Online Appendix

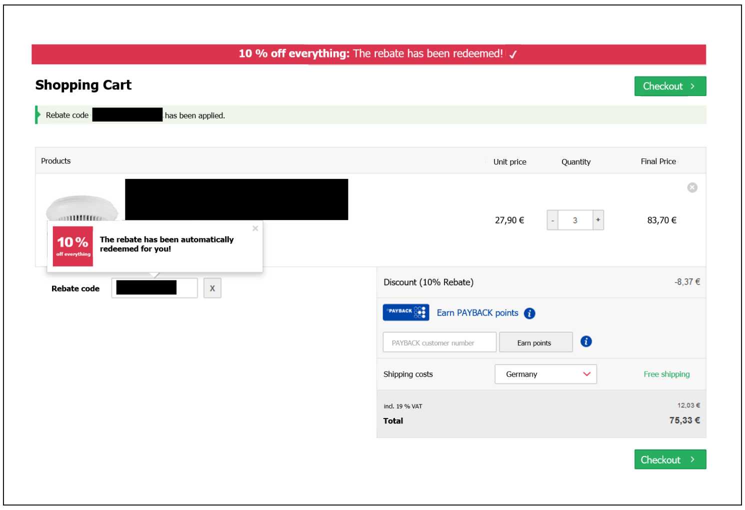
# Buy Baits and Consumer Sophistication: Field Evidence from Instant Rebates

by Matthias Rodemeier

## A Additional Tables and Figures

Figure A1: Additional Screenshots


(a) Subpage with Rebate Code in Group B.1, B.2a and B.2b




(b) Checkout Page in Group A: Automatic Redemption

**Notes:** Panel a) is the subpage showing the rebate code in experimental groups B.1, B2.a, and B2.b. Panel b) is the checkout page in group A, in which the rebate code is automatically applied to the purchase value.

Table A1: Summary Table

| Variable | A<br>10%, automatic | B.1<br>10%, w/o reminder | B.2a<br>10%, w/ reminder | B.2b<br>10%, w/ reminder<br>+ announcement | C.1<br>15%, w/o reminder | C.2a<br>15%, w/ reminder | C.2b<br>15%, w/ reminder<br>+ announcement | D<br>Control |
|---|---|---|---|---|---|---|---|---|
| Desktop user (Yes=1) | 0.351 | 0.351 | 0.351 | 0.352 | 0.350 | 0.352 | 0.349 | 0.353 |
| | (0.477) | (0.477) | (0.477) | (0.478) | (0.477) | (0.477) | (0.477) | (0.478) |
| Mobile phone user (Yes=1) | 0.563 | 0.564 | 0.562 | 0.563 | 0.564 | 0.562 | 0.566 | 0.563 |
| | (0.496) | (0.496) | (0.496) | (0.496) | (0.496) | (0.496) | (0.496) | (0.496) |
| Tablet user (Yes=1) | 0.086 | 0.086 | 0.087 | 0.085 | 0.085 | 0.086 | 0.085 | 0.085 |
| | (0.281) | (0.280) | (0.282) | (0.279) | (0.279) | (0.280) | (0.278) | (0.278) |
| Number of sessions | 1.336 | 1.338 | 1.337 | 1.337 | 1.354 | 1.344 | 1.342 | 1.332 |
| | (1.123) | (1.128) | (1.109) | (1.084) | (1.175) | (1.106) | (1.101) | (1.077) |
| Made purchase (Yes=1) | 0.022 | 0.021 | 0.022 | 0.021 | 0.023 | 0.024 | 0.024 | 0.018 |
| | (0.146) | (0.143) | (0.147) | (0.144) | (0.149) | (0.153) | (0.152) | (0.132) |
| More than one purchase (Yes=1) | 1.041 | 1.024 | 1.059 | 1.032 | 1.027 | 1.027 | 1.024 | 1.025 |
| | (0.288) | (0.157) | (0.460) | (0.213) | (0.182) | (0.176) | (0.175) | (0.179) |
| Redeemed rebate (Yes=1) | 0.880 | 0.526 | 0.630 | 0.681 | 0.570 | 0.692 | 0.725 | 0.000 |
| | (0.325) | (0.499) | (0.483) | (0.466) | (0.495) | (0.462) | (0.447) | (0.000) |
| N | 76,243 | 76,401 | 75,676 | 76,347 | 75,884 | 76,774 | 76,337 | 75,806 |

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

# B Mathematical Appendix

## B.1 Proof of Proposition 1

The probability of buying at the store can be written as

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

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

$$\hat{\theta} \approx \frac{\frac{\partial}{\partial s} B(s,\hat{\theta},\hat{c})}{\frac{\partial}{\partial s} B(s,1,\hat{c})}. \qquad (2)$$

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 $\kappa$.

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:

$$\hat{c} \approx -\frac{\Delta_{\hat{c}} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial \hat{c}} B(s, \hat{\theta}, \hat{c})} \tag{3}$$

$$= \frac{\Delta_{\hat{c}} B(s, \hat{\theta}, \hat{c})}{\frac{\partial}{\partial s} B(s, \hat{\theta}, \hat{c})}. \tag{4}$$

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

$$R(s, \theta, c) = \theta \int^{s-c} dH(\kappa). \tag{5}$$

It immediately follows that

$$\theta = \frac{R(s, \theta, c)}{R(s, 1, c)}. \tag{6}$$

An alternative way to identify $\theta$ 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

$$\theta = \frac{\frac{\partial R(s,\theta,c)}{\partial s}}{\frac{\partial R(s,1,c)}{\partial s}}. \tag{7}$$

Hassle costs can be approximated to first order by

$$c \approx -\frac{\Delta_c R(s,\theta,c)}{\frac{\partial R(s,\theta,c)}{\partial c}} \tag{8}$$

$$= \frac{\Delta_c R(s,\theta,c)}{\frac{\partial R(s,\theta,c)}{\partial s}}, \tag{9}$$

where I have used the fact that

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

Recall that consumers are sophisticated if and only if $\hat{\theta} = \theta$ and $\hat{c} = c$. Comparing equation 2 with equation 6, and equation 4 with equation 9, implies consumers are sophisticated if and only if

$$\frac{\frac{\partial}{\partial s}B(s,\hat{\theta},\hat{c})}{\frac{\partial}{\partial s}B(s,1,\hat{c})} \approx \frac{R(s,\theta,c)}{R(s,1,c)} \tag{10}$$

and

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

This completes the proof. □

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

An intuitive approach to identifying the subjective redemption probability might be comparing the demand response to a rebate with the demand response to 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 a 1 USD reduction in price. In my empirical setting, we could then simply compare the demand response to the 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:

$$R(s, \hat{\theta}, \hat{c}) = \frac{\frac{\partial B(s,\hat{\theta},\hat{c})}{\partial s}}{\frac{\partial B(s,1,0)}{\partial s}}. \tag{12}$$

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 $R(s, \hat{\theta}, \hat{c})$:

$$\frac{\partial B(s, 1, 0)}{\partial s} \times R(s, \hat{\theta}, \hat{c}) = \frac{\partial B(s, \hat{\theta}, \hat{c})}{\partial s} \tag{13}$$

$$\iff \int^{s} f(s|\kappa)h(\kappa)d\kappa \times \hat{\theta} \int^{s-\hat{c}} h(\kappa)d\kappa = \hat{\theta} \int^{s-\hat{c}} f(s - \hat{c}|\kappa)h(\kappa)d\kappa \tag{14}$$

$$\iff \underbrace{f(s|\kappa \leq s)Pr(\kappa \leq s)}_{\substack{\text{marginal consumers intending} \\ \text{to redeem in absence of frictions}}} R(s, \hat{\theta}, \hat{c}) = \underbrace{\hat{\theta} f(s - \hat{c}|\kappa \leq s - \hat{c})Pr(\kappa \leq s - \hat{c})}_{\substack{\text{marginal consumers intending} \\ \text{to redeem in presence of frictions}}} \tag{15}$$

Thus, this equality only holds under a special distributional property, which depends both on $\epsilon$ and $\kappa$. 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 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 15 holds. The identification strategy in Proposition 1 does not require this additional assumption and is, therefore, more general.

## B.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, when they are independent of the taste parameters, perceived and true inattention are still identified by the same aggregate demand elasticities in Proposition 1. By contrast, hassle costs are only identified if they are roughly homogeneous.

It is important to highlight that heterogeneity only affects the identification of the structural parameters, not the reduced-form test of sophistication explained 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, $\kappa$ and $\epsilon$. $B(\hat{\theta})$ and $R(\theta)$ are now the buying and redemption probability for a given realization of $\hat{\theta}$ and $\theta$, respectively.

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

$$\mathbb{E}\left[\frac{\partial B(s,\hat{\theta},\hat{c})}{\partial s}\right] = \int \frac{\partial}{\partial s} B(s,\hat{\theta},\hat{c}) dL_{\hat{\theta}}(\hat{\theta}). \tag{16}$$

Using the same derivation to arrive at equation 2, it follows that the expectation of perceived inattention can be identified by aggregate demand elasticities:

$$\mathbb{E}[\hat{\theta}] \approx \frac{\mathbb{E}\left[\frac{\partial}{\partial s} B(s,\hat{\theta},\hat{c})\right]}{\frac{\partial}{\partial s} B(s,1,\hat{c})}. \tag{17}$$

Similarly, using equation 6, it immediately follows that the expectation of true inattention is identified by aggregate redemption probabilities:

$$\mathbb{E}\left[\theta\right] = \frac{\mathbb{E}[R(s,\theta,c)]}{R(s,1,c)}. \tag{18}$$

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

$$\mathbb{E}\left[\Delta_{\hat{c}} B(s, \hat{\theta}, \hat{c})\right] \approx \int \Delta \hat{c} \frac{\partial}{\partial \hat{c}} B(s, \hat{\theta}, \hat{c}) dP_{\hat{c}}(\hat{c}), \tag{19}$$

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 4. An analogous argument can be made when true hassle costs are heterogeneous by taking the expectation of both sides of equation 9: 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 of heterogeneity, but hassle costs are only identified structurally if they are approximately homogeneous. These results hold as long as redemption frictions are independent of the idiosyncratic taste parameters, i.e., of preferences.

## B.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. Unlike a lump-sum rebate, an ad valorem rebate changes the optimal consumption vector because it effectively changes the price of each good. Therefore, we 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

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

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

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

For convenience, let $w_1(t, \hat{\theta}, \hat{c}, \kappa) = \hat{\theta}\,\{(v(\mathbf{x_1}) - \mathbf{p}'\mathbf{x_1} + t\mathbf{p}'\mathbf{x_1} - \hat{c})\} + \left(1 - \hat{\theta}\right)\{v(\mathbf{x_0}) - \mathbf{p}'\mathbf{x_0} + \kappa\}$ 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

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

Note that

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

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,

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

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

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

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

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

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 $\Delta c = -c$.

## B.5 GMM Estimation

Denote the buying probability in the control group by $\beta_D$. Let the treatment effect on the buying probability by treatment $t \in \{A, B.1, B.2, C.1, C.2\}$ be denoted by $\beta_t$. Perceived inattention is approximated by comparing demand responses to a rebate with and without a reminder. This is can be identified by $\hat{\theta} \approx \frac{\beta_{C.1}-\beta_{B.1}}{\beta_{C.2}-\beta_{B.2}}$. In addition, we may also compare how a rebate with and without a reminder increases demand relative to control. With two rebate values, this yields two additional moments, i.e., $\hat{\theta} \approx \frac{\beta_{B.1}}{\beta_{B.2}}$ and $\hat{\theta} \approx \frac{\beta_{C.1}}{\beta_{C.2}}$.

Perceived hassle costs are approximated by

$$\hat{c} \approx \frac{\beta_A - \beta_{B.2}}{\frac{\partial}{\partial s}B(s,1,\hat{c})}. \tag{20}$$

A linear approximation of the demand derivative is given by $\frac{\partial}{\partial s}B(s,1,\hat{c}) \approx \beta_{B.2}$. Assuming linearity, we can thus re-write $\beta_A$ in terms of structural parameters:

$$\beta_A = \hat{c}\frac{\partial B(s,1,\hat{c})}{\partial s} + \beta_{B_2} = (\hat{c} + \Delta s_1)\frac{\partial B(s,1,\hat{c})}{\partial s}. \tag{21}$$

These reformulation results in the following moment conditions for demand:

$$\mathbb{E}\left[\mathbf{I}_i\left(Buy_i-\beta_D-(\hat{c}+\Delta s_1)\frac{\partial B(s,1,\hat{c})}{\partial s}\times A_i-\frac{\partial B(s,1,\hat{c})}{\partial s}\Delta s_1\hat{\theta}\times B.1_i-\frac{\partial B(s,1,\hat{c})}{\partial s}\Delta s_1 B.2_i\right.\right.$$
$$\left.\left.+\frac{\partial B(s,1,\hat{c})}{\partial s}\Delta s_2\hat{\theta}\times C.1_i-\frac{\partial B(s,1,\hat{c})}{\partial s}\Delta s_2 C.2_i\right)\right]=0, \quad (22)$$

where $\mathbf{I}_i = (A_i, B.1_i, B.2_i, C.1_i, C.2_i, D_i)$ is the $6 \times 1$ vector of instruments indicating the experimental group of subject $i$. The monetary changes in the rebate value are $\Delta s_1 = 9.60$EUR and $\Delta s_2 = 14.40$EUR. They are obtained by multiplying the rebate values by the median shopping basket value in group A, which is 96 EUR.[1] Since four parameters need to be estimated, the model is over-identified.

Next, I derive the identification of true redemption frictions. True hassle costs are identified by

$$c = \frac{\Delta_c R(s,1,c)}{\frac{\partial R(s,1,c)}{\partial s}} = \frac{-\tau_{B2}}{\frac{\partial R(s,1,c)}{\partial s}} \quad (23)$$
$$\Leftrightarrow \tau_{B2} = -c\frac{\partial R(s,1,c)}{\partial s}. \quad (24)$$
$$(25)$$

We can insert this into the expression for $\tau_{C2}$:

$$\tau_{C2} = \tau_{B2} + \Delta s\frac{\partial R(1,c,s)}{\partial s}$$
$$= \frac{\partial R(s,1,c)}{\partial s}(\Delta s - c)$$

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

[1] See Appendix B.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.

$$\theta = \frac{R(s,\theta,c)}{R(s,1,c)} \approx \frac{\tau_{B1}+\tau_A}{\tau_A+\tau_{B2}}$$
$$\Leftrightarrow \tau_{B1} \approx \theta(\tau_A+\tau_{B2}) - \tau_A$$
$$= \theta(\tau_A - c\frac{\partial R(s,1,c)}{\partial s}) - \tau_A.$$

A 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 $\tau_{C1}$:

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

where in the last line, I have substituted for $\tau_{B1}$.

These reforumlations yield the following five moment conditions:

$$\mathbb{E}\Bigg[\mathbf{J}_i\Bigg(Redeem_i - \tau_A - \left(\theta(\tau_A - c\frac{\partial R(s,1,c)}{\partial s}) - \tau_A\right) \times B.1_i + c\frac{\partial R(s,1,c)}{\partial s} \times B.2_i$$
$$-\left(\theta\left[\tau_A + \frac{\partial R(s,1,c)}{\partial s}(\Delta s - c)\right] - \tau_A\right) \times C.1_i - \left(\frac{\partial R(s,1,c)}{\partial s}(\Delta s - c)\right) \times C.2_i\Bigg)\Bigg] = 0, \quad (26)$$

with $\mathbf{J}_i = (A, B.1, B.2, C.1, C.2_i)$ denoting the vector of instruments, excluding the control group. Here, $\Delta s = 4.80$EUR is the difference between the 10%- and 15%-rebate. With five moments and four parameters, the model is overidentified and I use a two-step GMM estimator to find the optimal weight matrix.

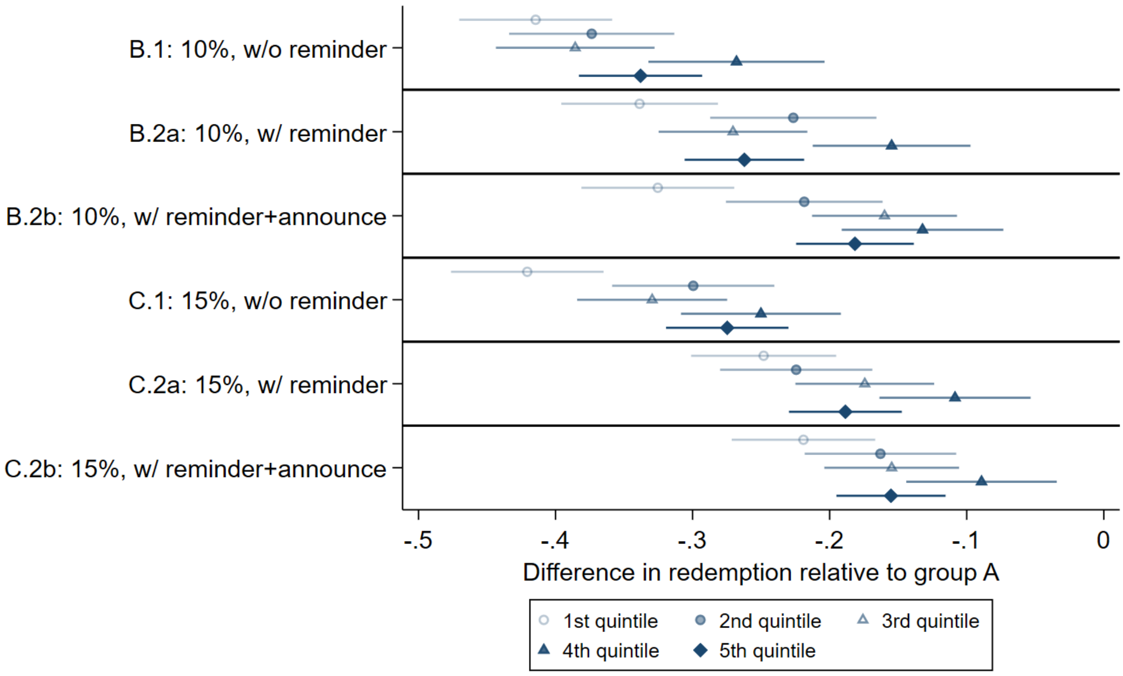
# C Heterogeneity Across Basket Value Distribution

In this Section, I study how redemption behavior varies with basket value. Since the rebates used in the experiment are ad valorem, how much a consumer receives as a discount depends on the total value of the purchased products. For example, redeeming a 10% rebate translates to a 10 EUR reduction in costs for consumers who buy 100 EUR worth of goods but only to 1 EUR for those whose basket value is 10 EUR. It is, therefore, interesting to know whether these two consumer types would respond differently to the rebate, as they receive largely different benefits from it.

In Figure C1, I plot differences in redemption rates between treatment groups across quintiles of the spending distribution. It should be noted that the results are to be interpreted with caution, however, because a promotion may simultaneously affect the value of the shopping basket and the probability of redemption. In this case, differences in redemption behavior associated with value of the basket may not have a causal interpretation. However, causality does not require that the promotions do not affect spending. It only requires that the relative position of each consumer in the spending distribution is the same between treatment and control.

With these limitations in mind, Figure C1 reveals interesting differences in redemption between low- and high-spending consumers that are consistent with a causal interpretation. In particular, redemption rates are roughly increasing in basket value. For the standard rebate (B.1), redemption is 41 percentage points lower compared to group A in the first quintile. The difference drops to around 38 percentage points for the second and third quintiles and even to 27 percentage points for the fourth quintile. For the last quintile, this pattern reverses as the redemption difference increases again to 34 percentage points. The same trend can be broadly observed for the other treatment groups: redemption increases in basket value but drops at the right tail of the value distribution. One potential explanation is that wealth increases disproportionally more than basket value as we move from the fourth to the fifth quintile. In this case, the benefits of a larger rebate value due to higher basket value are partially offset by

Figure C1: Redemption Probabilities across Basket Value Distribution




Note: This figure shows differences in redemption rates relative to group A (automatic redemption) across treatments for each quintile of the spending distribution. Darker shades of blue indicate higher spending. The error bars indicate 90% confidence intervals.

higher wealth. Redemption rates may then reverse as larger wealth decreases the incentive to pay attention to the rebate and go through the hassle of redemption. This interpretation is consistent with the German wealth distribution, where the top decile owns around 60% of total wealth.[2]

Differences across treatment groups suggest a similar pattern as our main results in Figure 4. Redemption rates increase with the reminder and the announcement, as well as with the rebate size.

[2]See Deutsche Bundesbank (2023), https://www.wsi.de/en/how-is-wealth-distributed-in-germany-14401.htm.

# D Browsing Behavior

To further understand consumer behavior, I obtain detailed firm data that tracks the browsing behavior of website visitors.[3] Column 1 of Table D1 reports the probability that a consumer starts browsing on the website, in percent relative to control.[4] If the firm offers a standard rebate, the probability to start browsing falls by 1.2% ($p < 0.05$) relative to control for the 10%-promotion. This negative effect of the rebate is almost fully eliminated when the firm offers a reminder. The fact that the reminder has a positive effect during the first minute of the website visit illustrates that some subjects make it to the checkout page quickly and see the reminder. If the reminder is immediately announced, the probability to stay on the website for more than 1 minute increases further by 1.08+1.96-1.22= 1.82% relative to control. Increasing the rebate value to 15% mitigates the negative effect of the standard rebate on the probability to start browsing. The announced reminder increases the probability to start browsing by 1.9% ($p < 0.01$). This provides evidence that some consumers anticipate their inattention at the very outset and select into the shop when their inattention is exogenously reduced by a reminder.

Column 2 shows differences in the probability to visit the checkout page at least once. The probability is 15% ($p < 0.01$) higher for the automatically-applied discount relative to control. For the equivalent rebate it is slightly lower, around 13.6%, and does not increase with the unannounced reminder. The latter result is to be expected since there is no difference between group B.1 and B.2a before they visit the checkout page. If the rebate is announced at the outset, the probability to visit checkout increases by 5% ($p > 0.1$) and 9% ($p < 0.01$) for the 10%- and 15% rebates, respectively.

In column 3, I report the number of subsequent visits on the checkout page *conditional on having visited the checkout page at least once*.[5] Importantly, the announcement of the reminder has a negative sign, both for the 10%- and 15%-rebate, and is statistically significant for the latter ($p < 0.05$). Subsequent visits are still higher with the announced reminder than in the standard rebate group (by 6.7%-1.9%=4.8% and 8.6%-4.4%=4.2%, respectively), but the increase is lower than if the reminder comes as a surprise during checkout. Thus, subjects that

[3]In the data, I observe for each link that a subject clicked on the following: a time stamp that includes the date, the hour, and the minute but not seconds. The time a subject spends on a link is calculated as the difference in minutes between the moment she clicks on that link and the moment she clicks on the next link.

[4]The outcome variable is a dummy equal to one if the subject either stays on the website for more than 1 minute or clicks on more than one weblink.

[5]These differences are not necessarily causal and should be understood as suggestive evidence to explore mechanisms underlying the main results.

receive the announcement (B.2b) already anticipate the reminder when visiting checkout, such that its impact is attenuated. Conversely, subjects that are not informed at the outset (B.2a) learn about the reminder during the first visit on the checkout page and, as a result, become more likely to continue shopping and redeem the promotion.[6] This implies that although the first point of contact with the reminder is different between B.2a and B.2b, in both groups marginal consumers see the reminder and anticipate that they will be attentive during the final checkout when they click on the "buy now"-button. This explains why the announcement has no incremental positive effect on the buying probability (recall Table 1).

Finally, column 4 examines how the browsing time differs among consumers that managed to redeem the rebate, relative to group A. Redeemers spend 13% more time to finish their purchase if they need to redeem themselves ($p < 0.01$). Redeemers who receive a reminder or announcement do not spend significantly less time to finish their purchase. This provides suggestive evidence that, as intended, the reminder did not change hassle costs but only inattention.

In sum, the browsing data provides evidence that i) consumers anticipate their inattention and are more likely to start browsing if a reminder is offered, ii) consumers who do not receive the announcement learn about the reminder during the first visit to the shopping basket, iii) the reminder changes inattention but leaves hassle costs unaffected.

[6] Note further that subsequent checkout visits are positively associated with B.1 but not A because in B.1 subjects need to move back and forth between checkout and the rebate page in order to redeem the rebate. If an unannounced reminder is shown at checkout, subsequent checkout visits increase further by 6.7%, highlighting that otherwise inattentive subjects now also go back to the rebate page in order to redeem the promotion. The same behavior is replicated for the 15%-promotion.

Table D1: Browsing Behavior

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Pr(Start browsing) | Pr(visit checkout) | # of subsequent checkout visits after 1st checkout visit | Duration among redeemers |
| A: 10%, discount | 0.8368 | 14.8971 | 0.1286 | |
| | (0.5349) | (2.3639) | (1.9602) | |
| B: 10%, rebate | -1.2174 | 13.5619 | 5.2779 | 12.8482 |
| | (0.5333) | (2.3535) | (1.9667) | (4.4816) |
| × reminder | 1.0842 | 0.7333 | 6.6844 | -3.3829 |
| | (0.5335) | (2.4292) | (2.0331) | (4.8465) |
| × announcement | 1.9615 | 4.5966 | -1.9000 | -3.7628 |
| | (0.5352) | (2.4580) | (2.0800) | (4.3904) |
| C: 15%, rebate | -0.0892 | 27.3394 | 8.6167 | 13.1690 |
| | (0.5335) | (2.4225) | (1.9875) | (4.1556) |
| × reminder | 0.3915 | -2.0683 | 8.6122 | 1.5184 |
| | (0.5309) | (2.5360) | (2.0164) | (4.4580) |
| × announcement | 1.8849 | 9.0972 | -4.3671 | -5.7717 |
| | (0.5317) | (2.5694) | (2.0264) | (4.2163) |
| N | 814,730 | 814,730 | 35,257 | 7,862 |

Note: This table reports different measures of browsing behavior across treatments. All coefficients are in percent relative to control. Column 1 reports the probability to start browsing, which is defined as either being on the website for more than 1 minute or visiting more than one page on the website. Column 2 reports the probability to visit the checkout page, which is also the shopping basket. The outcome variable in column 3 is the number of subsequent visits on the checkout page conditional on having visited checkout at least once. Column 4 shows differences in the time spent to finish the purchase among buyers that redeemed the rebate.

# 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, 1.2% 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 probability of buying again is 0.47 percentage points lower for a rebate than for an equivalent discount. Interestingly, the coefficient for the reminder is positive and almost as large in absolute size as the standard rebate. This could suggest that the negative effect of the rebate on customer loyalty is partially offset if the firm offers a reminder. The result would be consistent with the interpretation that consumers are aware of their inattention but remain mostly 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.

Table E1: Probability to Buy More Than Once

| | (1)<br>Probability to buy more than once (in %) |
|---|---|
| B: 10%, rebate | -0.4728 |
| | (0.3436) |
| × reminder | 0.3348 |
| | (0.3299) |
| × announcement | 0.0406 |
| | (0.3631) |
| C: 15%, rebate | -0.2850 |
| | (0.3547) |
| × reminder | 0.2792 |
| | (0.3439) |
| × announcement | -0.1554 |
| | (0.3498) |
| D: control | -0.1170 |
| | (0.3909) |
| Constant (A: 10%, automatic) | 1.1816*** |
| | (0.2696) |
| N | 12,895 |

Note: This 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.

# F Sample Selection Model

A large literature in econometrics has developed techniques to address bias resulting from sample selection building on the influential 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.

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.

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).

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

$$u_i = \gamma \mathbf{Z}_i + \iota \mathbf{X}_i + \eta_i, \tag{27}$$

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 $\eta_i$. 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 the region from which the subject is visiting. 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

$$v_i = \omega \mathbf{T}_i + \chi \mathbf{X}_i + \zeta_i, \tag{28}$$

where $\zeta_i$ 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:

$$r_i = \begin{cases} 1 \text{ if } v_i > 0 \text{ and } b_i = 1 \\ 0 \text{ if } v_i \leq 0 \text{ and } b_i = 1 \\ 0 \text{ if } b_i = 0. \end{cases} \tag{29}$$

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:

$$lnL = \sum_{i=1}^{N} \{b_i r_i ln\Phi_2(T\omega, Z\gamma, \rho) + b_i(1-r_i)ln[\Phi(T\omega) - \Phi_2(T\omega, Z\gamma; \rho)] + (1-b_i)r_i ln[\Phi(Z\gamma) \\ - \Phi_2(T\omega, Z\gamma; \rho)] + (1-b_i)(1-r_i)ln[1 - \Phi(T\omega) - \Phi(Z\gamma) - \Phi_2(T\omega, Z\gamma; \rho)]\}, \tag{30}$$

where I denote the standard normal distribution by $\Phi$ and the joint distribution by $\Phi_2$. In the estimation, I maximize the likelihood function in equation 30.[7] If the correlation between residuals is zero, this likelihood simply equals the sum of the likelihoods of two independent probit models.

[7] 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, $\rho$, and then compare the log-likelihood values with the one when $\rho$ is estimated. This exercise confirms that the global maximum has a log-likelihood value of $-63, 854$ and a correlation of residuals of around 0.4.

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

Table F1: Effect of Internet Outages on Buying Probability

| | (1)<br>Buying Probability $\times$100 | (2)<br>Buying Probability $\times$100 |
|---|---|---|
| Internet outage | -0.1734*** | -0.1738*** |
| | (0.0513) | (0.0513) |
| A: 10%, discount | | 0.3970*** |
| | | (0.0736) |
| B: 10%, rebate | | 0.3054*** |
| | | (0.0721) |
| $\times$ reminder | | 0.1364* |
| | | (0.0774) |
| $\times$ announcement | | -0.1056 |
| | | (0.0779) |
| C: 15%, rebate | | 0.4847*** |
| | | (0.0740) |
| $\times$ reminder | | 0.1276 |
| | | (0.0791) |
| $\times$ announcement | | -0.0248 |
| | | (0.0798) |
| Constant | 2.1990*** | 1.8031*** |
| | (0.0211) | (0.0499) |
| N | 609,468 | 609,468 |

Note: This 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.

Before estimating the selection model, I first analyze whether internet outages have a significant effect on the buying probability. Table F1 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 and highly statistically significant decrease in the buying probability by 7.9%, or 0.17 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 F2 reports the main estimation results. The log-likelihood value is –63,854 and corresponds to the global maximum, as I show in the appendix. The correlation between residuals is 0.41, 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 of 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 2 in the main text 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 as in 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.

Table F2: Estimation Results from Sample Selection Model

| | Redemption Probability (in %) |
|---|---|
| B.1: 10%, w/o reminder | -36.4649<br>(13.0971) |
| B.2a: 10%, w/ reminder | -29.0568<br>(8.7967) |
| B.2b: 10%, w/ reminder+announce | -24.9844<br>(6.6225) |
| C.1: 15%, w/o reminder | -33.2969<br>(11.3087) |
| C.2a: 15%, w/ reminder | -23.4082<br>(6.3906) |
| C.2b: 15%, w/ reminder+announce | -20.3822<br>(5.1640) |
| $\rho$ | 0.4122<br>(0.2583) |
| Fisher's Z-transformation | 0.4383<br>(0.3111) |
| Log likelihood | -63,854.278 |
| N | 533,662 |

Note: This table reports estimation results from the sample selection model in equation 30. 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 $\rho$. Fisher's Z transformation is the inverse hyperbolic tangent of $\rho$ and asymptotically normally distributed. Standard errors are in parentheses.

# G Endogenous outside option

This Appendix explains why the identification strategy in the main part of the paper is robust to the possibility that consumers have endogenous outside options, e.g., a gift card. The crucial distinction is that i) the *choice* of the outside option is endogenous while ii) the *value* of the outside option is orthogonal to the treatments (by random assignment). As explained in the main text, the identification strategy of the experimental design only requires ii) and is robust to i). As long as the distribution of gift cards is identical across the two groups, the effect of removing hassle from redeeming the rebate is identified.

Figure G1 provides a graphical illustration of this point by plotting a hypothetical distribution of $\kappa$. For all values of $\kappa$ that are below $s - c$, consumers in group B.2 redeem the rebate. The redemption probability is equal to the blue-shaded area. If we eliminate hassle costs, consumers with a gift card value above $s - c$ but below $s$ now decide to redeem the rebate instead of the gift card. This increases the redemption rate by $\Delta_c R$, which is the area shaded in red. The remaining consumers with a gift card value above $s$ continue to use the gift card, such that the redemption rate stays below 1. The change in rebate value that induces the redemption rate in B.2 to increase by the red area is equal to $-c$.

We also know how consumers would respond to an increase in rebate value, i.e. $\frac{\partial R(s,1,c)}{\partial s}$, from comparing B.1 with C.1 (or B.2 with C.2). For another example, consider a consumer that comes to the shop with a gift card of 13% and is randomized into group B.1 in which she is offered a 10% rebate. Assume further that she is fully attentive and faces no hassle costs from finding and typing in the rebate code. Then, her only "hassle costs" of redeeming the rebate is the value of not redeeming the gift card, i.e. 13% of basket value. She then chooses to redeem the gift card as 13%>10%. If, instead, she is randomized into group C.1, she decides to redeem the rebate because the value is now 15% and, therefore, higher than the gift card. Thus, this variation identifies the redemption elasticity with respect to rebate value, given some outside option in the form of a 13% gift card. In the model, this is $\frac{\partial R(s,\theta,c)}{\partial s}$. The fact that consumers endogenously switch from using a gift card to using the rebate is not a threat to identification of the redemption elasticity. Whether the consumer chooses the outside option is endogenous, but whether she is in group B.1 or C.1 is exogenous, implying that the distribution of gift cards is the same in group B.1 and C.1.

These examples illustrate that the experimental design identifies both the effect of removing hassle and the effect of increasing the rebate value. Together, this identifies our estimate of money-metric hassle costs $c = \Delta_c R(s, 1, c)/\frac{\partial R(s,1,c)}{\partial s}$.

These arguments are also illustrated more formally in the model. Proposition 1 is derived without any restrictions on the distribution of $\kappa$ other than smoothness. Consequently, Proposition 1 identifies redemption frictions even if consumers have gift cards of different sizes and switch endogenously from the gift card to the rebate for sufficiently low rebate values.

To study the role of the endogenous outside option further, Table G1 provides the analysis excluding subjects that did not redeem the rebate in group A. Relative to the main results in Table 2 of the paper, hassle costs appear larger as the difference between A and B.1 is now 47% instead of 35%. This result is mechanical because I dropped subjects from group A without dropping their statistical twins from group B.1. A result that is *not* mechanical

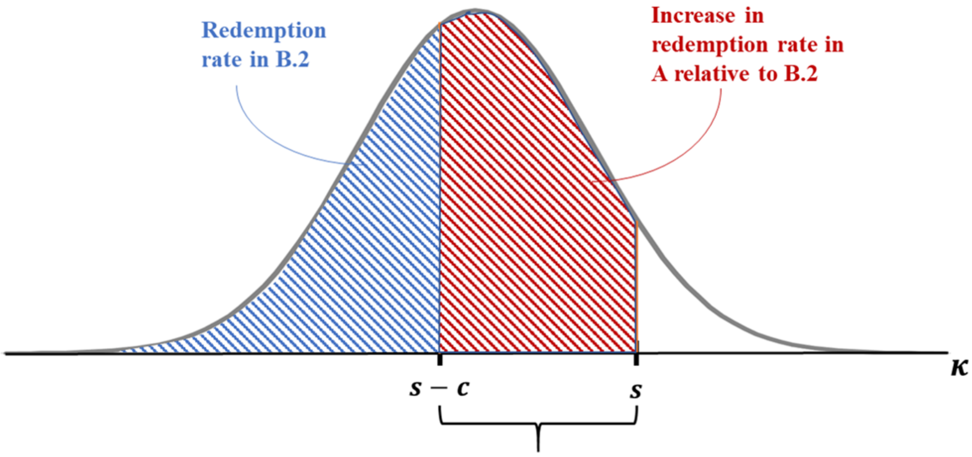


Figure G1: Distribution of Outside Option

but very reassuring is that all other treatment effects are identical to the main results: i) the effect of the reminder, ii) the effect of the announcement, and importantly iii) the effect of increasing the stakes from 10% to 15%. In Table G2, I show how these results translate into structural estimates: inattention is unchanged at 80%, $\frac{dR}{ds}$ is unchanged at $0.011$, and hassle costs increase (again mechanically) from 20 to 30 EUR.

Table G1: Redemption Rates Excluding Non-Redeemers in A

| | (1)<br>Redemption Probability $\times 100$ |
|---|---|
| $\times$ reminder | 10.3579 |
| | (1.7677) |
| $\times$ announcement | 5.1248 |
| | (1.7064) |
| C: 15%, rebate | 4.3895 |
| | (1.7483) |
| $\times$ reminder | 12.1582 |
| | (1.6325) |
| $\times$ announcement | 3.3309 |
| | (1.5315) |
| Constant (B: 10%, rebate) | 52.6117 |
| | (1.2623) |
| N | 10,226 |

Table G2: Structural Estimates Excluding Non-Redeemers in A

| | Redemption Decision |
|---|---|
| *Inattention and Hassle Costs:* | |
| $\theta$ | 0.8043 |
| | ( 0.0130) |
| $c$ (in EUR) | 30.9225 |
| | ( 5.0932) |
| *Other Parameters:* | |
| $\frac{dR}{ds} \times 100$ | 1.1168 |
| | ( 0.1949) |
| N | 11,683 |

# References


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

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.

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

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.