\hat{c}} \frac{\partial Q(s, \hat{\theta}, \hat{c}, \kappa)}{\partial s} d\kappa \\
&= \int^{s-\hat{c}} \frac{\partial Q(s, \hat{\theta}, \hat{c}, \kappa)}{\partial s} d\kappa \\
&= \hat{\theta} \int^{s-\hat{c}} f(\hat{\theta}(s-\hat{c}) + (1-\hat{\theta})\kappa|\kappa) h(\kappa) d\kappa \\
&\approx \hat{\theta} \int^{s-\hat{c}} f(s-\hat{c}|\kappa) h(\kappa) d\kappa,
\end{aligned}$$

which implies

$$\hat{\theta} \approx \frac{\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial s} B(s, 1, \hat{c})}. \tag{17}$$

The approximation requires that $f(\epsilon|\kappa)$ is roughly constant on $[\hat{\theta}(s-\hat{c}) + (1-\hat{\theta})\kappa, s-\hat{c}]$ for all κ.

Next, I derive sufficient statistics for perceived hassle costs. To a first-order approximation,

$$\Delta_{\hat{c}} B(s, \hat{\theta}, \hat{c}) \approx \Delta \hat{c} \frac{\partial}{\partial \hat{c}} B(s, \hat{\theta}, \hat{c}).$$

If the treatment fully eliminates hassle costs, then $\Delta \hat{c} = 0 - \hat{c}$, and to first order:

$$\hat{c} \approx -\frac{\Delta_{\hat{c}} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial \hat{c}} B(s, \hat{\theta}, \hat{c})} \tag{18}$$

$$= \frac{\Delta_{\hat{c}} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c})}. \tag{19}$$

To go from the first to the second line, I have used the fact that

$$\begin{aligned}\frac{\partial}{\partial \hat{c}} B(s, \hat{\theta}, \hat{c}) &= Q(s, \hat{\theta}, \hat{c}, s - \hat{c}) + \int^{s-\hat{c}} \frac{\partial Q(s, \hat{\theta}, \hat{c}, \kappa)}{\partial \hat{c}} d\kappa - M(s - \hat{c}) \\ &= -\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c}).\end{aligned}$$

This proves the first part of the proposition. To derive the sufficient statistics for the true redemption frictions, recall that the unconditional redemption probability is given by

$$R(s, \theta, c) = \theta \int^{s-c} dH(\kappa). \tag{20}$$

It immediately follows that

$$\theta = \frac{R(s, \theta, c)}{R(s, 1, c)}. \tag{21}$$

An alternative way to identify θ relies on a comparison of redemption elasticities with and without inattention. Note that a very small change in the rebate size changes the redemption probability by

$$\frac{\partial R(s, \theta, c)}{\partial s} = \theta h(s - c),$$

which implies that

$$\theta = \frac{\frac{\partial R(s,\theta,c)}{\partial s}}{\frac{\partial R(s,1,c)}{\partial s}}. \tag{22}$$

Hassle costs can be approximated to first order by

$$c \approx -\frac{\Delta_c R(s, \theta, c)}{\frac{\partial R(s,\theta,c)}{\partial c}} \tag{23}$$

$$= \frac{\Delta_c R(s, \theta, c)}{\frac{\partial R(s,\theta,c)}{\partial s}}, \tag{24}$$

where I have used the fact that

$$\frac{\partial R(s,\theta,c)}{\partial c} = -\theta h(s-c)$$
$$= -\frac{\partial R(s,\theta,c)}{\partial s}.$$

Recall that consumers are sophisticated if and only if $\hat{\theta} = \theta$ and $\hat{c} = c$. Comparing equation 17 with equation 21, and equation 19 with equation 24, implies consumers are sophisticated if and only if

$$\frac{\frac{\partial}{\partial s} B(s,\hat{\theta},\hat{c})}{\frac{\partial}{\partial s} B(s,1,\hat{c})} \approx \frac{R(s,\theta,c)}{R(s,1,c)} \tag{25}$$

and

$$\frac{\Delta_{\hat{c}} B(s,1,\hat{c})}{\frac{\partial}{\partial r} B(s,1,\hat{c})} \approx \frac{\Delta_c R(s,1,c)}{\frac{\partial}{\partial s} R(s,1,c)}. \tag{26}$$

This completes the proof. □

A.2 Identifying the Subjective Redemption Probability from Demand Responses to Rebates and Price Reductions

An intuitive approach of identifying the subjective redemption probability might be to compare the demand response to a rebate with that of a price reduction. As explained in the main text, the intuition would be that a 2 USD rebate with $R(s,\hat{\theta},\hat{c}) = 0.5$ should increase demand by the same amount as 1 USD reduction price. In my empirical setting, we could then simply compare the demand response to the a typical rebate with the demand response to the automatically applied discount. In the notation of the model, we simply invert the relationship above and approximate $R(s,\hat{\theta},\hat{c})$ by the ratio of demand responses:

$$R(s,\hat{\theta},\hat{c}) = \frac{\frac{\partial B(s,\hat{\theta},\hat{c})}{\partial s}}{\frac{\partial B(s,1,0)}{\partial s}}. \tag{27}$$

However, as can be verified below this identification strategy relies on an implicit and potentially strong assumption about the distribution of marginal consumers.

Formally, the claim is that the following relationship can be used to identify $R(s,\hat{\theta},\hat{c})$:

$$\frac{\partial B(s,1,0)}{\partial s} \times R(s,\hat{\theta},\hat{c}) = \frac{\partial B(s,\hat{\theta},\hat{c})}{\partial s} \tag{28}$$

$$\iff \int^{s} f(s|\kappa)h(\kappa)d\kappa \times \hat{\theta}\int^{s-\hat{c}} h(\kappa)d\kappa = \hat{\theta}\int^{s-\hat{c}} f(s-\hat{c}|\kappa)h(\kappa)d\kappa \tag{29}$$

$$\iff \underbrace{f(s|\kappa \leq s)Pr(\kappa \leq s)}_{\substack{\text{marginal consumers intending}\\ \text{to redeem in absence of frictions}}} R(s,\hat{\theta},\hat{c}) = \underbrace{\hat{\theta}f(s-\hat{c}|\kappa \leq s-\hat{c})Pr(\kappa \leq s-\hat{c})}_{\substack{\text{marginal consumers intending}\\ \text{to redeem in presence of frictions}}} \tag{30}$$

Thus, this equality only holds under a special distributional property, which depends both on ϵ and κ. The left-hand side consists of two parts. The first part is the density of consumers who think they will redeem the rebate in the absence of redemption frictions ($\hat{\theta} = 1$, $\hat{c} = 0$) and who, at the same time, are at the margin to the automatically applied rebate, i.e. have $\epsilon = s$. The second part is simply the subjective redemption probability. The right-hand side is the density of consumers who both think they redeem the rebate in the presence of redemption frictions and who are marginal to this rebate. Thus, the equation says that the density of marginal consumers thinking they redeem in the absence of redemption frictions, weighted by the subjective redemption probability, must equal the density of marginal consumers thinking they redeem in the presence of redemption frictions. The subjective redemption probability can only be identified if the condition in equation 30 holds. The identification strategy in Proposition 1 does not require this additional assumption and is, therefore, more general.

A.3 Model with Heterogeneity in Redemption Frictions

In the main part of the paper, behavioral frictions are homogeneous. In this section, I introduce heterogeneity in perceived and true inattention and hassle costs, respectively. I show that perceived and true inattention are still identified by the same aggregate demand elasticities in Proposition 1, when they are independent of the taste parameters. By contrast, hassle costs are only identified if they are roughly homogenous.

It is important to highlight that heterogeneity only affects the identification of the structural parameters, not the reduced-form test of sophistication explained in Section 4.5.1 in the main part of the paper.

To introduce heterogeneity in inattention, let $L_{\hat{\theta}}(\hat{\theta})$ and $P_{\theta}(\theta)$ denote the marginal distributions of perceived and true inattention, respectively. Assume that both distributions are smooth and that perceived and true inattention are independent of the idiosyncratic taste parameters, κ and ϵ. $B(\hat{\theta})$ and $R(\theta)$ are now the buying and redemption probability for a given realization of $\hat{\theta}$ and θ, respectively.

The effect of a small change in the rebate value on aggregate demand is therefore

$$\mathbb{E}\left[\frac{\partial B(s,\hat{\theta},\hat{c})}{\partial s}\right] = \int \frac{\partial}{\partial s} B(s,\hat{\theta},\hat{c}) dL_{\hat{\theta}}(\hat{\theta}). \tag{31}$$

Using the same derivation to arrive at equation 17, it follows that the expectation of perceived inattention can be identified by aggregate demand elasticities:

$$\mathbb{E}[\hat{\theta}] \approx \frac{\mathbb{E}\left[\frac{\partial}{\partial s} B(s,\hat{\theta},\hat{c})\right]}{\frac{\partial}{\partial s} B(s,1,\hat{c})}. \tag{32}$$

Similarly, using equation 21, it immediately follows that the expectation of true inattention is identified by aggregate redemption probabilities:

$$\mathbb{E}\left[\theta\right] = \frac{\mathbb{E}[R(s,\theta,c)]}{R(s,1,c)}. \tag{33}$$

These results show that perceived and true inattention are identified by the same aggregate buying and redemption behavior as in Proposition 1.

Next, consider the case in which hassle costs are heterogeneous. Let $L_{\hat{c}}(\hat{c})$ and $P_c(c)$ denote the marginal distribution of $\hat{c}$ and c, respectively, and assume that both distributions are smooth and independent to the idiosyncratic taste parameters. The aggregate demand response to a change in perceived hassle costs is approximated to first order by

$$\mathbb{E}\left[\Delta_{\hat{c}} B(s,\hat{\theta},\hat{c})\right] \approx \int \Delta\hat{c} \frac{\partial}{\partial \hat{c}} B(s,\hat{\theta},\hat{c}) dP_{\hat{c}}(\hat{c}), \tag{34}$$

which is generally not equal to $\mathbb{E}[\Delta\hat{c}]\mathbb{E}[\frac{\partial}{\partial \hat{c}} B(s,\hat{\theta},\hat{c})]$. The demand response for consumer types with a given $\hat{c}$ depends on both the type-specific change in their perceived hassle costs and the type-specific buying elasticity. Since both $\Delta\hat{c}$ and $\frac{\partial}{\partial \hat{c}} B(s,\hat{\theta},\hat{c})$ vary with $\hat{c}$, the expectation of the product is not equal to the product of the individual expectations. Thus, we cannot re-arrange terms and use the same identification strategy as in equation 19. An analogous argument can be made when true hassle costs are heterogeneous by taking the expectation of both sides of equation 24: changes in aggregate redemption probabilities are not sufficient to identify expected hassle costs.

In sum, the structural identification strategy of perceived and true inattention is robust to the introduction redemption frictions, but hassle costs are only identified structurally if they are approximately homogenous. These results hold as long as redemption frictions are independent of the idiosyncratic taste parameters, i.e. of preferences.

A.4 Model with Ad Valorem Rebate

Proposition 1 was derived using a lump sum rebate of value s, whereas the experimental design uses an ad valorem rebate. In this section, I show that the same predictions from Proposition 1 can be derived with an ad valorem rebate. The difference between the two types of rebates is that it is more involved to model an ad valorem rebate because the rebate value depends on the endogenous purchase value of the consumer.

Let t denote an ad valorem rebate. The value of the rebate is given by $t\mathbf{p}'\mathbf{x}$ where $\mathbf{p} = (p^1, p^2, ..., p^J)$ is the vector of prices and $\mathbf{x} = (x^1, x^2, ..., x^J)$ the consumption vector. Different to a lump-sum rebate, an ad valorem rebate changes the optimal consumption vector because it effectively changes the price of each good. We therefore need to model a third margin where the consumption vector is a function of the rebate.

Let $\mathbf{x}_r$ be the chosen consumption vector given redemption choices r. Given the consumer buys at the store and is attentive, she chooses

$$\mathbf{x}_r = \arg\max_{\mathbf{x}} \left\{ v(\mathbf{x}) - \mathbf{p}'\mathbf{x} + r\left(t\mathbf{p}'\mathbf{x} - \hat{c}\right) \right\}$$

If she is not attentive, she chooses the same consumption vector as if she was attentive but decided not to redeem the rebate, i.e. $\mathbf{x}_0$. The first-order conditions are

$$\frac{\partial v}{\partial x^j} - p^j + rtp^j = 0$$

for every good j.

Given the consumer buys at the store and is attentive, she chooses $r = 1$ if and only if

$$\begin{gathered} v(\mathbf{x_1}) - \mathbf{p}'\mathbf{x_1} + t\mathbf{p}'\mathbf{x_1} - \hat{c} \geq v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0} + \kappa \\ \Leftrightarrow u(t, \hat{c}) \geq \kappa \end{gathered}$$

with $u(t, \hat{c}) = v(\mathbf{x_1}) - \mathbf{p}'\mathbf{x_1} + t\mathbf{p}'\mathbf{x_1} - \hat{c} - (v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0})$.

She chooses to buy at the store if and only if

$$\hat{\theta}\left\{ r\left(v(\mathbf{x_1}) - \mathbf{p}'\mathbf{x_1} + t\mathbf{p}'\mathbf{x_1} - \hat{c}\right) + (1-r)\left(v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0} + \kappa\right)\right\} + \left(1 - \hat{\theta}\right)\left\{v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0} + \kappa\right\} \geq \epsilon.$$

For convenience, let $w_1(t, \hat{\theta}, \hat{c}, \kappa) = \hat{\theta}\left\{\left(v(\mathbf{x_1}) - \mathbf{p}'\mathbf{x_1} + t\mathbf{p}'\mathbf{x_1} - \hat{c}\right)\right\} + \left(1 - \hat{\theta}\right)\left\{v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0} + \kappa\right\}$ and $w_0(\kappa) = v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0} + \kappa$. The probability to buy at the store can be expressed by

$$B(t, \hat{\theta}, \hat{c}) = \int^{u(t,\hat{c})} \int^{w_1(t,\hat{\theta},\hat{c},\kappa)} f(\epsilon|\kappa) d\epsilon h(\kappa) d\kappa + \int_{u(t,\hat{c})} \int^{w_0(\kappa)} f(\epsilon|\kappa) d\epsilon h(\kappa) d\kappa.$$

Let $Q(t,\hat{\theta},\hat{c},\kappa)=\int^{w_1(t,\hat{\theta},\hat{c},\kappa)} f(\epsilon|\kappa)d\epsilon h(\kappa)$ and $M(\kappa)=\int^{w_0(\kappa)} f(\epsilon|\kappa)d\epsilon h(\kappa)$. The effect of a very small change in the rebate value on the buying probability is given by

$$\begin{aligned}\frac{\partial}{\partial t}B(t,\hat{\theta},\hat{c}) &= \frac{\partial u}{\partial t}Q(t,\hat{\theta},\hat{c},u)+\int^{u}\frac{\partial Q(t,\hat{\theta},\hat{c},\kappa)}{\partial t}d\kappa-\frac{\partial u}{\partial t}M(u)\\ &= \frac{\partial u}{\partial t}\Big(Q(t,\hat{\theta},\hat{c},u)-M(u)\Big)+\int^{u}\frac{\partial Q(t,\hat{\theta},\hat{c},\kappa)}{\partial t}d\kappa.\end{aligned}$$

Note that

$$\begin{aligned}w_1(t,\hat{\theta},\hat{c},u) &= \hat{\theta}\left\{(v(\mathbf{x_1})-\mathbf{p}'\mathbf{x_1}+t\mathbf{p}'\mathbf{x_1}-\hat{c})\right\}+\left(1-\hat{\theta}\right)\left\{v(\mathbf{x_0})-\mathbf{p}'\mathbf{x_0}+u\right\}\\ &= \hat{\theta}\left\{(v(\mathbf{x_1})-\mathbf{p}'\mathbf{x_1}+t\mathbf{p}'\mathbf{x_1}-\hat{c})\right\}+\left(1-\hat{\theta}\right)\left\{[v(\mathbf{x_1})-\mathbf{p}'\mathbf{x_1}+t\mathbf{p}'\mathbf{x_1}-\hat{c}]\right\}\\ &= v(\mathbf{x_1})-\mathbf{p}'\mathbf{x_1}+t\mathbf{p}'\mathbf{x_1}-\hat{c}\\ &= v(\mathbf{x_0})-\mathbf{p}'\mathbf{x_0}+u\\ &= w_0(u).\end{aligned}$$

Therefore, $Q(t,\hat{\theta},\hat{c},u)-M(u)=0$ and

$$\begin{aligned}\frac{\partial}{\partial t}B(t,\hat{\theta},\hat{c}) &= \int^{u}\frac{\partial Q(t,\hat{\theta},\hat{c},\kappa)}{\partial t}d\kappa\\ &= \int^{u}\frac{\partial w_1(t,\hat{\theta},\hat{c},\kappa)}{\partial t}f(w_1(t,\hat{\theta},\hat{c},\kappa)|\kappa)h(\kappa)d\kappa\\ &= \int^{u}\hat{\theta}\left((v_{\mathbf{x}_1}-\mathbf{p}')\frac{\partial\mathbf{x}_1}{\partial t}+(t\mathbf{p}'\frac{\partial\mathbf{x}_1}{\partial t}+\mathbf{p}'\mathbf{x}_1)\right)f(w_1(t,\hat{\theta},\hat{c},\kappa)|\kappa)h(\kappa)d\kappa\\ &= \hat{\theta}\mathbf{p}'\mathbf{x}_1\int^{u}f(w_1(t,\hat{\theta},\hat{c},\kappa)|\kappa)h(\kappa)d\kappa.\end{aligned}$$

If f is roughly constant on the interval $[w_1(t,\hat{\theta},\hat{c},\kappa),w_1(t,1,\hat{c},\kappa)]$, then

$$\hat{\theta}\approx\frac{\frac{\partial}{\partial t}B(t,\hat{\theta},\hat{c})}{\frac{\partial}{\partial t}B(t,1,\hat{c})}.$$

To derive the sufficient statistics for perceived hassle costs, first note that a small change in perceived hassle costs changes the buying probability by

$$\begin{aligned}\frac{\partial}{\partial\hat{c}}B &= \frac{\partial u}{\partial\hat{c}}Q(t,\hat{\theta},\hat{c},u(\cdot))+\int^{u}\frac{\partial Q(t,\hat{\theta},\hat{c},\kappa)}{\partial\hat{c}}d\kappa-\frac{\partial u}{\partial\hat{c}}M(u(\cdot))\\ &= -\hat{\theta}\int^{u}f(w_1|\kappa)h(\kappa)d\kappa.\end{aligned}$$

This implies that

$$\frac{\partial}{\partial t}B(t,\hat{\theta},\hat{c}) = -\frac{\partial}{\partial \hat{c}}B(t,\hat{\theta},\hat{c})\mathbf{p}'\mathbf{x}_1.$$

To a first-order approximation,

$$\begin{aligned}\Delta_{\hat{c}}B(t,\hat{\theta},\hat{c}) &\approx \frac{\partial}{\partial \hat{c}}B(t,\hat{\theta},\hat{c})\Delta\hat{c}\\ \Delta\hat{c} &\approx \frac{\Delta_{\hat{c}}B(t,\hat{\theta},\hat{c})}{\frac{\partial B(t,\hat{\theta},\hat{c})}{\partial \hat{c}}}.\end{aligned}$$

If $\Delta\hat{c} = 0 - \hat{c}$, then:

$$\begin{aligned}\hat{c} &\approx -\frac{\Delta_{\hat{c}}B(t,\hat{\theta},\hat{c})}{\frac{\partial B(t,\hat{\theta},\hat{c})}{\partial \hat{c}}}\\ &= \frac{\Delta_{\hat{c}}B(t,\hat{\theta},\hat{c})}{\frac{\partial B(t,\hat{\theta},\hat{c})}{\partial t}}\mathbf{p}'\mathbf{x_1}.\end{aligned}$$

As we can see, both $\hat{\theta}$ and $\hat{c}$ are identified in the same way as in Proposition 1 but s is replaced by $t\mathbf{p}'\mathbf{x}_1$.

Next, I derive the sufficient statistics for the true redemption frictions. The redemption probability is given by:

$$R(t,\theta,c) = \theta\int^{u} dH(\kappa)$$

such that inattention is identified by

$$\theta = \frac{R(t,\theta,c)}{R(t,1,c)}.$$

To identify true hassle costs first note that

$$\begin{aligned}\frac{\partial R(t,\theta,c)}{\partial c} &= -\theta h(u)\\ &= -\frac{\partial R(t,\theta,c)}{\partial t}(\mathbf{p}'\mathbf{x_1})^{-1}\end{aligned}$$

which implies that, to first order,

$$\begin{aligned}
\Delta_c R(t,\theta,c) &\approx \frac{\partial R(t,\theta,c)}{\partial c}\Delta c \\
&= -\frac{\partial R(t,\theta,c)}{\partial t}(\mathbf{p}'\mathbf{x}_1)^{-1}\Delta c \\
\Leftrightarrow c &\approx \frac{\Delta_c R(t,\theta,c)}{\frac{\partial R(t,\theta,c)}{\partial t}}\mathbf{p}'\mathbf{x}_1
\end{aligned}$$

when $\Delta c = -c$.

A.5 GMM Estimation

A.5.1 Estimation

To obtain values of the perceived redemption frictions, I estimate the structural parameters by substituting the regression coefficients in equation 14 with the structural parameters and then solve a set of moment conditions.

Denote the buying probability in the control group by β_D. Let the treatment effect on the buying probability by treatment $t \in \{A, B.1, B.2b, C.1, C.2b\}$ be denoted by β_t. As I show below we can rewrite these reduced-form treatment effects in terms of the structural parameters. This reformulation results in the following six moment conditions for demand:

$$\mathbb{E}\left[\mathbf{I}_i\left(Buy_i - \beta_A \times A_i - \beta_{B_1} \times B.1_i - \beta_{B_2.b}B.2b_i - \left(\frac{\hat{\theta}\Delta s}{\hat{c}}(\beta_A - \beta_{B.2b}) + \beta_{B.1}\right) \times C.1_i \right.\right.$$
$$\left.\left. - \left(\frac{\Delta s}{\hat{c}}(\beta_A - \beta_{B.2b}) + \beta_{B.2b}\right) \times C.2b_i - \beta_D\right)\right] = 0, \quad (35)$$

where $\mathbf{I}_i = (A_i, B.1_i, B.2b_i, C.1_i, C.2b_i, D_i)$ is the 6×1 vector of instruments indicating the experimental group of subject i. The monetary change in the rebate value, Δs, is calculated by multiplying the additional five percentage points (of the 15%- relative to the 10%-rebate) by the median shopping basket value in group A, which is 96 EUR. Thus, $\Delta s = 4.80$ EUR.[40] Since the number of parameters to be estimated is also six, the model is exactly identified.

Next, true inattention and hassle costs are estimated using moments of redemption behavior. Again assuming independence between the residuals of the buying and redemption equation, I can

[40] See Appendix A.4 for a formal proof that it is possible to translate an ad-valorem rebate to a lump-sum rebate in this way. Instead of the mean I use the median to adjust for outliers with very large shopping basket values.

rewrite the reduced-form treatment effect parameters in terms of the underlying structural choice parameters, which yields the following five moment conditions:

$$\mathbb{E}\Bigg[\mathbf{J}_i\Bigg(Redeem_i - \tau_A - \left(\theta(\tau_A - c\frac{\partial R(s,1,c)}{\partial s}) - \tau_A\right) \times B.1_i + c\frac{\partial R(s,1,c)}{\partial s} \times B.2b_i$$
$$- \left(\theta\left[\tau_A + \frac{\partial R(s,1,c)}{\partial s}(\Delta s - c)\right] - \tau_A\right) \times C.1_i - \left(\frac{\partial R(s,1,c)}{\partial s}(\Delta s - c)\right) \times C.2b_i\Bigg)\Bigg] = 0, \tag{36}$$

with $\mathbf{J}_i = (A, B.1, B.2b, C.1, C.2b_i)$ denoting the vector of instruments, excluding the control group. With five moments and four parameters, the model is overidentified and I use a two-step GMM estimator to find the optimal weight matrix.

A.5.2 Derivation of Moment Conditions

Perceived hassle costs are approximated by

$$\hat{c} \approx \frac{\Delta_{\hat{c}} B(s,1,\hat{c})}{\frac{\partial}{\partial s} B(s,1,\hat{c})} \approx \frac{\beta_A - \beta_{B.2b}}{\frac{\beta_{C.2b} - \beta_{B.2b}}{\Delta s}} \tag{37}$$

$$\Leftrightarrow \beta_{C.2b} \approx \frac{\Delta s}{\hat{c}}(\beta_A - \beta_{B.2b}) + \beta_{B.2b}. \tag{38}$$

Perceived inattention is approximated by

$$\hat{\theta} \approx \frac{\beta_{C.1} - \beta_{B.1}}{\beta_{C.2b} - \beta_{B.2b}} \tag{39}$$

$$\Leftrightarrow \beta_{C.1} \approx \hat{\theta}(\beta_{C.2b} - \beta_{B.2b}) + \beta_{B.1} \tag{40}$$

$$= \hat{\theta}\frac{\Delta s}{\hat{c}}(\beta_A - \beta_{B.2b}) + \beta_{B.1}, \tag{41}$$

where the last line follows from substituting $\beta_{C.2b}$ from equation 38.

Under this assumption, true hassle costs are identified by

$$c = \frac{\Delta_c R(s,1,c)}{\frac{\partial R(s,1,c)}{\partial s}} = \frac{-\tau_{B2b}}{\frac{\partial R(s,1,c)}{\partial s}} \tag{42}$$

$$\Leftrightarrow \tau_{B2b} = -c\frac{\partial R(s,1,c)}{\partial s}. \tag{43}$$

$$\tag{44}$$

We can insert this into the expression for τ_{C2b}:

$$\begin{aligned}
\tau_{C2b} &= \tau_{B2b} + \Delta s \frac{\partial R(1,c,s)}{\partial s} \\
&= \frac{\partial R(s,1,c)}{\partial s} (\Delta s - c)
\end{aligned}$$

This yields the moment conditions in equation 35.

True inattention can be identified in multiple ways. The first identification strategy relies on the comparison between redemption probabilities with and without inattention:

$$\begin{aligned}
\theta &= \frac{R(s,\theta,c)}{R(s,1,c)} \approx \frac{\tau_{B1} + \tau_A}{\tau_A + \tau_{B2b}} \\
\Leftrightarrow \tau_{B1} &\approx \theta(\tau_A + \tau_{B2b}) - \tau_A \\
&= \theta(\tau_A - c\frac{\partial R(s,1,c)}{\partial s}) - \tau_A.
\end{aligned}$$

The second identification strategy of inattention relies on the comparison of demand derivatives:

$$\theta = \frac{\frac{\partial}{\partial s} R(s,\theta,c)}{\frac{\partial}{\partial s} R(s,1,c).}$$

We can insert this condition into the expression for τ_{C1}:

$$\begin{aligned}
\tau_{C1} &= \tau_{B1} + \Delta s \frac{\partial R(s,\theta,c)}{\partial s} \\
&= \tau_{B1} + \theta \Delta s \frac{\partial R(s,1,c)}{\partial s} \\
&= \theta(\tau_A - c\frac{\partial R(s,1,c)}{\partial s}) - \tau_A + \theta \Delta s \frac{\partial R(s,1,c)}{\partial s} \\
&= \theta \left[\tau_A + \frac{\partial R(s,1,c)}{\partial s} (\Delta s - c)\right] - \tau_A,
\end{aligned}$$

where in the last line, I have substituted for τ_{B1}. Rewriting τ_{B1} and τ_{C1} as above produces the moment conditions in equation 36.

B Pilot Studies

The first pilot study was implemented between July and August 2018 for a period of three weeks and with a smaller sample of 13,204 website visitors in the United Kingdom. Different from the main experiment, the study did not include users with mobile devices and tablets. Only subjects who used a desktop were randomized into one of the experimental groups. The study also only included a 10%- but not a 15%-rebate. The experiment included the following three treatment groups and one control group: group A, B.1, B.2a, and D. The treatments in which the reminder was explicitly announced at the start of the visit (B.2b and C.2b) were not included.

The second pilot study took place in August 2019 for a period of one week and with a larger sample of 52,302 consumers in the German online shop of the company. Just as in the first pilot study, the study included only desktop users, and the experimental design consisted of the groups A, B.1, B.2a, and D.

In both pilot studies the design of the banners and the rebate code differed on some dimensions in comparison to the one presented in the main part of the paper. The reason for the divergence is that the marketing department of the company sometimes changes the promotion design. The experiments I ran effectively changed certain features of these promotions but not the entire visual design. For privacy reasons, I cannot display the banners of the pilot studies in this paper. Overall, the conceptual design of the banners was similar to the banners presented in the main body of the paper.

Table B1 documents the results from linear probability models of the outcomes of interest on the treatment indicators. The baseline buying probability in the first pilot study with UK customers is around 3.4%, and therefore higher than for the sample analyzed in the main experiment of the paper. The automatically applied 10%-discount increases the buying probability by around 1.04 percentage points. Introducing hassle costs lowers this effect to 0.86 percentage points, but the difference in treatment coefficients is not statistically significant. Increasing inattention by removing the reminder further reduces the positive effect of the rebate down to 0.42 percentage points.

Looking at redemption rates, I observe a redemption probability of 95% for subjects who receive the automatically applied discount. Redemption rates fall much more steeply than in the main experiment of the paper. The introduction of hassle costs is associated with a decrease of 73 percentage points. Removing the reminder is associated with an additional decrease of 5 percentage points.

Even though the coefficients are different than for the main experiment of the paper, the qualitative results are similar: consumers substantially reduce demand when the firm tries to exploit inattention and removes the reminder. Hassle costs only slightly, and not significantly, decrease demand even though they are a much bigger redemption friction than inattention. Thus, inattention is *perceived* to be a substantially larger friction than hassle costs by consumers, even though the opposite is true.

Another interesting observation is that even though redemption frictions of the rebate appear to be much larger in this experiment, the demand response to the rebate is also substantially smaller than in the main experiment: requiring consumers to redeem the rebate reduces the positive effect of the promotion on demand by 60% (from 1.04 to 0.42 percentage points). Although differences across experiments are only correlational, this behavior is consistent with the interpretation that rebates with lower redemption probabilities cause smaller positive effects on demand.

In the second pilot study with subjects in Germany, the baseline buying probability equals 4.9%, and an automatically applied discount increases demand by 0.73 percentage points. Again, introducing hassle costs has no significantly different effect on demand. The coefficient is even slightly larger than the one in group A, but the difference in coefficients is likely a result of sampling variation. Removing the reminder substantially lowers the demand response to the rebate from 0.87 to 0.1 percentage points.

The baseline redemption probability is 77% and falls by 30 percentage points with the introduction of hassle costs. Removing the reminder has an additional negative effect of 13 percentage points.

Thus, also in this experiment, hassle costs and inattention both reduce redemption rates, but consumers only anticipate their inattention. While hassle costs are a larger friction than inattention, consumers believe the opposite.

Although the effects in the pilot studies are quantitatively not the same as the ones presented in the main experiment, the qualitative results are the same. Differences in point estimates are to be expected because the samples differed along many dimensions, such as the type of device consumers used to visit the website, the user's country of origin, and the year of the experiment. Other idiosyncratic differences are the visual designs of the promotions created by the marketing department.

Table B1: Buying and Redemption Probabilities in Pilot Studies

	Pilot Study 1		Pilot Study 2	
	(1)	(2)	(3)	(4)
	Buying Probability ×100	Redemption Probability ×100	Buying Probability ×100	Redemption Probability ×100
A: 10%, automatic	1.040**	95.172***	0.728***	77.166***
	(0.481)	(1.787)	(0.277)	(1.558)
B.1: 10%, w/o reminder	0.417	-77.850***	0.099	-43.376***
	(0.460)	(3.815)	(0.268)	(2.416)
B.2a: 10%, w/ reminder	0.862*	-72.795***	0.873***	-30.341***
	(0.472)	(3.928)	(0.277)	(2.393)
D: control	3.403***		4.889***	
	(0.317)		(0.189)	
Regression constant	D	A	D	A
N	13,204	415	52,302	2,140
Country	United Kingdom		Germany	
Year	2018		2019	
Sample	Desktop users only		Desktop users only	

Note: The table shows average treatment effects for the two pilot studies that preceded the main experiment in the paper. Average treatment effects are estimated from a linear probability model of the buying and redemption probability on the treatment indicators. The control group is excluded in columns 2 and 4 because control group subjects could not redeem the rebate by construction. Robust standard errors are in parentheses and clustered at the subject level. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

C Heterogeneity Based on Zip-Code Level Income Data

Table C1: Heterogeneity by Income

	(1) Buying Probability $\times 100$	(2) Redemption Probability $\times 100$
A: 10%, discount	0.452***	89.509***
	(0.125)	(1.259)
B: 10%, rebate	0.238**	-34.324***
	(0.122)	(2.493)
$\times$ reminder	0.037	12.833***
	(0.109)	(2.579)
C: 15%, rebate	0.426***	-32.439***
	(0.125)	(2.434)
$\times$ reminder	0.058	14.418***
	(0.115)	(2.468)
D: Control	1.793***	
	(0.084)	
Above median income	0.069	-1.781
	(0.119)	(1.947)
Interaction effects:		
A: 10%, discount $\times$ above median income	-0.031	
	(0.179)	
B: 10%, rebate $\times$ above median income	0.088	-1.048
	(0.175)	(3.591)
$\times$ reminder	0.066	-0.910
	(0.158)	(3.619)
C: 15%, rebate $\times$ above median income	0.002	4.578
	(0.179)	(3.486)
$\times$ reminder	0.173	-5.025
	(0.164)	(3.433)
N	420,857	8,305

Note: The table reports treatment effects for consumers from regions with an income below and above the sample median. Income data is based on the zip-code level, as described in the main part of the paper. Robust standard errors are in parentheses and clustered on the subject level. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

D Correlation Between Redemption Rates and Spending

Since the rebates used in the experiment are ad valorem, how much a consumers receives as a discount depends on the shopping basket value. For example, a 10% rebate that is redeemed translates to a lump sum rebate of 10 EUR for consumers who buy goods in the value of 100 EUR, but only to 1 EUR for those whose basket value is 10 EUR. It may, therefore, be interesting to know whether these two consumer types would respond differently to the rebate, as they receive largely different benefits from it.

The empirical challenge in answering this question is that I do not observe how much each consumer would have spent in the absence of the rebate. If a consumer buys goods in the value of 100 EUR, this *may* be the amount she would have spent in the absence of the promotion, but it is likely to be larger because goods demand may increase with the rebate. In fact, results in Table 4 suggest the latter as shopping basket values increase with rebates. Thus, it is not possible to identify whether consumers who have higher shopping basket values in the absence of the rebate exhibit larger treatment effects to the rebate.

For completeness, I report the *correlation* between shopping basket value and redemption rates across treatments. In particular, I re-run regression 15 but interact each treatment with a dummy that indicates whether the shopping basket value is above the median shopping basket value in the control group. Table D1 provides results.

In group A, the redemption rate for consumers with high spending is economically and statistically indistinguishable from consumers with low spending. For the 10% rebate, the redemption rate is 11 percentage points larger for high spending consumers. Similarly, it is 8 percentage points larger for high spending consumers at the 15% rebate. Interaction terms with the reminder and the announcement are statistically zero at both the 10%- and 15%-rebate value.

Consistent with basic economic principles, these patterns may suggest that consumers for whom the (lump sum) value of the rebate is larger are also the once who are more likely to redeem the rebate. An alternative interpretation is that the interaction terms are entirely driven by the fact that rebates *cause* the shopping basket value to increase, and that those who redeem the rebate are also those who have been induced to buy more.

Table D1: Redemption Rates for Consumers with High and Low Shopping Basket Values

	(1) Redemption Probability $\times 100$
A: 10% discount	87.558***
	(1.207)
B: 10%, rebate	-40.061***
	(2.139)
× reminder	10.270***
	(2.552)
× announcement	3.862
	(2.494)
C: 15%, rebate	-35.189***
	(2.137)
× reminder	11.396***
	(2.380)
× announcement	3.500
	(2.249)
Above basket value of control group median	0.351
	(1.664)
Interaction effects:	
B: 10%, rebate × above basket value of control group median	11.006***
	(2.999)
× reminder	-1.082
	(3.495)
× announcement	3.225
	(3.354)
C: 15%, rebate × above basket value of control group median	8.435***
	(2.924)
× reminder	1.903
	(3.219)
× announcement	-0.776
	(2.999)
Regression constant	A
N	11,872

Note: The table reports results from a linear regression that adds interaction terms between the treatments and a dummy indicating whether the shopping basket value is above the median shopping basket value of the control group. Robust standard errors are in parentheses and clustered on the subject level. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

E Customer Loyalty

I estimate the effects of the treatment on the probability of buying more than once during the experimental period. The outcome variable is a dummy equal to 1 if the consumer purchased twice or more, and 0 otherwise. In the regression, the constant represents the mean of group A that received the automatically applied discount. All treatment coefficients are therefore interpreted relative to an automatically-applied discount.

Table E1 reports the results. In the group with the automatically applied discount, 2.8% of all buyers make a second purchase during the experimental period. All other coefficients are statistically insignificant. There is no clear indication that rebates have a negative effect on customer loyalty relative to discounts.

However, some effect sizes are relatively large. The standard 10%-rebate (B.1) has a negative coefficient suggesting that the repurchase probability is 0.51 percentage points lower for a rebate than for an equivalent discount. Interestingly, the coefficient for the reminder is positive and almost equal in absolute size to the standard rebate. This could suggest that the negative effect of the rebate on customer loyalty is completely offset if the firm offers a reminder. The result would be consistent with the interpretation that consumers are aware of their inattention but remain naive about hassle costs even after this naiveté has been exploited.

The directional effects are the same for the 15%-rebate but different in magnitude: the negative effect of the rebate is smaller, while the positive effect of the reminder is also smaller.

Finally, an important result is that the most negative coefficient is the one of the control group where subjects did not receive a promotion. This suggests that even though rebates may have negative effects on customer loyalty relative to an automatically applied discount, they still increase customer loyalty relative to offering no price promotion.

Table E1: Repurchase Probabilities

	Probability to Buy more than once (in %)
B: 10%, rebate	-0.510
	(0.557)
× reminder	0.539
	(0.557)
× announcement	-0.215
	(0.572)
C: 15%, rebate	-0.438
	(0.552)
× reminder	0.158
	(0.523)
× announcement	-0.282
	(0.510)
D: control	-0.623
	(0.572)
A (constant)	2.842***
	(0.409)
N	13,224

Note: The table reports average treatment effects on the probability of purchasing more than once. The regression constant is the mean of group A. Robust standard errors are in parentheses. *,**,***: significant at $p < 0.1$, $p < 0.05$, $p < 0.01$, respectively.

F Effect of Announcement on Engagement with the Rebate

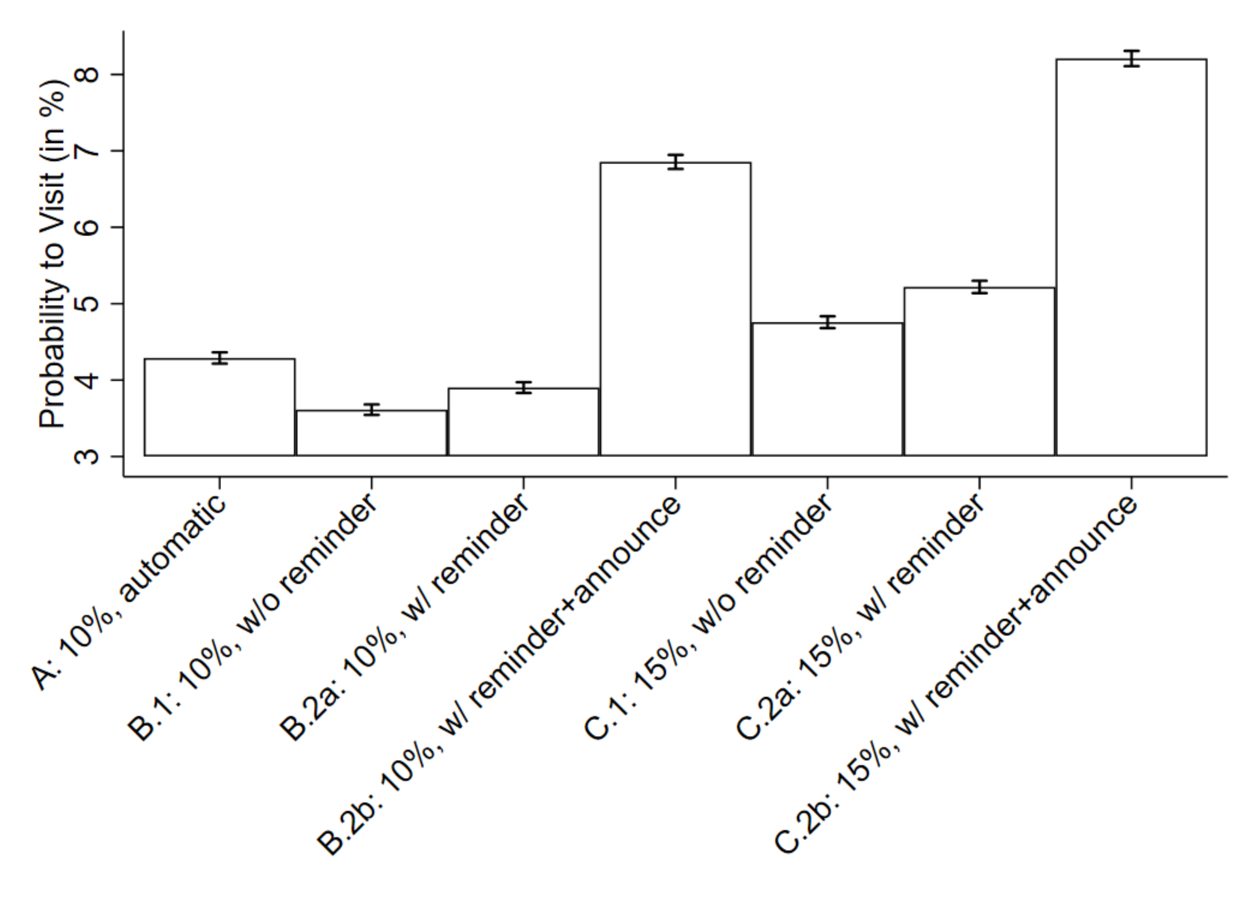
Figure F1 plots the probability of visiting the rebate page at least once that involves the necessary information about how to redeem the rebate, including the rebate code. For groups B.1, B.2a, and B.2b, this is the page shown in Figure 3. For groups C.1, C2.a, and C2.b, the page looks the same but the rebate size is larger. Group A can also visit a rebate page, but the page only informs them again that the rebate is automatically redeemed, and therefore involves no additional information in comparison to the information button already provided in the banner.

In group B.1, around 3.6% of website visitors browse to the rebate page. The number is slightly larger for group B.2a, where 3.9% of subjects visit the rebate page. When the reminder is announced, this number increases to 6.9%. The announcement also increases the probability of visiting the rebate page for the 15%-discount. The probability increases from 5.23% (C.2a) to 8.21% (C.2b) in response to the announcement. The absolute treatment effect of the announcement is highly statistically significant ($p < 0.01$) at both rebate sizes and remarkably similar in absolute size across stakes (around +2.9 percentage points for both rebate values).

Figure F2 plots the probability of visiting the checkout page and provides complementary evidence. The announcement substantially increases the probability of clicking on the checkout page and, thereby, viewing the reminder.

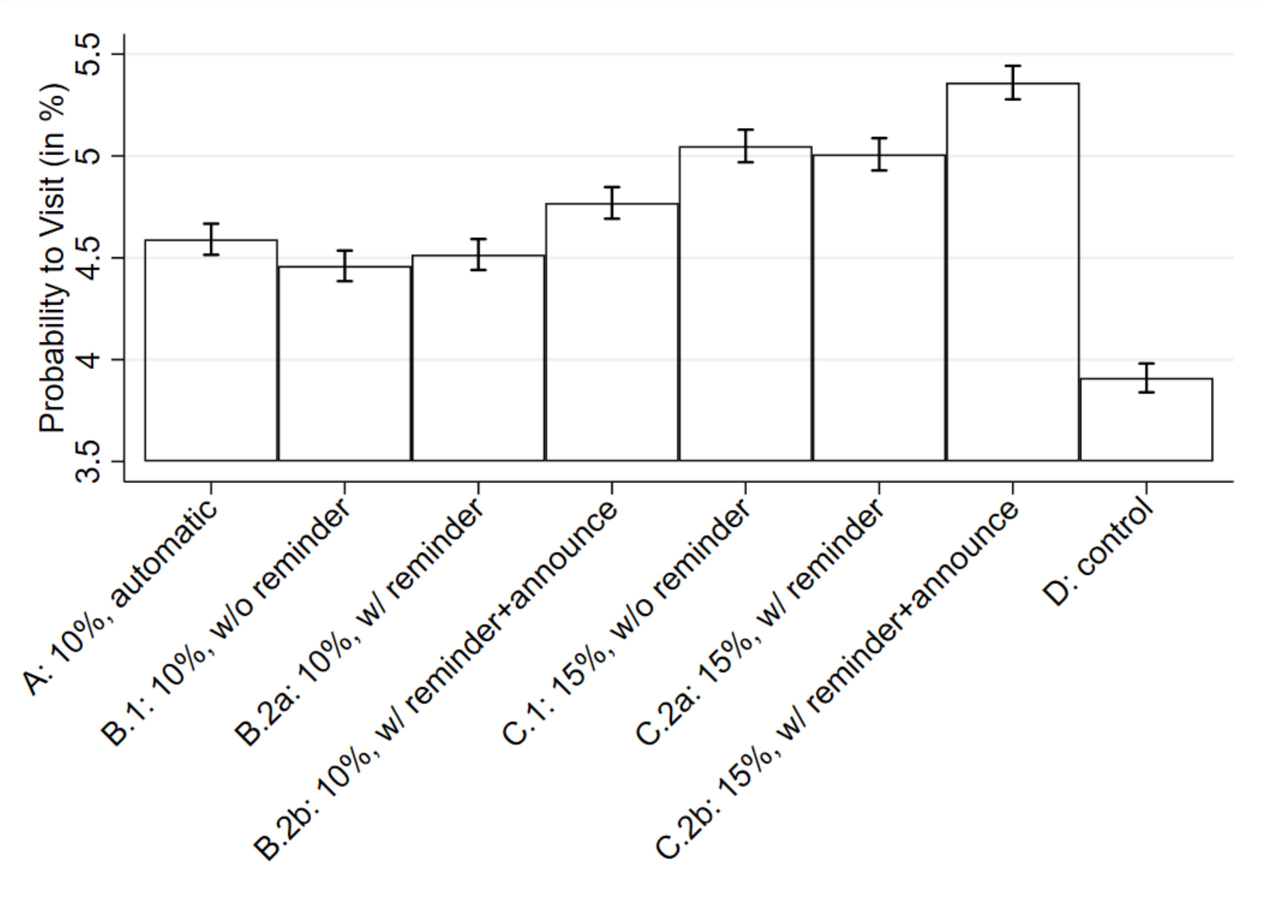
Subjects who receive the announcement are, therefore, substantially more likely to engage with the rebate and pay attention to the reminder.

Figure F1: Probability to Visit Rebate Page



Note: This figure shows the probability of visiting the rebate page at least once. The error bars represent standard errors.

Figure F2: Probability to Visit Checkout Page



Note: This figure shows the probability of visiting the checkout page at least once. The error bars represent standard errors.

G Sample Selection Model

The selection model uses regional and temporal variation in internet outages as an exclusion restriction. I use publicly available data on internet outages from Heise Online, a platform that documents user complaints about internet outages received by phone across the country. The dataset includes, among other variables, the area code and the duration of the outage. For the experimental observations, I only observe the city of each website visitor and not the area code. To merge internet outages with the dataset from the experiment, I use geo data from OpenGeoDB to assign each area code to a respective city. This approach allows me to assign internet outages collected from Heise Online to website visitors in the experiment.

One could use various approaches to construct a dummy variable that indicates whether a city experienced a major internet outage. In constructing the variable, I closely follow Müller and Schwarz (2020), who have used outages as exogenous variation in a different setting. Specifically, they study the effect of social media utilization on hate crime, and use internet outages as exogenous variation for access to social media. Following their approach, I count the total number of internet outages that occurred in the city of the website visitor. Because larger cities will have more internet outages mechanically, the authors normalize the number of internet outages by the number of inhabitants of each city, and I follow their approach. I then create a dummy variable that indicates whether the subject's area experienced a major internet outage. I define major internet outages as the 90th percentile of total internet outages normalized by the number of inhabitants. Because internet outages may also affect whether subjects even appear in my dataset (another level of sample selection), I only count internet outages that happened after the subject's *first* visit to the website during the experimental period. To avoid that subjects who visit at a later point in time have a lower number of outages mechanically, I count internet outages for each subject seven days after their first visit. Thus, even for subjects whose first visit was during the last day of the experiment, the following seven days are accounted for in terms of outages.

The sample selection model follows the standard setup introduced by Van de Ven and Van Praag (1981) when the dependent variables of both the selection and the outcome equation are binary. With some abuse of notation, I denote the buying decision of subject i by b_i and her rebate redemption choice by r_i. The utility from buying at the shop is given by

$$u_i = \gamma \mathbf{Z}_i + \iota \mathbf{X}_i + \eta_i, \tag{45}$$

where $\mathbf{X}_i$ is a vector of control variables, including a dummy for the device the subject uses (desktop, tablet, or smartphone) and date fixed effects. The latent utility component is denoted by η_i. The vector $\mathbf{Z}_i$ includes an indicator for each treatment and the instrument indicating whether the city of subject i experiences a major internet outage. In addition, the vector includes interaction terms between the instrument and the average income of region the subject is visiting from. Including interaction terms is important because it reduces the degree of collinearity between the treatment regressors in the outcome equation and the correction term. A high degree of collinearity is a well-known disadvantage of sample selection models, which causes inflated standard errors. Collinearity is a particular limitation in my application, because all treatments need to appear in both the selection and outcome equation. Allowing for the effect of internet outages to vary by income group adds a substantial degree of flexibility and increases precision of the point estimates on the intensive margin.

Utility from rebate redemption equals

$$v_i = \omega \mathbf{T}_i + \chi \mathbf{X}_i + \zeta_i, \tag{46}$$

where ζ_i is the unobserved utility from rebate redemption and $\mathbf{X}_i$ includes the same control variables as on the extensive margin.The vector $\mathbf{T}_i$ includes the treatment dummies and does *not* include internet outages.

Subject i's buying decision is given by

$$b_i = \begin{cases} 1 \text{ if } u_i > 0 \\ 0 \text{ otherwise }. \end{cases}$$

Her redemption choice is determined by the intensive margin utility and only observed if she buys:

$$r_i = \begin{cases} 1 \text{ if } v_i > 0 \text{ and } b_i = 1 \\ 0 \text{ if } v_i \leq 0 \text{ and } b_i = 1 \\ 0 \text{ if } b_i = 0. \end{cases} \tag{47}$$

Selection arises when $cov(\eta, \zeta) \neq 0$. I make the standard assumption that each error term follows a standard normal distribution, $\eta \sim N(0, 1)$ and $\zeta \sim N(0, 1)$, with correlation between the residuals given by $\rho = corr(\eta, \zeta)$. Monte Carlo simulations show that when these distributional assumptions are violated, the model still performs well in many cases, as long as a valid exclusion restriction exists (Cook and Siddiqui 2020).

To estimate the parameters of interest, I maximize the well-known form of the log-likelihood function that can be derived from the model above:

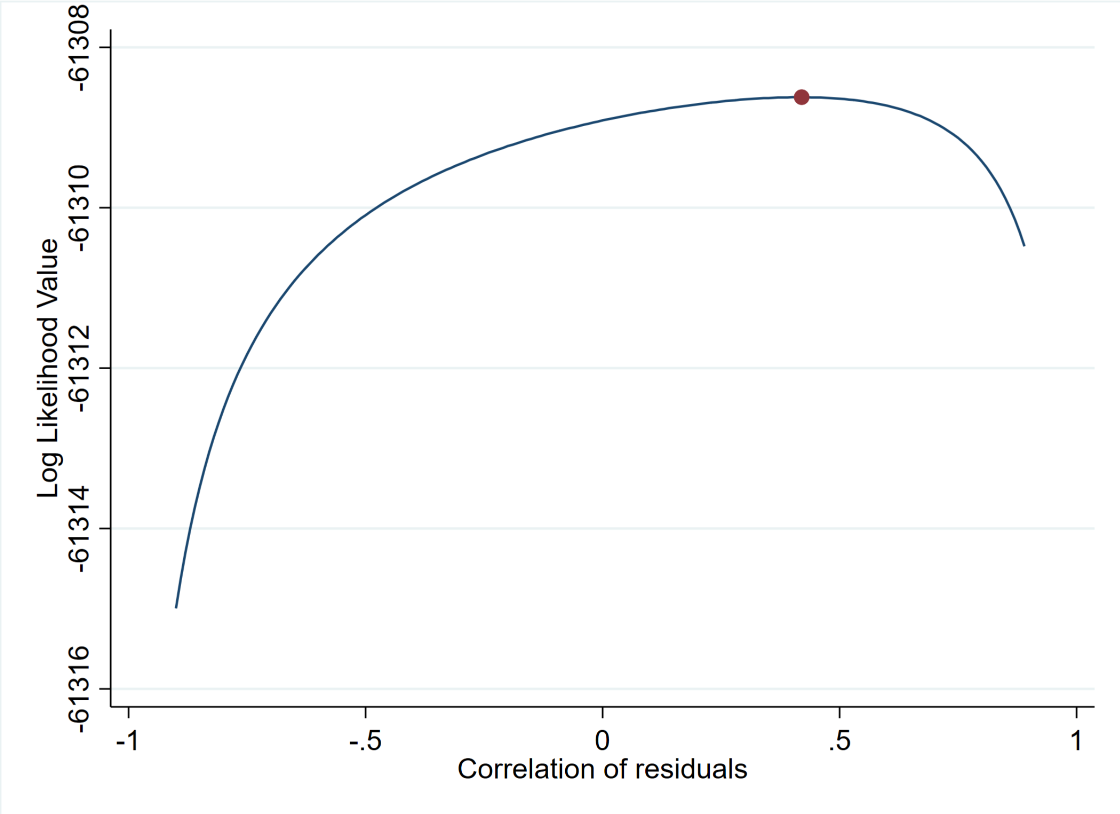
$$lnL = \sum_{i=1}^{N}\{b_i r_i ln\Phi_2(T\omega, Z\gamma, \rho) + b_i(1-r_i)ln[\Phi(T\omega) - \Phi_2(T\omega, Z\gamma; \rho)] + (1-b_i)r_i ln[\Phi(Z\gamma) - \Phi_2(T\omega, Z\gamma; \rho)] + (1-b_i)(1-r_i)ln[1 - \Phi(T\omega) - \Phi(Z\gamma) - \Phi_2(T\omega, Z\gamma; \rho)]\}, \quad (48)$$

where I denote the standard normal distribution by Φ and the joint distribution by Φ_2. If the correlation between residuals is zero, this likelihood simply equals the sum of the likelihoods of two independent probit models.

Given the structure of the model, the differences in redemption rates between experimental conditions, that is, the coefficients in ω, have a causal interpretation.

Next, I maximize the likelihood function in equation 48. To ensure I have found the global, instead of a local, maximum, I estimate the model for various *given* values of the correlation between residuals, ρ, and then compare the log likelihood values with the one when ρ is estimated. I estimate the log likelihood for given values of the correlation between the residuals using the code developed by Cook, Newberger, and Lee (2020). Figure G1 reports results by plotting the log-likelihood value for given values of the correlation. The red dot indicates the global maximum with a log-likelihood value of $-61,308$ and a correlation of residuals of around 0.4. These are exactly the values produced by the maximum likelihood estimation shown in Table 6 in the main part of the paper. This exercise confirms that the reported correlation of residuals corresponds to the global maximum.

Figure G1: Log Likelihood Values for Given Values of ρ



H Additional Figures

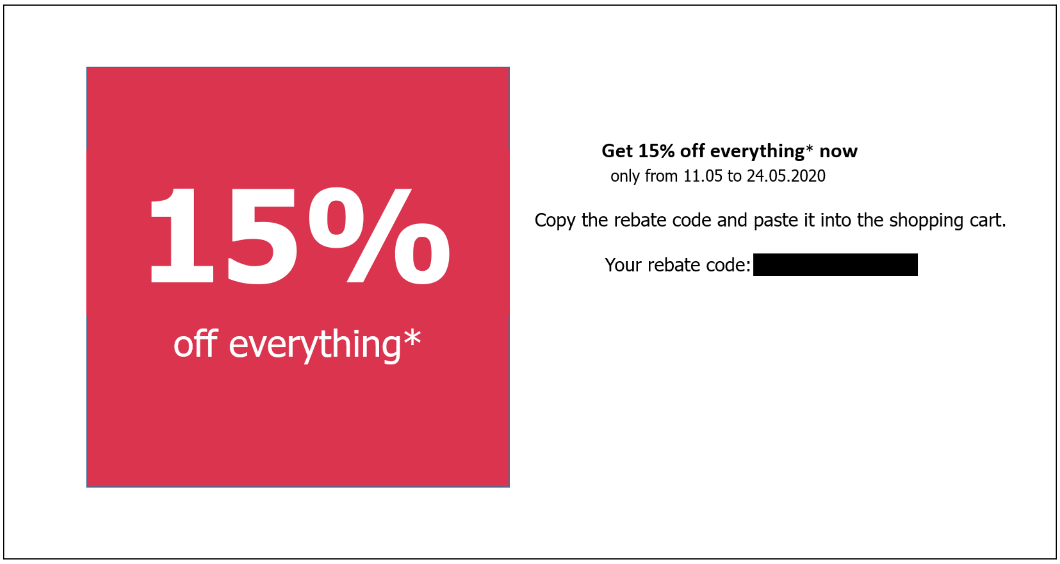
Figure H1: Banner in Groups C.1 and C.2a: 15% Rebate, Active Redemption

Note: This figure shows an English translation of the banner displayed in experimental groups C.1 and C.2a.

Figure H2: Banner in Groups C.2b: 15% Rebate, Active Redemption with Announcement of Reminder

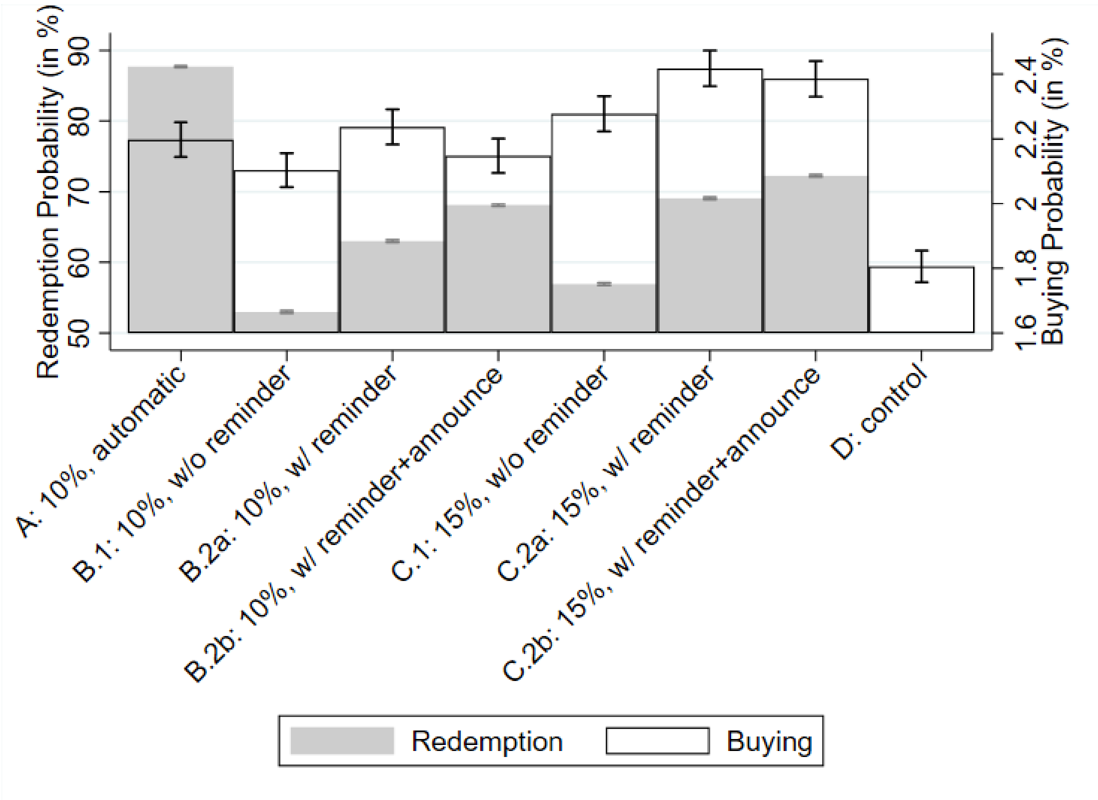
Note: This figure shows an English translation of the pop-up box that explicitly announces the reminder upon visiting the website in experimental group C2.b.

Figure H3: Subpage with Rebate Code in Group C.1, C.2a and C.2b



Note: This figure shows an English translation of the subpage showing the rebate code in experimental groups C.1, C2.a, and C2.b.

Figure H4: Buying and Redemption Probabilities: Disaggregated Treatments



Note: This figure shows the buying probability for the entire sample and the redemption probability conditional on buying. The error bars represent standard errors.