CESifo WORKING PAPERS

6761
2017
November 2017

Deadlines and Cognitive Limitations

Steffen Altmann, Christian Traxler, Philipp Weinschenk

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6761*
Category 13: Behavioural Economics

Deadlines and Cognitive Limitations

Abstract

This paper studies the interplay between deadlines and cognitive limitations. We analyze an agent's decision to complete a one-off task under a deadline. Postponing the task can be beneficial for the agent; missing the deadline, however, leads to a drop in the agent's rewards. If the agent exhibits cognitive limitations, postponing increases the risk of becoming inattentive and failing to complete the task in time. Our framework provides a rich set of predictions on the behavioral implications of deadlines. We test these predictions in a field experiment at a dental clinic, in which we exogenously vary deadlines and rewards for arranging check-up appointments. The empirical results underline the behavioral relevance of cognitive limitations. Imposing relatively tight deadlines induces patients to act earlier and at a persistently higher frequency than without a deadline. Evidence from a follow-up experiment and complementary surveys supports the notion that deadlines may serve as a powerful instrument when individuals' cognitive capacity is limited.

JEL-Codes: C930, D030, D910.

Keywords: deadlines, cognitive limitations, limited memory, field experiment.

Steffen Altmann
University of Copenhagen
Copenhagen / Denmark
steffen.altmann@econ.ku.dk

Christian Traxler
Hertie School of Governance
Germany - 10117 Berlin
traxler@hertie-school.org

Philipp Weinschenk
University of Kaiserslautern
Kaiserslautern / Germany
weinschenk@wiwi.uni-kl.de

November 1, 2017
We thank Michele Belot, Keith M. Ericson, Kyle Hyndman, David Laibson, Dmitry Taubinsky, Klaus Wälde, and Joachim Winter for valuable comments and suggestions. The paper also benefited from fruitful discussions with seminar and conference participants in Bergamo, Berlin, Copenhagen, Düsseldorf, Frankfurt, Gothenburg, Harvard, Kiel, London, Mannheim, Mainz, Marburg, Munich, Prague, Rotterdam, and San Diego. Financial support from Volkswagen Foundation and the Max Planck Institute for Collective Goods is gratefully acknowledged. We thank Ann-Kathrin Crede, Nicolas Kaufung, and Kyle Ott for excellent research assistance.

1 Introduction

Deadlines are a pervasive feature of our lives. As consumers, we are regularly offered time-limited coupons, rebates, and 'special deals'. After buying a product, our consumer rights are governed by deadlines regarding payment dates, refund and return policies, or product warranties. As academics, we face deadlines for submitting student grades, grant applications, or conference papers. We need to keep track of the check-in deadlines for our next travels, and the upcoming cutoff dates for switching our health insurance or retirement savings plan. We receive a bonus for finishing an expert report before the due date—and we ought to take care of this extra income before tax day.

From an economic perspective, deadlines impose constraints on our intertemporal choice sets. At the same time, we commonly face further constraints in terms of our cognitive and perceptual capacities (Simon 1955, Kahneman 2003). In fact, a growing body of literature indicates that individuals exhibit limitations in memory and attention (e.g., Chetty et al. 2009, Ericson 2011, Taubinsky 2014, Haushofer 2015, Ericson 2017). Such limitations have been found to affect consumers' purchasing behavior (Hossain and Morgan, 2006; Lacetera et al., 2012; Englmaier et al., 2017), responses to taxation (Chetty et al. 2009, Taubinsky and Rees-Jones 2017), or the take-up of social benefits (Bhargava and Manoli 2015). A related strand of research shows that people focus disproportionately on particularly salient aspects of financial decisions, work, or consumption choices, while paying little or no attention to less salient ones (Bordalo et al. 2012, Bordalo et al. 2013, Kőszegi and Szeidl 2013).

In this paper, we study the interplay between deadlines and cognitive limitations in intertemporal decisions. We start from a simple theoretical framework, in which a decision maker faces a deadline for completing a one-off task. Timely task completion is beneficial for the agent, who receives a reward if the task is carried out before the deadline. In any given period, however, the agent potentially has an incentive to postpone the task: as the costs for task completion vary stochastically over time, postponing entails an option value of realizing lower task completion costs at a later point in time. Without cognitive limitations, the agent's likelihood of completing the task depends positively on the length of the deadline. Specifically, longer deadlines increase the chance that the task is completed, both in terms of pre-deadline and long-run (post-deadline) task completion rates. At the same time, the conditional probability of completing the task in a given period—the hazard rate—is monotonically increasing towards the deadline. This is the case because the option value of postponing the task becomes smaller and smaller as the deadline comes closer. Notably, this option-value effect (and thus the hazard rate) depends only on how much time is left until the deadline, but not on how much time has passed since the beginning of the game.

When people exhibit limitations in memory or attention, all of these predictions may reverse. In particular, we show that task completion rates for agents with cognitive limitations may be higher under shorter deadlines than under longer ones. Intuitively, shorter deadlines increase the likelihood that the task is completed in early periods—when the agent is more likely to have it still on the top of her mind. With a longer (or no) deadline, in contrast, the option value of waiting is high and incentives to act early are weakened. Hence, the agent is more likely to postpone the task to later periods, which in turn increases the risk of becoming inattentive.

Our analysis further shows that cognitive limitations work against the option-value effect described above. Depending on the severity of cognitive limitations, pre-deadline hazard rates can thus be U-shaped or monotonically decreasing over time. Finally, hazard rates at a given point before the deadline do not only depend on the 'time left' but also (negatively) on the overall length of the deadline.

In the second and main part of the paper, we test the competing theoretical predictions in a natural field experiment at a dental clinic. The task that people face in our experiment consists of calling the clinic to arrange a preventive check-up appointment—a task that should be conducted regularly, but at a relatively low frequency, as it is commonly the case for other chores like adapting one's health insurance or retirement savings plan (see, e.g., Chetty et al., 2014a; Heiss et al., 2016). Building upon the clinic's pre-existing system of postcard reminders, our experiment introduces and randomly varies the length of deadlines for making a check-up appointment. Specifically, patients in our main experiment are eligible for receiving a reward if they contact the dentist within either one or three weeks after our intervention. For a third group of patients, we do not set any deadline. In a smaller follow-up experiment, we augment the number of treatments to test, among others, the impact of longer deadlines of six and ten weeks.

Our empirical analysis draws on roughly 3,600 patient-treatment observations collected over more than three years. To evaluate how different deadlines affect the timing and frequency at which patients arrange check-up appointments, we make use of the records of all inbound phone calls at the dental clinic during this period. In addition, our data includes further information on actual check-up attendance as well as a rich set of individual background characteristics.

Our results point out that cognitive limitations play an important role for people's responses to deadlines. In particular, the data clearly reject the prediction of increasing pre-deadline hazard rates that follows from the benchmark model without cognitive constraints. We also find no evidence for the benchmark prediction of time-invariant hazards in the no-deadline treatment. Hazard rates are U-shaped or decreasing over time, supporting the notion that individuals exhibit (intermediate levels of) cognitive limitations. Our data further show that, at a given period before the deadline, hazard rates are negatively related to the length of the deadline. This finding is again inconsistent with the benchmark model, according to which hazard rates should only depend on the time left until the deadline.

The treatment differences in hazard rates also translate into strong and systematic deadline effects in terms of task completion, i.e., the cumulative frequencies at which individuals contact the dentist for arranging an appointment. During the first week, for instance, the one-week deadline induces significantly higher completion rates than the other two treatments in our main experiment. After three weeks, completion rates are similar under the one- and three-week deadline. However, under both deadlines, the three-week completion rates are significantly higher than in the no-deadline treatment. Notably, the gap between the treatments with and without a deadline persists after the deadlines have expired: even after 100 days, the fraction of patients who have arranged a check-up appointment is around 10% higher in the deadline treatments than in the no-deadline case. Hence, imposing a deadline leads to higher task completion rates than facing no deadline.

The key results from our empirical analysis are corroborated in the follow-up experiment. In addition, we provide evidence suggesting that shorter deadlines can indeed trigger higher within-deadline task completion rates than longer deadlines. In particular, we find that the fraction of patients who arranges an appointment by the three-week deadline is higher than the fraction of patients that does so during the twice as long time window of a six-week deadline. While the difference is not statistically significant, it is consistent with the idea that shorter deadlines may not only accelerate but also boost overall task completion. In this sense, shorter deadlines can yield a 'double dividend'.

Complementary evidence from an online survey experiment and a post-experimental survey at the dental clinic further underlines the potential benefits of short deadlines. Almost all survey participants indicate that they perceive deadlines as helpful to avoid problems related to 'postponing and forgetting' and a majority of respondents generally prefer relatively tight deadlines, suggesting that they are at least partially aware of the risk of becoming inattentive.

A number of recent papers have shown theoretically that cognitive constraints (Taubinsky, 2014), present bias (O'Donoghue and Rabin, 1999; Bisin and Hyndman, 2014), or the interplay between the two (Ericson, 2017) can have important implications for the functioning of deadlines. Some of these studies also illustrate the empirical relevance of these mechanisms. Ericson (2017), for instance, shows that neglecting cognitive limitations can lead to severely biased preference estimates from dynamic choice data; Taubinsky (2014) documents that reminders about upcoming deadlines can help to increase task completion; Bisin and Hyndman (2014) find that students exhibit a demand for self-imposed deadlines, even if these do ultimately not help them to fulfill their tasks.[1]

We add to this literature in several important ways. First, we provide evidence from a large, natural field experiment that allows us to exogenously vary a rich set of deadlines and observe behavior over a relatively long time frame. Second, to the best of our knowledge, our paper is the first to document that imposing a deadline can lead to persistently higher task completion rates than an otherwise identical no-deadline environment. Third, we expand the focus from task completion to hazard rates. Similar as Heffetz et al. (2016), who study how variation in the timing of reminder messages affects the payment of parking tickets, we analyze hazard rates as an additional metric to evaluate the role of cognitive limitations. This also differentiates our model framework from others: our theoretical analysis, which is otherwise closely related to Taubinsky (2014) and Ericson (2017), offers several testable predictions on the shape of hazard rates. Together with the evidence on cumulative task completion, the empirically observed hazards provide a coherent picture on how cognitive limitations can shape people's reactions to deadlines.

Our data also allow us to examine a potential concern regarding the analysis of hazard rates in the presence of unobserved type heterogeneity (see, e.g., Salant 1977 and, for a recent discussion, Heffetz et al. 2016). With systematic heterogeneity in patients' propensity to respond, our interpretation of decreasing hazards may be misleading, as it abstracts from differential sorting of patients over time. To assess this concern, we can draw on a rich set of individual-level characteristics as well as detailed data on patients' behavior

[1] An interesting policy perspective on the optimal design of deadlines is discussed in Rees-Jones and Taubinsky (2016). The behavioral consequences of deadlines have also been examined in a number of other studies in economics (e.g., Bertrand et al., 2010; Chetty et al., 2014b; Damgaard and Gravert, 2017) and psychology (e.g., Ariely and Wertenbroch, 2002; Shu and Gneezy, 2010). In contrast to our paper and those mentioned above, these latter studies either consider deadline effects along with a range of other interventions, or do not study the micro-level mechanisms reflected in the observed aggregate outcomes.

prior to our intervention.[2] We use these data to predict which 'types' of patients are most likely to quickly respond to our intervention. Comparing hazards among subgroups with high vs. low predicted response propensities, we observe only modest differences in the level and shape of hazard rates. In particular, we do not find any evidence of monotonically increasing pre-deadline hazards. This does naturally not preclude that there may remain relevant residual heterogeneity within the subgroups. However, our evidence shows at least that response patterns under different deadlines are remarkably similar for a wide range of different groups. Moreover, the subgroup analysis also sheds some light on who 'benefits' most from deadlines in terms of increased task completion. We find that deadlines have a larger impact on task completion rates of groups with lower baseline response rates in the no-deadline condition. Put differently, deadlines seem to be particularly well suited to foster task completion among groups with lower baseline propensities to execute the task at hand.

The idea that shorter deadlines may be beneficial for task completion has also been discussed in a number of earlier theory contributions. In this literature, tight deadlines enhance incentives to start engaging in long-term projects (Saez-Marti and Sjögren, 2008), help to overcome free-riding in teams (Campbell et al., 2014; Weinschenk, 2016), or reduce procrastination by present-biased agents (O'Donoghue and Rabin, 1999; Herweg and Müller, 2011). Complementing this perspective, our results indicate that—independent of these motives—tight deadlines may have positive effects for task completion when agents are subject to cognitive limitations.

On a more general level, our findings add to a better understanding of the behavioral consequences of scarcity in economic and mental resources (Mani et al., 2013; Mullainathan and Shafir, 2013; Wälde, 2015; Carvalho et al., 2016; Schilbach et al., 2016). Some of our results also speak to the growing literature on planning prompts (surveyed in Rogers et al. 2015 and Beshears et al. 2016). Specifically, in a subset of our treatments, we observe that deadlines affect behavior even if there is no explicit reward attached to them, suggesting that they may help individuals to better plan and structure their tasks. From a practical perspective, this also suggests that deadlines can be a powerful and seemingly 'cheap' policy tool. Caution is warranted, however, as the gains from imposing a deadline in one domain might imply costs for fulfilling other tasks (compare the notion of 'focus dividend' vs. 'bandwith tax' in Mullainathan and Shafir, 2013). Assessing the nature and potential consequences of these interdependencies is an important task for future research.

The paper proceeds as follows: in the following section, we introduce our theoretical framework and derive behavioral predictions for agents with limited and unlimited cognitive capacities. Section 3 discusses the setup and procedures of our experiment, and Section 4 presents our empirical results. Section 5 concludes.

[2]These data were collected during an earlier study, in which we analyzed how patients at the same clinic respond to differently framed reminder messages (Altmann and Traxler, 2014). Consistent with other work showing that reminders affect behavior in a variety of settings (e.g., Apesteguia et al. 2013, Karlan et al. 2016, Calzolari and Nardotto 2017), we found that reminders increase check-up rates relative to a no-reminder control group, but that the framing of the messages matters little for patients' responses. Note that the evidence on reminder effects also alludes to the relevance of cognitive limitations in different applications.

2 The Model

An agent is confronted with a task that can be carried out at any time t before a deadline $T \in \mathbb{N}$.[3] The costs of completing the task in period t are $c_t \in \mathbb{R}^+$. Costs are stochastic and drawn at the beginning of each period from a continuous distribution F with support $[0, \bar{k}]$. The model is set in discrete time. In every period t, the agent learns about the realization of costs and subsequently decides whether to complete the task in this period or not. If the agent completes the task, she obtains a reward Y with a present value of y, with $0 < y < \bar{k}$. If she decides not to complete the task and there is time left, $t < T$, the agent advances to the next period. The game ends if the task is completed or the deadline is reached without task completion.

In the following, we analyze the effects of finite deadlines on the decision to complete the task. We will also explore the case without a deadline, which is isomorphic to having an infinite deadline. As a benchmark, we first consider the case of an agent who has no cognitive limitations.

2.1 Agent without Cognitive Limitations

Completing the task in period t yields a payoff of $y - c_t$,[4] while not doing so in t yields δV_{t+1}, where $\delta \in (0, 1)$ denotes the discount factor and V_{t+1} the expected discounted payoff of reaching period $t + 1$, which can be interpreted as an option value. Completing the task in period t is optimal for the agent if and only if the costs of doing so are sufficiently low:[5]

$$y - c_t \leq \delta V_{t+1} \iff c_t \leq \hat{c}_t := y - \delta V_{t+1}. \tag{1}$$

The option values are endogenous. Under the optimal decision rule from equation (1),

$$V_t = \int_0^{\hat{c}_t} (y - c) dF(c) + (1 - F(\hat{c}_t)) \, \delta V_{t+1}. \tag{2}$$

The option value of reaching period t thus equals the probability of completing the task in t times the expected payoff of doing so *plus* the probability of not completing the task in t times the discounted option value of reaching period $t + 1$.

We next determine the time structure of the option values, which in turn determines the optimal decisions. In Appendix A.1, we show that the option value is decreasing over time:

$$V_t > V_{t+1} \text{ for all } t \in \{1, ..., T\}. \tag{3}$$

[3]Extending our framework to situations where the agent can still complete the task after the deadline is straightforward. We further discuss this case in Section 3.3 and Appendix A.7.

[4]Recall that y denotes the present value of the reward. If the reward Y is received $\ell > 0$ periods after task completion, $y = \delta^\ell Y$. The literature typically considers a one-period delay. In this case, we have $y = \delta Y$ and the cost threshold in equation (1) would be $\hat{c}_t = \delta(Y - V_{t+1})$.

[5]This threshold property is reminiscent of search models in the tradition of Stigler (1961), McCall (1970), and Mortensen (1970): the agent 'stops searching'—i.e., she exerts the task in our setting—as soon as the (net) payoff from doing so is sufficiently high.

This property is intuitive: the more time is left until the deadline, the higher is the agent's expected discounted payoff from postponing the task, i.e., the option value. Furthermore, the option value V_t is bounded and approaches the bound $\bar{V}$ as the deadline goes to infinity (see Appendix A.2):

$$V_t < \bar{V} \text{ for all } T < \infty \text{ and } \lim_{T \to \infty} V_t = \bar{V}. \tag{4}$$

From these properties one can derive a first set of results.

Hazard Rate over Time. The probability that the agent completes the task in period t, given that she has not done so before, is given by

$$h_t = F(\hat{c}_t). \tag{5}$$

To illustrate the link to our later empirical analysis—and in slight misuse of terminology given our discrete-time framework—we refer to h_t as the *hazard rate of task completion* in what follows.

How does h_t evolve over time? Since the option value V_t is decreasing over time (see inequality 3) and the cost-threshold $\hat{c}_t$ is decreasing in the option value (see equation 1), the cost-threshold is increasing over time. Thus, it follows from (5) that the hazard rate is increasing over time:

$$h_t < h_{t+1} \text{ for all } t \in \{1, ..., T-1\}. \tag{6}$$

Intuitively, the closer the deadline, the higher are the costs that the agent is willing to bear to get the task done. This in turn translates into a higher hazard rate. Note that this result contrasts with the case where the agent faces no deadline for completing the task. Without a deadline, $T \to \infty$, the environment is time-invariant and the hazard rate is constant over time: $h_t = F\left(y - \delta\bar{V}\right)$.

HYPOTHESIS 1 (H1): *(i) Under a finite deadline T, the hazard rate is increasing over time until T. (ii) When the agent faces no deadline, the hazard rate is constant over time.*

Hazard Rate and Distance to the Deadline. We next explore how the option value and the hazard rate depend on how much time there is left until the deadline. Since the option value of reaching the period after the deadline is zero ($V_{T+1} = 0$), (1) and (2) imply that the option value in the period immediately before the deadline, V_T, is the same for all possible deadlines T. Repeating this argument further shows that the option values in preceding periods, $V_{T-1}, V_{T-2}, ...$, are the same for all T, too. Hence, if we vary the length of the deadline, the option value only depends on the time left until the deadline but *not* on the absolute value of time:

$$V_t|_{T=T'} = V_{t+s}|_{T=T'+s} \text{ for all } T' \in \mathbb{N} \text{ and } s \in \mathbb{N}. \tag{7}$$

Using (1), we can make the same statement for the cost threshold and, hence, for the hazard rate:

$$h_t|_{T=T'} = h_{t+s}|_{T=T'+s} \text{ for all } T' \in \mathbb{N} \text{ and } s \in \mathbb{N}. \tag{8}$$

The intuition behind this property is again simple: if, at a given point in time, the agent has not yet completed the task, her incentives for completing it in the present period depend on how much time there is left until the deadline, but not on how much time has already passed. We summarize this result in

HYPOTHESIS 2 (H2): *Under a finite deadline T, the hazard rate depends only on the time left until the deadline, but not on the absolute value of time.*

Completion Rates. The probability that the agent has completed the task until period t—the *period-t completion rate*—is given by

$$p(t) := 1 - \Pi_{s=1}^{t}(1 - h_s), \tag{9}$$

where $\Pi_{s=1}^{t}(1-h_s)$ measures the probability of not completing the task until t. The probability of completing the task before the deadline—which we refer to as *within-deadline T completion rate*—is $p(T)$. Making use of (8), it is straightforward to see that the within-deadline completion rate is increasing in the length of the deadline. Formally, for all pairs of deadlines T' and T'', with $T' < T''$, it holds that[6]

$$p(T'') > p(T'). \tag{10}$$

Intuitively, during the last T' periods before the deadline, the agent's optimal strategy is exactly the same under the short and the long deadline. For the longer deadline T'', however, the agent may have completed the task already in periods $t \leq T'' - T'$. The overall probability that the agent completes the task within the deadline is therefore higher under the long deadline T'' than under the short deadline T', and it is highest if $T \to \infty$.

HYPOTHESIS 3 (H3): *The within-deadline completion rate $p(T)$ is monotonically increasing in the length of the deadline T.*

Finally, the model also provides insights on how completion rates at any given point in time vary with the length of the deadline T. In particular, as shown in Appendix A.3, for all pairs of deadlines T' and T'', with $T' < T''$, there exists a date $\tau \in (T', T'')$ such that

$$p(t)|_{T=T'} > p(t)|_{T=T''} \text{ for all } t < \tau \text{ and } p(t)|_{T=T'} < p(t)|_{T=T''} \text{ for all } t > \tau. \tag{11}$$

For any period before the threshold τ, period-t completion rates are higher under the short deadline T' than under the long deadline T''. Vice versa, period-t completion rates under T'' are higher than those under T' for all periods after the date τ. Intuitively, while a short deadline provides stronger incentives for task

[6]Since hazard rates are positive and smaller than one, i.e., $h_t = F(\hat{c}_t) \in (0, 1)$ for all $t \in \{1, ..., T\}$ (see Appendix A.2), (8) implies that $p(T') = 1 - \Pi_{t=1}^{T'}(1 - h_t|_{T=T'}) = 1 - \Pi_{t=T''-T'+1}^{T''}(1 - h_t|_{T=T''}) < 1 - \Pi_{t=1}^{T''}(1 - h_t|_{T=T''}) = p(T'')$. Note that this also holds if $T'' \to \infty$, i.e., for the case without a deadline.

completion in the short run, the long deadline does provide incentives over a longer time horizon; completion rates therefore 'catch up' and ultimately overtake those under the short deadline.[7]

Let us emphasize that this result holds independently of whether T'' is finite or infinite. Therefore, comparing a scenario in which the agent faces a deadline (finite T') to one where she does not ($T'' \to \infty$), we expect completion rates to be higher in the finite deadline scenario for all periods before the finite deadline, whereas after some later point in time, $\tau > T'$, completion rates become higher in the no-deadline scenario.

HYPOTHESIS 4 (H4): *For all pairs of deadlines T' and T'', with $T' < T''$, the period-t completion rate $p(t)$ is higher [lower] under the deadline T' than under the deadline T'' in any period $t < \tau$ [$t > \tau$], with $T' < \tau < T''$.*

2.2 Cognitive Limitations

We now consider an agent with cognitive limitations. More specifically, the agent might exhibit limited memory (Ericson, 2017), becomes inattentive (Taubinsky, 2014), or no longer has the task 'on the top of her mind' because of other cognitive constraints (Karlan et al., 2016). We formalize such cognitive limitations in the simplest possible way: at the end of each period, the agent turns inattentive with probability $\gamma \in (0, 1]$. Once she is inattentive, she will never complete the task.[8] For the moment, we further assume that the agent is naive about her cognitive limitations. The case of a sophisticated agent is discussed in Section 2.3 below.

Being naive, the agent does not foresee the possibility that she might become inattentive. She therefore uses the same option values and cost thresholds as in the case without cognitive limitations. From the perspective of the beginning of the game, the probability that the agent completes the task in period t is then given by

$$q_t = F(\hat{c}_t)(1-\gamma)^{t-1} \prod_{s=1}^{t-1}(1 - F(\hat{c}_s)). \tag{12}$$

The probability that the agent completes the task in period t thus depends on (i) the probability of the cost draw in this period being low enough, (ii) the probability of the agent still being attentive in period t, and (iii) the probability of not having completed the task in some earlier period $s < t$. The hazard rate for the cognitively limited agent is then given by

$$\tilde{h}_t = \frac{q_t}{1 - \sum_{s=1}^{t-1} q_s}. \tag{13}$$

[7] A sufficient condition that (11) holds for situations where the agent can complete the task *after* the deadline (yielding a reward that is smaller than the one for within deadline completion) requires post-deadline hazard rates to be lower than pre-deadline hazards. In Appendix A.7 we show that this condition indeed holds.

[8] A similar assumption (on memory limitations) is made in Ericson (2017). Allowing for time-dependent probabilities γ_t or the possibility of 'recalling' the task again at some later point in time complicates notation without changing the key qualitative insights (see Section 2.3 below). Attention dynamics in settings involving repeated tasks are further analyzed in Taubinsky (2014).

Hazard Rate over Time. To determine how the hazard rate evolves over time, we consider the hazard ratio

$$\frac{\tilde{h}_{t+1}}{\tilde{h}_t} = \frac{\frac{q_{t+1}}{1-\sum_{s=1}^{t} q_s}}{\frac{q_t}{1-\sum_{s=1}^{t-1} q_s}}, \tag{14}$$

which can be rewritten as

$$\frac{\tilde{h}_{t+1}}{\tilde{h}_t} = \frac{F(\hat{c}_{t+1})}{F(\hat{c}_t)}(1-\gamma)(1-F(\hat{c}_t))\frac{1-\sum_{s=1}^{t-1} F(\hat{c}_s)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r))}{1-\sum_{s=1}^{t} F(\hat{c}_s)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r))}. \tag{15}$$

This expression illustrates two countervailing effects that shape the behavior of a cognitively limited agent. The first fraction on the right-hand side captures the *option-value effect*. Since the option value is decreasing over time (see inequality 3), the cost threshold is higher in $t+1$ than in t and hence $F(\hat{c}_{t+1}) > F(\hat{c}_t)$. This effect contributes towards increasing the hazard over time. The remaining terms on the right-hand side—which are all positive and below one—capture the *cognitive-limitation effect*, which works into the opposite direction: due to the agent's cognitive limitations, the likelihood of having the task on the top of her mind is decreasing from one period to the next. The interplay of this effect with the former determines how the hazard evolves over time.

Specifically, when the probability of becoming inattentive is sufficiently high, the cognitive-limitation effect dominates and the hazard rate is locally decreasing. Vice versa, if the probability of becoming inattentive is sufficiently low, the option-value effect dominates and the hazard rate is locally increasing. More formally, there exists a threshold $\hat{\gamma}_t \in (0,1]$ such that $\tilde{h}_{t+1} < \tilde{h}_t$ whenever $\gamma > \hat{\gamma}_t$, and $\tilde{h}_{t+1} > \tilde{h}_t$ whenever $\gamma < \hat{\gamma}_t$ (see Appendix A.4).

One can further show that the difference in option values V_t and V_{t+1} is larger, the lower is V_{t+1}, i.e., the closer period t is to the deadline T.[9] This implies that the option-value effect is relatively strong in periods close to the deadline, whereas it is weaker in early periods where the deadline is still far away. The option-value effect is thus more likely to be dominated by the cognitive-limitation effect in periods long before the deadline. During early periods of the game, hazard rates are thus more likely to be locally decreasing. Vice versa, the option-value effect is more likely to dominate in periods close to the deadline. As a result, agents with cognitive limitations may exhibit U-shaped hazard functions that are initially decreasing but then increase in later periods once the deadline comes closer.

Finally, for the case without a deadline, $T \to \infty$, the option value is constant over time and, consequently, the option-value effect plays no role. As a result, the cognitive-limitation effect always dominates and the hazard rate is monotonically decreasing over time.

ALTERNATIVE HYPOTHESIS 1 (H1'): *(i) Under a finite deadline T, the hazard rate is locally decreasing if the probability of becoming inattentive is sufficiently high, whereas it is locally increasing if γ is sufficiently low. (ii) When the agent faces no deadline, $T \to \infty$, the hazard rate is decreasing over time.*

[9] See equation (27) and the subsequent arguments in the Appendix.

Hazard Rate and Distance to the Deadline. In a next step, we analyze how the hazard rate under cognitive limitations depends on the absolute value of time as well as on the time left until the deadline. In contrast to the case without cognitive limitations, the hazard rate for a cognitively limited agent depends negatively on the absolute value of time, holding the distance to the deadline fixed. Formally, as shown in Appendix A.5,

$$\tilde{h}_t\Big|_{T=T'} > \tilde{h}_{t+s}\Big|_{T=T'+s} \quad \text{for all } \; T' \in \mathbb{N} \; \text{ and } \; s \in \mathbb{N}. \tag{16}$$

This result is due to the cognitive-limitation effect discussed above: holding the distance to the deadline fixed, the agent is more likely to having turned inattentive and, hence, less likely to complete the task at a given point in time before a late deadline, relative to the corresponding point in time before a shorter deadline.

ALTERNATIVE HYPOTHESIS 2 (H2'): *Under a finite deadline T, the hazard rate depends on the time left until the deadline and negatively on the absolute value of time.*

Completion Rates. A further interesting deviation from our benchmark results is obtained when we examine completion rates of a cognitively limited agent. Consider first the case where γ is small, i.e., $\gamma \to 0$. In this case, the hazard rates with cognitive limitations, $(\tilde{h}_1, ..., \tilde{h}_T)$, approach those without cognitive limitations, $(h_1, ..., h_T)$. As in the case with no cognitive limitations, the within-deadline completion rate $p(T)$ is then increasing in the length of the deadline. The result looks quite different, however, when γ is large, $\gamma \to 1$. In this case, the hazard rate is positive in the first period and approaches zero in all subsequent periods. The within-deadline completion rate $p(T)$ thus approaches the first-period hazard rate $F(\hat{c}_1)$. As the cost threshold $\hat{c}_1$ is decreasing in the length of the deadline T, the within-deadline completion rate $p(T)$ is decreasing in T, too. With cognitive limitations, shorter deadlines may therefore cause *higher* within-deadline completion rates.[10]

ALTERNATIVE HYPOTHESIS 3 (H3'): *The within-deadline completion rate $p(T)$ is increasing in the length of the deadline for low levels of γ, but decreasing in the deadline length for high levels of γ.*

In addition to this result—which parallels earlier findings by Taubinsky (2014) and Ericson (2017)—one can derive further results on period-t completion rates, $p(t)$. In particular, as in the framework without cognitive limitation, $p(t)$ is higher under a short deadline T' than under a long deadline T'' for any period $t \leq T' < T''$. This holds since in all these periods the cost threshold is higher under the short than under the long deadline. In contrast to the case without cognitive limitations, however, it may no longer be true that the period-t completion rate under a long deadline catches up in later periods and overtakes the one under a short deadline after some threshold period τ (where $T' < \tau < T''$). In particular, for the case where γ is high enough to yield $p(T') > p(T'')$, the completion rate under the long deadline never fully catches up, and all period-t completion rates are higher under the short deadline T' than under the long deadline T'',

[10]To illustrate this result, let $c \sim U[0,1]$, $y \in (0,1)$, $\delta = 1$, and T be 1 or 2. Then $p(T = 1) = y$ and $p(T = 2) = 1 - (1 - y + y^2/2)(1 - (1 - \gamma)y)$. Thus, $p(T = 1) > [=, <] p(T = 2)$ if $\gamma > [=, <] \frac{y^2-3y+2}{y^2-2y+2}$.

i.e., $p(t)|_{T=T'} < p(t)|_{T=T''}$ holds for any t. Hence, with sufficiently strong cognitive limitations, period-t completion rates are persistently higher with a short deadline than under a long (possibly infinite) deadline.

ALTERNATIVE HYPOTHESIS 4 (H4'): *For all pairs of deadlines T' and T'', with $T' < T''$, the period-t completion rate $p(t)$ is higher under the deadline T' than under the deadline T'' in any period $t \leq T'$. For periods $t > T'$, $p(t)$ remains persistently higher under the deadline T' than under T'' if γ is sufficiently high.*

2.3 Extensions

Our framework can be extended in a number of different directions. In what follows, we examine how an agent's sophistication about her cognitive limitations, present bias, or overoptimism regarding task completion costs affect our results. We also discuss the possibility of recall and implications of deadlines for the agent's well-being.

Sophistication. So far, we have assumed that the agent is naive about her cognitive limitation, i.e., she does not anticipate that the task might drop off the top of her mind (compare Holman and Zaidi, 2010). If we instead assume that the agent exhibits cognitive limitations, but is perfectly aware of them—i.e., she is fully sophisticated—this alters our results as follows.[11]

If the agent does not complete the task in period t, the present value of her payoff is $(1-\gamma)\delta V_{t+1}^S$. Hence, her optimal decision rule is to complete the task in period t if and only if

$$c_t \leq \hat{c}_t^S := y - (1-\gamma)\delta V_{t+1}^S. \tag{17}$$

In contrast to a naive agent, the sophisticated agent thus adjusts her decision rule to the inattention rate γ. In particular, when the agent decides on whether she completes the task in period t, she takes into account that it might otherwise drop off her mind. Her cost threshold for completing the task in t is thus increasing in the inattention rate. All else equal, the cost threshold of a sophisticated agent is therefore higher than for a naive agent or an agent without cognitive limitations, who both decide based on $\gamma = 0$.

Intuitively, one might expect that sophistication completely revokes the effects of cognitive limitations, such that the hazard rate $\tilde{h}_t^S$ is again increasing for all $t \in \{1, ..., T\}$. This conjecture, however, turns out to be wrong. Let us illustrate this point for the case of $T = 2$. In the second period, the agent's cost threshold for completing the task is $\hat{c}_2 = y$. This holds for both the naive and sophisticated agent. The second-period hazard rate is then $\tilde{h}_2 = \tilde{h}_2^S = (1-\gamma)F(y)$, which is again independent of the agent's naivete or sophistication. From above we know that in period $t = 1$, the sophisticated agent anticipates the possibility that she might become inattentive at the end of the period. She therefore applies a higher cost threshold in $t = 1$ than a naive agent. Consequently, the first-period hazard rate is higher for a sophisticated than the one for a naive agent.[12]

[11]The arguments are similar for the case of partial sophistication, i.e., when the agent's perceived probability of becoming inattentive, $\hat{\gamma}$, is positive but lower than the true rate of inattention, $0 < \hat{\gamma} < \gamma$.

[12]To be precise, $\hat{c}_1^S = y - (1-\gamma)\delta \int_0^y (y-c)dF(c)$ while $\hat{c}_1 = y - \delta \int_0^y (y-c)dF(c)$, such that $\tilde{h}_1^S = F(\hat{c}_1^S) > F(\hat{c}_1) = \tilde{h}_1$.

Hence, whenever the hazard rate is decreasing over time for a naive agent, the same must also hold true for a sophisticated agent. Formally, if $\tilde{h}_1 > \tilde{h}_2$ then $\tilde{h}_1^S > \tilde{h}_2^S$. A more detailed discussion of sophistication is provided in Appendix A.6.

Present Bias. Suppose now that the agent has time-inconsistent, present-biased preferences as captured by the well-known $\beta - \delta$ approach (e.g., Laibson, 1997). If the agent bears the costs of completing the task immediately but receives the reward with an ℓ-period delay, she completes the task in period t if and only if

$$c_t \leq \hat{c}_t^B := \beta\left(\delta^\ell Y - \delta V_{t+1}\right), \tag{18}$$

where the present bias parameter β is between zero and one. Comparing the cost thresholds in (18) and (1), we see that a present-biased agent uses a lower cost threshold than a time-consistent one. Except for inducing lower cost thresholds, however, the analysis is not affected by the present bias. In particular, if a present-biased agent has no (further) cognitive limitations (i.e., if $\gamma = 0$), hazard rates are still increasing over time and only depend on the time left until the deadline but not on the absolute value of time.[13]

Overoptimism. We next analyze what happens when the agent is overoptimistic regarding the costs of task completion, in the sense that she overestimates the likelihood of low cost realizations and underestimates the likelihood of high cost realizations. Formally, an overoptimistic agent bases her decisions on a perceived cost distribution F^o, while her costs are actually drawn from distribution F, where $F^o(c) > F(c)$ for all $c \in (0, \bar{k})$, $F^o(0) = F(0) = 0$, and $F^o(\bar{k}) = F(\bar{k}) = 1$. The overoptimistic agent completes the task in period t if and only if

$$c_t \leq \hat{c}_t^o = y - \delta V_t^o. \tag{19}$$

In period T, the perceived option value of an optimistic agent is higher than the option value of an agent with realistic beliefs:

$$V_T^o = \int_0^{\hat{c}_T^o = y} (y - c)dF^o(c) > \int_0^{\hat{c}_T = y} (y - c)dF(c) = V_T. \tag{20}$$

Since

$$V_t^o = \int_0^{\hat{c}_t^o} (y - c)dF^o(c) + (1 - F^o(\hat{c}_t^o))\,\delta V_{t+1}^o > \int_0^{\hat{c}_t} (y - c)dF(c) + (1 - F(\hat{c}_t))\,\delta V_{t+1} = V_t, \tag{21}$$

the same holds true in period $T - 1$ and, by induction, also in all previous periods. Hence, overoptimism causes the agent to perceive the option value to be strictly higher than its true value in all periods $t \leq T$. From (1) and (19), it then follows that an overoptimistic agent uses lower cost thresholds in all $t \leq T - 1$, compared to an agent with realistic beliefs and no cognitive limitations. Similar to a present-biased agent, the overoptimistic one therefore exhibits, *ceteris paribus*, a higher likelihood to postpone the task in any given

[13]For an analysis capturing possible interactions between present-bias and memory limitations see Ericson (2017).

period. Beyond these quantitative differences, however, results remain qualitatively unchanged relative to the no-cognitive-limitations case.

Recall. By continuity, our results are also robust to the case where—after becoming inattentive—the agent may later return to processing the task with a small positive probability. More generally, one can show that, concerning both hazard and task completion rates, a model with positive probability of 'recall' is isomorphic to the model with 'zero recall' and time-dependent probabilities of becoming inattentive $(\gamma_1, ..., \gamma_T)$.

Deadlines and the Agent's Well-being. An interesting implication of the above analysis is that the expected payoff of an agent *without* cognitive limitations is monotonically increasing in the length of the deadline. Formally, the first-period option value V_1 measures the agent's expected payoff, which is increasing in T by (3) and (7). When we allow for cognitive limitations, however, this result may revert. To illustrate this point, let γ be large, $\gamma \to 1$. The cognitively limited agent's expected payoff is then given by $\int_0^{\hat{c}_1}(y - c)dF(c)$. Since the threshold $\hat{c}_1$ is below the value of the reward y and decreases in the length of the deadline T, the expected payoff is *decreasing* in T. Cognitively limited agents might therefore be better off under a shorter deadline, whereas agents without cognitive limitations always benefits from the higher flexibility associated with longer deadlines.

3 The Experiments

To assess the competing predictions from our theoretical analysis empirically, we conducted two field experiments at a large dental clinic. The experiments study how deadlines shape the timing and frequency at which patients arrange appointments for preventive check-ups. Several features of our setting make it ideally suited to explore the interplay between deadlines, cognitive limitations, and task completion. First, check-ups should be conducted regularly, but at a relatively low frequency.[14] Furthermore, patients at our collaboration partner's clinic have to actively contact the dentist to arrange a new check-up appointment. Hence, we study a setting in which people have to actively carry out a task, but have no well-defined date for when to do so. This is similar to a variety of other tasks, like adapting one's insurance or savings plans, switching energy or phone providers, etc.

Anticipating that patients frequently fail to regularly schedule check-up appointments,[15] the dentist uses a reminder system. Whenever a patient is up for the next check-up according to her recommended check-up interval (see fn. 14), the system sends a postcard reminding her to arrange an appointment. After receiving the postcard, patients have the task on the top of their mind—just as in the first period of our theoretical

[14] In the case of dental health, it is widely agreed that regular check-ups are a key factor of effective health prevention, since many dental diseases are asymptomatic in their early stages but can be discovered and treated through professional check-ups (see, e.g., Lang et al. 1994). Dentists in Germany typically suggest check-up intervals of 6 months and somewhat shorter intervals of 3–4 months for patients with an elevated risk of developing certain dental diseases.

[15] In an independent online survey that we conducted with roughly 3,000 individuals in Germany, 56% of respondents indicate that they aim at having (at least) two dental check-ups per year. Among this group, 43% did not act accordingly during the past 2 years. Compare Section 4.3 and Table C.5 in the Online Appendix for further details.

framework. The costs of subsequently carrying out the task, however, may fluctuate, depending for example on other duties that a patient faces. Over time, the task might drop from the top of a patient's mind if her cognitive and attentional capacities are limited.

3.1 Main Experiment

The institutional setting at the dental clinic give us the opportunity to systematically study how deadlines affect individuals' behavior.[16] Specifically, we used the postcard reminders to introduce and exogenously vary deadlines faced by patients. If a patient contacted the dentist until a specified deadline, she was eligible to receive a reward. In our main experiment, we varied the length of the deadline across three different treatment conditions. In the *D1* treatment, participants faced a short deadline that was set one week after the sending day of the postcards. In treatment *D3*, we implemented a longer deadline of three weeks. Finally, in the *ND* treatment, the postcards contained no deadline. Hence, the reward was, in principle, available for an unlimited time period.

We implemented two different reward levels: a 'small present' and a large reward. As a small present, participants could choose between various dental care products, such as toothpaste, dental floss, etc. The large reward consisted of a voucher for a free professional dental cleaning, which would otherwise cost the average patient about 70 euros.[17] In addition, we also had treatments that involved no explicit reward, but simply asked the patients to contact the dentist before the communicated deadline. Interacting the three deadline treatments (D1, D3 and ND) with the three reward levels leaves us with a total of 9 treatment cells in a 3×3 factorial design.[18] Since we are mainly interested in the impact of deadlines, we will focus on the deadline dimension when discussing our results in Section 4. Most of our analysis, therefore, pools observations with identical deadline lengths that involve different rewards. Note that this approach is valid as we assign people randomly and independently to the different treatment cells. Throughout our empirical analysis, however, we will also present results that control for differences in reward levels. We discuss additional results for the individual treatment cells in Section 4.3.

Procedures and Outcomes. The main experiment was conducted between 2011 and 2013. During the experiment, the dentist sent out reminders every second Friday, implying that the postcards are delivered to participants on Saturday or Monday, at the latest. The deadlines of the D1 and the D3 treatment were set for the Friday one and three weeks after the sending day, respectively. Together with the biweekly sending waves,

[16] From a methodological perspective, the pre-existing reminder system is also ideal in another respect: it provides a natural environment that is unlikely to raise questions about experimenter scrutiny or an ongoing experiment. In particular, patients of the clinic are already used to receiving reminder messages, and they are also used to receiving slightly different messages over time, as the dentist regularly varied the design and content of the reminder postcards in the past (see Altmann and Traxler 2014).

[17] One might wonder whether patients perceive a professional dental cleaning as a reward. Reassuringly, 64% of the participants in our online survey (see fn. 15 and Section 4.3) state that they do not find the procedure to be unpleasant; 93% would 'rather' or 'definitely' make use of the procedure if it was offered free of charge by their health insurance plan. Regarding the smaller reward, 73% of respondents in our online survey (correctly) expect a 'small present' in this context to be worth roughly 2.50 euros (see Table C.5 in the Online Appendix for further details).

[18] An overview of the treatment matrix and an example of a postcard from the experiment is provided in the Online Appendix.

this implies that the deadline date of the D3 treatment from wave W coincides with the deadline from the D1 treatment in wave $W + 1$. We will exploit this feature below.

In each sending wave W, we randomly assigned patients to treatments who were up for their next check-up at the time of the sending wave.[19] Members of the same household were assigned to the same treatment in each wave. This procedure aims at minimizing the risk of possible treatment spillovers. In total, we observe N = 2,661 patient-treatment observations across 43 randomization waves. The observations cover 1,175 individuals living in 1,015 different households. These numbers reflect that—due to the long overall duration of the experiment and the recommended check-up interval of typically six months (see fn. 14)—most patients were treated more than once. Our procedure ensures that treatments are randomized independently, each time a patient is up for a new appointment.

Our main dependent variable is the date at which a patient contacts the dentist to arrange a check-up appointment. We derive this variable from a data set that covers all inbound phone calls at the dental clinic between 2011 and mid 2014. As indicated above, the date at which a participant first contacts the dentist to make the appointment (rather than the actual date of the appointment) is the relevant date which determines whether she is eligible for receiving the reward. Moreover, it is also a natural outcome variable for evaluating the impact of our treatments, since it is the first observable response of participants. Finally, using this response measure instead of the date of the check-up avoids potential issues of congestion in the dentist's schedule. We nevertheless do measure whether a participant missed (or re-scheduled) an appointment. Such 'no-shows' are rare in our sample—the overall no-show rate is 4.6%—and do not systematically vary across treatments.

Our data analysis focuses on participants' responses during the first 100 days after the sending date of each wave. Studying longer response periods is potentially problematic, because some participants (e.g., diabetes mellitus patients and other patient groups who exhibit an increased risk of developing certain dental diseases) will already receive their next check-up reminder after 3–4 months. In addition to measuring individuals' responses to our intervention, we also observe a broad range of sociodemographic characteristics and health-related variables, such as patients' health insurance status, history of major dental treatments, and their dental health classification ('at-risk'). Table 1 provides summary statistics for these background characteristics across treatments (see table notes for a definition of the individual variables). The last column of the table reports results from validation checks for random treatment assignment. F-tests from separate regressions of the individual characteristics on dummies for the different deadline conditions indicate no significant treatment differences for nine out of ten variables. Hence, apart from some minor differences (which we will account for in our later analysis), randomization succeeded in generating samples that are well balanced across treatments.

[19] Note that some patients might 'select out' of the experiment by already arranging a check-up well before their next due date. Our data suggests that this selection effect is, at best, modest.

Table 1: Summary Statistics and Balancing Tests (Main Experiment)

	D1	D3	ND	F-Stats
Age	35.70	35.75	36.24	0.311
	(15.88)	(16.48)	(15.98)	[0.732]
Female	0.56	0.59	0.54	2.074
	(0.50)	(0.49)	(0.50)	[0.126]
Distance	12.75	10.10	11.63	1.572
	(43.94)	(34.27)	(42.06)	[0.208]
Housing price	8.02	8.06	8.01	0.454
	(1.14)	(1.13)	(1.10)	[0.635]
Private HI	0.19	0.19	0.21	0.606
	(0.39)	(0.39)	(0.41)	[0.546]
Family	0.17	0.15	0.16	0.462
	(0.38)	(0.36)	(0.37)	[0.630]
At risk	0.11	0.11	0.10	0.593
	(0.32)	(0.31)	(0.30)	[0.553]
Patient retention	4.51	4.11	4.27	4.886
	(2.74)	(2.65)	(2.58)	[0.008]
Past showup$^\flat$	0.73	0.70	0.74	1.455
	(0.45)	(0.46)	(0.44)	[0.228]
Pain$^\flat$	0.22	0.25	0.27	2.029
	(0.42)	(0.44)	(0.44)	[0.132]
N	907	879	875	

Notes: The table presents summary statistics (mean values, standard deviations in parentheses) for individuals' background characteristics, separately for each of the three deadline conditions. The overall sample size is N = 2,661. 'Distance' denotes the great-circle kilometer distance between a person's home address and the dentist. 'Housing price' is the average rental price at the participant's home address (euros per square meter, excluding utilities and dues). 'Family' is a dummy indicating that several household members took part in the study. 'Private HI' is a dummy indicating whether a patient is covered by private health insurance (0 for public health insurance). 'At risk' indicates that a participant is recommended a shortened check-up interval of 3 or 4 months. 'Patient retention' is the number of years since a participant first visited the dentist. The dummy 'Past showup' indicates whether a participant had at least one check-up in the year prior to a given treatment. 'Pain' indicates if a participant was exposed to a painful dental treatment in the past. The last two variables (labeled with $\flat$) are only available for 2,189 [Past showup] and 1,990 [Pain] observations, respectively. The final column reports F-statistics [and the corresponding p-values in brackets] from tests for treatment differences based on separate regressions of each of the characteristics on dummies for the different deadline conditions.

3.2 Follow-up Experiment

Following the main experiment described above, we conducted a second experiment in the same institutional setting. Besides assessing the general robustness of our main results, the smaller follow-up experiment aimed at exploring whether our key findings remain qualitatively stable under longer deadlines. To this end, the follow-up experiment included—in addition to the D1, D3, and ND treatment—treatments with a 6-week (*D6*), a 10-week (*D10*), and an 'end-of-year' deadline (*EoY*), in which the deadline was set to December 31.[20] As a reward, the follow-up experiment involved a dental-care kit worth 10 euros. In all other respects, the procedures were identical to those in our main experiment.

[20]This corresponds to a deadline length of 5–7 months, depending on the randomization wave. Note that the *EoY* treatment was rolled out only in May 2013, resulting in a slightly smaller treatment cell for this treatment.

The follow-up experiment was conducted during the 2nd and 3rd quarter of 2013. Table B.1 in the Appendix presents an overview of participants' background characteristics and randomization checks. The overall number of patient-treatment observations is $N = 927$ (including 798 individuals from 642 households). Together with the higher number of deadlines, the lower number of observations implies that treatment cells in the follow-up experiment are smaller than the ones in our main experiment. Consequently, the statistical power in our analyses of the follow-up experiment will be limited. As pointed out above, the second experiment should thus primarily be seen as a qualitative replication and extension of our main experiment.

3.3 Linking up with the Theoretical Framework

Before turning to the empirical results, it is worth reconsidering our theoretical predictions in light of our empirical setting. As discussed in Section 2, deadlines matter both in our baseline model as well as in the framework with cognitive limitations. However, the empirically observed response patterns under different deadlines will help us shed light on the relevance of cognitive limitations. The analysis from Section 2 pointed to four (interrelated) metrics that are of particular interest in this respect: (i) the shape of the hazard rates over time, (ii) whether hazard rates depend only on the time left until the deadline or also on the absolute value of time, (iii) the rate at which agents complete the task within the deadline, and (iv) the fraction of agents who have completed the task until a given point in time t.

The data from our experiments provide direct measures for all these metrics. In particular, the date at which patients contact the dentist to arrange a check-up appointment allows us to calculate hazard rates for the different treatments (i.e., measuring the patients' probability to respond, conditional on not having responded to the postcard before). In the same vein, we readily observe 'period-t completion rates' (i.e., the cumulative fraction of participants who have contacted the dentist within t days after the intervention) and 'within-deadline completion rates' (i.e., the fraction of patients who have called before the deadline stipulated in their treatment).

Despite the direct linkage between the variables of interest from our theoretical analysis and their empirical counterparts, there are a few details that differ between our theoretical framework and the empirical setting. First, subjects in the experiment can (and do) call the clinic after the deadline has passed. This is due to the fact that, in addition to the explicit rewards that are tied to the deadlines, patients in the experiment face further implicit incentives to arrange a check-up appointment (e.g., related to expected improvements in future health).[21] Hence, there exist positive 'rewards' for executing the task after the deadline. For the sake of simplicity, our theoretical analysis abstracts from this possibility. In Section A.7 in the Appendix, we extend the model to allow for positive rewards and non-zero task completion rates after the deadline. Doing so is straightforward and does not change the model's key qualitative predictions.

[21] For certain groups of patients, there also exist further economic incentives to arrange check-ups regularly. Participants covered by the German public health insurance (more than 80% of our sample), for instance, receive a 20% [30%] allowance on treatment costs if they had at least one check-up per year for the last 5 [10] years. Note that these other explicit and implicit incentives are, by construction, orthogonal to our treatment intervention.

Secondly, our setting does not only include costs for *arranging* a check-up but also costs for attending the appointment. One has to keep in mind, however, that the primary task in the experiment (the action that is tied to the deadline) is the first step, i.e., calling to arrange an appointment. Irrespective of this observation, the theoretical model can be readily extended to capture costs of arranging and attending an appointment (the idea is that in the first decision, the agent compares the costs of arranging to the reward net of the (expected) attendance costs).

A third point worth noting concerns our treatments with no explicit rewards. Without a loss in rewards, one would, in principle, expect no drop in hazard rates around the deadlines. Patients, however, might also perceive a drop in some of the implicit incentives involved in our setting. This could, for instance, be the case if patients exhibit social image concerns (e.g., they call in a timely manner to signal to the dentist that they care about their dental health), or if they perceive the deadline as a medical recommendation. We will return to this point in Section 4.3 below, in which we will also discuss additional survey evidence on how patients perceive the deadlines and underlying incentives in our setting.

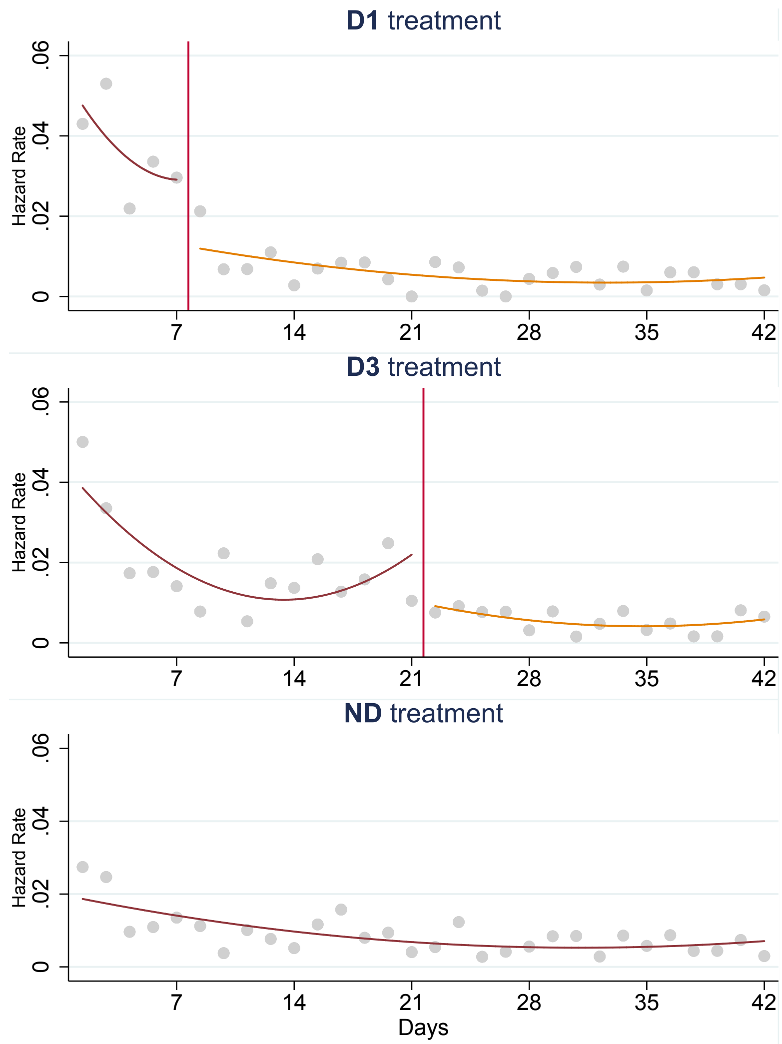
4 Empirical Results

4.1 Main Experiment

Hazard Rates over Time. In a first step, we analyze how deadlines affect the shape of hazard rates across treatments. Figure 1 plots hazard rates for the D1, D3, and the ND treatment. The gray dots depict the empirically observed hazard rates on a daily basis together with quadratic fits, estimated separately for the pre- and the post-deadline period. For both the D1 and D3 treatment, the evidence contrasts sharply with the prediction of monotonically increasing hazard rates (Hypothesis 1). Instead, under both deadlines, hazard rates tend to decline during the first week of the experiment. Going from the second to the third week, we observe a modest increase in hazard rates in the D3 treatment (see middle panel of Figure 1). Over the three weeks before the deadline, the hazard function in this treatment thus exhibits a U-shaped pattern. Both the initially declining hazards and the increase in hazard rates close to the deadline are consistent with the predictions of the cognitive-limitations model (Hypothesis 1'), in which the interplay between the cognitive-limitation effect and the option-value effect can give rise to U-shaped hazard rates.

For the no-deadline treatment (displayed in the bottom panel of Figure 1), the data also show a declining pattern in hazard rates. This observation—which is further corroborated when considering hazard rates over a longer time horizon (see Figure B.1 in the Appendix)—is again in line with the cognitive-limitations model (Hypothesis 1') but in conflict with the constant hazard rate predicted by the baseline model for the no-deadline case (Hypothesis 1).

Figure 1: Daily Hazard Rates (Main Experiment)



Notes: The figure displays daily hazard rates for the treatments with a one-week deadline (top panel), three-week deadline (middle panel) and no deadline (bottom panel) over a period of six weeks (42 days) after sending the mailings.

The fitted curves in the three panels of Figure 1 indicate that hazard rates in all treatments are decreasing initially.[22] To examine the shape of the hazard rates in more detail, we next conduct a duration analysis and investigate time-varying treatment effects on the hazard rates. We start from a proportional hazards model,

$$h_t = h_{0,t} \exp(\beta_0 + \beta_1 \,\mathrm{D1} + \beta_2 \mathrm{D3}), \tag{22}$$

where D1 and D3 are dummies for the corresponding treatments. The Cox proportional hazards model leaves the baseline hazard function $h_{0,t}$ unspecified (estimating it non-parametrically). In addition to the Cox model,

[22]The estimates for the (linear) slope coefficient underlying the pre-deadline fits are significantly negative (at the 1%-level) for the D3 and the ND treatment. For the D1 treatment, this coefficient from the quadratic model is statistically insignificant. It becomes significantly negative in a linear model (which, for the D1 treatment, performs better than a quadratic fit in terms of AIC).

we also examine a parametric specification of $h_{0,t}$. More specifically, we make use of the Weibull model where $h_{0,t} = \rho t^{\rho-1}$. Depending on the level of ρ—the (estimated) shape parameter of the Weibull distribution—the model can accommodate a baseline hazard that is either decreasing ($0 < \rho < 1$), constant ($\rho = 1$), or increasing ($\rho > 1$). The Weibull model thus allows for a straightforward test of Hypothesis 1, which predicts a constant hazard rate for the ND treatment, i.e., for D1 $=$ D3 $= 0$, in which case the hazard rate is $\rho t^{\rho-1} \exp(\beta_0)$.

Table 2: Duration Analysis (Main Experiment)

	(1)	(2)	(3)
D1	1.135	1.157	1.158
	[0.103]	[0.066]	[0.074]
D3	1.180	1.215	1.223
	[0.028]	[0.012]	[0.012]
Small Incent		1.085	1.086
		[0.336]	[0.346]
Large Incent		1.429	1.441
		[0.000]	[0.000]
ρ (shape parameter)			0.565
			[0.000]
Controls	–	Yes	Yes

Notes: The table presents hazard ratios estimated with Cox (Columns 1 and 2) and Weibull (Column 3) proportional hazards models (N = 2,661). Specifications (2) and (3) include dummies for different incentive levels, controls for individual characteristics, and dummies absorbing wave specific effects. p-values—based on robust standard errors, clustered at the household level—are reported in brackets. ρ is the estimated shape parameter of the Weibull distribution, with $\rho < 1$ indicating a decreasing hazard rate.

Estimated hazard ratios for equation (22)—based on all responses during the first 100 days after the intervention—are presented in Table 2. Columns (1) and (2) depict results from Cox proportional hazards models. The estimated hazard ratios for both deadline treatments are larger than one. Compared to the treatment with no deadline, we observe roughly 15 and 20% higher hazards in the D1 and D3 treatment, respectively. When we add controls for the different incentive levels, individuals' background characteristics and dummies absorbing sending-wave specific effects, the estimates become slightly more precise and turn statistically significant for both deadlines ($p = 0.066$ and $p = 0.012$, respectively; see Column (2) of Table 2). Column (3) reports the estimated hazard ratios from the Weibull model. The estimates are hardly distinguishable from those obtained in the Cox model reported in Column (2). The estimated shape parameter of the Weibull distribution is significantly smaller than one ($\rho = 0.565$, $p < 0.001$).[23] Hence, the estimates clearly reject the prediction of a constant hazard rate for the no-deadline treatment (Hypothesis 1). Instead,

[23]Estimating specification (3) for the ND treatment alone, one obtains $\rho = 0.672$, which is again significantly different from one ($p < 0.001$).

the estimated ρ supports the prediction of decreasing hazard rates from the cognitive-limitations model (Hypothesis 1').

Moving beyond the impact of deadlines on average hazard rates, we next estimate augmented models that allow for time-varying treatment effects. Specifically, we estimate treatment effects separately for each of the first three weeks after the intervention. This approach allows us to test further predictions from our theoretical analysis. Estimation results are presented in Table 3. The first specification estimates time-specific treatment effects by including separate treatment dummies for the first, second, and third week after the intervention, as well as a treatment dummy for all later weeks; Column (2) adds controls for the different incentive levels, individual background characteristics and sending-wave dummies; Column (3) further accounts for (calendar) week×year specific effects. Note that the last specification is quite conservative, to the extent that the combination of wave and week×year dummies absorbs some of the time-varying impact of the treatments that we want to capture.[24]

The estimation results document strong and highly significant effects of the different deadlines. Let us first consider the treatment with the short deadline. During the first week, the hazard rate in the D1 treatment is more than three times larger than in the ND treatment (see estimated hazard ratio for D1^{w1}). However, as already observed in Figure 1, the strong effect of the short deadline is fully concentrated in the first week. For none of the later weeks do we estimate a significant increase in the hazard relative to the ND treatment. In fact, for the third and all later weeks (weeks 4–15), the estimated hazard ratios are smaller than one, suggesting that response rates in the ND treatment catch up relative to the D1 treatment in this phase—a point that we will discuss further below.

For the three-week deadline, the estimates again show a strong impact of the deadline during the first week of the experiment, with a highly significant hazard ratio of about 2.5. Moreover, the results confirm the U-shape in the hazard rates that we already noted in Figure 1: moving from the first to the second week, the hazard ratio shrinks substantially; from the second to the third week, hazards increase again. Both the initial decline as well as the later increase are statistically significant (Wald tests, $p = 0.026$ and 0.062, respectively). When adding additional controls, this pattern remains qualitatively robust. Only for the heavily saturated specification from Column (3) of Table 3, the U-shape looses statistical significance.

RESULT 1: *(i) Neither under the one-week deadline nor under the three-week deadline do we find monotonically increasing pre-deadline hazards: hazard rates exhibit a decreasing (D1 treatment) or U-shaped pattern (D3 tretment) before T. (ii) Without a deadline (ND treatment), hazard rates are decreasing over time. The evidence rejects Hypothesis 1 and supports Hypothesis 1'.*

[24] Specification (3) seems to be overly saturated to non-parametrically estimate the baseline hazard. For the sake of comparability, we therefore estimate all specifications with a Weibull rather than a Cox model. Note further that we replicated all results for time-varying treatment effects with binary response models (in particular: complementary log-log regressions as detailed, e.g., in Sueyoshi, 1995) for discrete time duration analysis.

Table 3: Duration Analysis with Time-Varying Treatment Effects (Main Experiment)

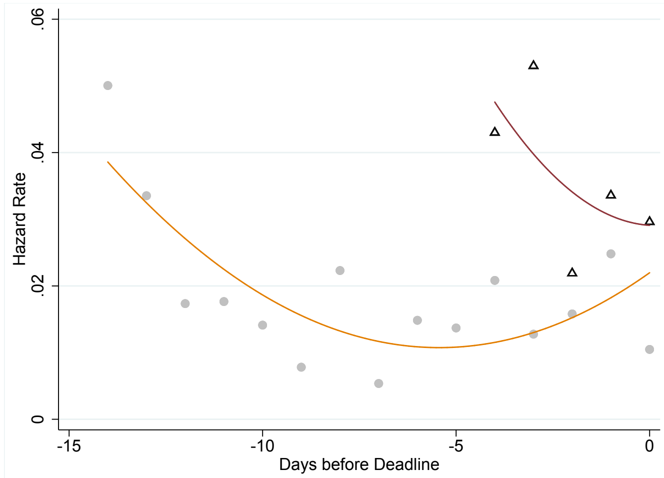
	(1)	(2)	(3)
$D1^{w1}$	3.506	3.574	3.312
	[0.000]	[0.000]	[0.000]
$D1^{w2}$	1.266	1.299	1.237
	[0.193]	[0.154]	[0.290]
$D1^{w3}$	0.819	0.835	0.615
	[0.375]	[0.426]	[0.038]
$D1^{w4+}$	0.672	0.675	0.726
	[0.000]	[0.000]	[0.002]
$D3^{w1}$	2.588	2.632	2.400
	[0.000]	[0.000]	[0.000]
$D3^{w2}$	1.667	1.717	1.693
	[0.003]	[0.002]	[0.005]
$D3^{w3}$	2.444	2.505	1.805
	[0.000]	[0.000]	[0.000]
$D3^{w4+}$	0.678	0.699	0.752
	[0.000]	[0.001]	[0.011]
Controls	–	Yes^a	Yes^b
Post-Estimation Tests:			
$D3^{w1} = D3^{w2}$	0.026	0.030	0.114
$D3^{w2} = D3^{w3}$	0.062	0.063	0.778
$D1^{w1} = D3^{w3}$	0.041	0.046	0.001

Notes: The table presents hazard ratios estimated with Weibull proportional hazards model (N = 2,661). Specification (2) includes dummies for different incentive levels, controls for individual characteristics, and dummies absorbing wave specific effects. Specification (3) adds further dummies that account for week×year specific effects. p-values—based on robust standard errors, clustered at the household level—are reported in brackets. The lower part of the table reports p-values from Wald tests.

Hazard Rates and Distance to the Deadline. Next we examine how hazard rates depend on the distance to the deadline and the absolute value of time. Remember that, absent cognitive limitations, the hazard rate in a given period should only depend on how much time there is left until the deadline, but not on the absolute value of time (Hypothesis 2). The hazard rate in the first week of the D1 treatment should therefore be identical to the third-week hazard in the D3 treatment. We can reject this hypothesis for all three specifications reported in Table 3 (with $p = 0.041$, $p = 0.046$ and $p = 0.001$ respectively): under the short deadline, the observed hazard in the pre-deadline week (i.e., during week 1) is always higher than the hazard during the pre-deadline week of the D3 treatment (i.e., week 3). Hence, the evidence on pre-deadline hazards is in conflict with Hypothesis 2 from our baseline model but consistent with the cognitive-limitations model, in which longer deadlines imply that hazard rates are lower in a given pre-deadline period (Hypothesis 2').

One might question the comparability of the pre-deadline weeks between the short and the long deadline, since for the D1 treatment this period also coincides with the first week after the intervention. Note first

Figure 2: Daily Hazard Rates (Main Experiment)



Notes: The figure plots daily hazard rates for the pre-deadline period, comparing the D1 (indicated by the triangles, △) and the D3 treatment (gray dots, ●).

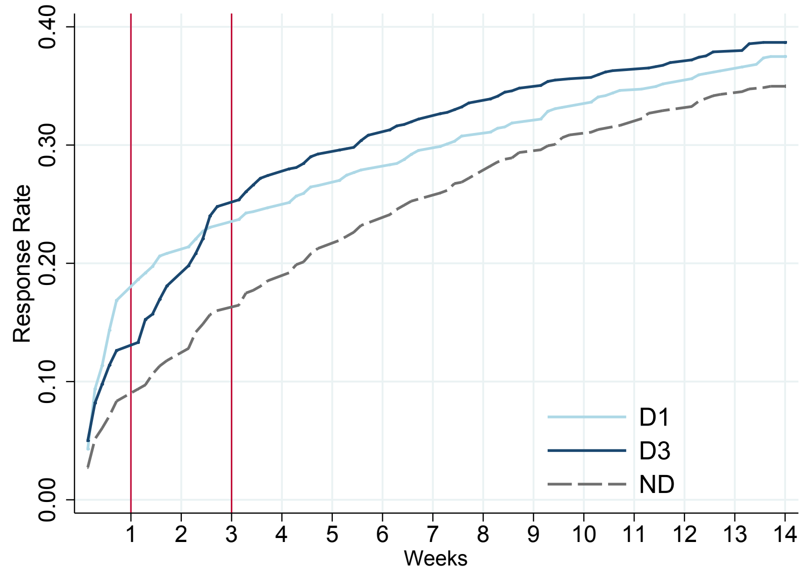
that this issue is irrelevant from a theoretical perspective. In addition, we can address this point further by comparing differences in *daily* rather than *weekly* hazards. Doing so yields similar results as before. This point is illustrated in Figure 2 (a re-scaled version of Figure 1), which compares the daily pre-deadline hazards between the two deadline treatments, centered around the day of the deadline. The figure demonstrates that the empirically observed hazard rates for each of the five days before the deadline are higher in the D1 than in the D3 treatment. To further back this observation, we estimate models in the spirit of those reported in Table 3 on a daily rather than weekly basis (see Table B.2 in the Appendix). The results show that hazards for the last day (or the last two days) before the deadline are higher in the D1 than in the D3 treatment.[25] These findings lead us to

RESULT 2: *Holding the distance to the deadline constant, the observed hazard rates depend negatively on the absolute value of time: for a given pre-deadline period, hazard rates are lower under the three-week deadline than under the one-week deadline. This rejects Hypothesis 2 and supports Hypothesis 2'.*

Completion Rates. In a next step, we examine how the observed differences in hazard rates translate into treatment differences in the cumulative response (or 'task completion') rates. Figure 3 shows the evolution of response rates under the different deadlines over time. More specifically, the figure depicts the (cumulative) fraction of individuals who have called to make an appointment until a given point in time, comparing individuals who were confronted with a one-week deadline (light blue line), three-week deadline (dark blue line), or no deadline (dashed grey line).

[25]While the differences in hazard ratios are substantial, the underlying coefficients for the daily hazards are, naturally, less precisely estimated than in the model with weekly hazards, rendering some of the treatment differences in daily hazards statistically insignificant or only weakly significant (see Table B.2 in the Appendix).

Figure 3: Cumulative Response Rates (Main Experiment)



Notes: The figure presents the empirically observed cumulative response rates over 14 weeks.

Comparing these response rates at a given point in time—the *period-t completion rate* in the language of our model—reveals several interesting facts: first, as already indicated by the hazard rate analysis, we observe that completion rates during the first week of the intervention are highest under the short deadline. At the end of the first week, the cumulative response rate is roughly 4 percentage points higher in the D1 than in the D3 treatment. Estimates from linear probability models, which are presented in Table 4, show that this difference is statistically significant ($p = 0.017$ and $p = 0.011$ in specifications with and without controls, respectively; see the post-estimation tests for Columns (1) and (2) of Table 4). Compared to the no-deadline treatment, the one-week completion rate is more than twice as high in the D1 treatment (16.8% vs. 8.3%, $p < 0.001$; see Columns (1) and (2) of Table 4). This illustrates again the strong impact of imposing a short deadline on responses during the first week.

At the end of the third week, the completion rate is about 2 percentage points higher in the D3 than in the D1 treatment. The difference is modest and statistically insignificant ($p = 0.401$ and $p = 0.352$; see post-estimation tests for Columns (3) and (4) of Table 4). However, the numbers for both deadline treatments lie significantly above the 3-week completion rate of 16% observed in the no-deadline treatment: the gap ranges from 7 (D1, $p < 0.001$) to almost 9 percentage points (D3, $p < 0.001$). Hence, imposing a deadline increases the task completion rate within the first three weeks by 45–55% relative to the no-deadline case.

A further interesting metric from our theoretical analysis is the *within-deadline completion rate*. Comparing these rates between the D1 and D3 treatment, we find that 16.9% of participants respond within the one-week deadline of D1, whereas 24.8% do so within the three-week deadline in D3. This difference of almost 8 percentage points is large and statistically significant (t-test, $p < 0.001$), suggesting that—for the

Table 4: Treatment Effects on Period-t Completion Rates (Main Experiment)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	7 Days		21 Days		100 Days		100 Days	
D1	0.085***	0.086***	0.070***	0.069***	0.028	0.028		
	(0.017)	(0.017)	(0.020)	(0.019)	(0.023)	(0.022)		
D3	0.043***	0.044***	0.088***	0.088***	0.042*	0.042*		
	(0.015)	(0.014)	(0.019)	(0.019)	(0.023)	(0.023)		
D1 or D3							0.035*	0.036*
							(0.020)	(0.020)
Large Incent		0.087***		0.090***		0.097***		0.096***
		(0.019)		(0.022)		(0.026)		(0.026)
Small Incent		-0.011		-0.022		0.028		0.025
		(0.014)		(0.018)		(0.022)		(0.022)
Constant	0.083***	0.020	0.160***	0.007	0.353***	-0.042	0.353***	0.028
	(0.010)	(0.074)	(0.014)	(0.097)	(0.019)	(0.121)	(0.019)	(0.119)
Post-Estimation Tests: (p-value)								
D1=D3	0.017	0.011	0.401	0.352	0.564	0.502	–	–
Controls	–	Yes	–	Yes	–	Yes	–	Yes

Notes: The table presents LPM estimates of the treatment effects on the cumulative response rate (the probability of calling) within 7, 21, and 100 days (N = 2,661). The lower part reports the p-values from Wald tests. Every second specification includes individual control variables and dummies absorbing wave-specific effects. Robust standard errors, clustered at the household level, are reported in parentheses. ***, **, * indicates significance at the 1%-, 5%-, 10%-level, respectively.

case of a one- vs. three-week deadline—the longer deadline triggers a higher within-deadline completion rate in our setting.

RESULT 3: *The within-deadline completion rate is higher in the D3 than in the D1 treatment.*

The evidence on within-deadline completion rates is qualitatively in line with the predictions of our baseline model (Hypothesis 3) and the cognitive-limitations framework (Hypothesis 3') with modest cognitive constraints (i.e., a modest level of γ). A similar picture emerges from the comparison of period-t completion rates during the second and third week after the intervention. During this period, cumulative response rates in the D3 treatment catch up and, after about 2.5 weeks, overtake the response observed for the D1 treatment. This pattern is again consistent with both Hypothesis 4 and 4' (and, in particular, with the prediction that cumulative responses intersect in a period $T' < \tau < T''$).

For the period after the third week, however, our baseline model predicts that completion rates in the no-deadline treatment should catch up, and eventually—after some threshold period $\tau > 3$ weeks—overtake those from the D1 and D3 treatment (see Hypothesis 4 and the discussion in Appendix A.7). The evidence from Figure 3 and the estimates reported in Table 4 suggest that this is not the case, at least not within the first 100 days after the intervention. While there is some catch-up of response rates in the ND treatment—as it is also indicated in the estimates from Table 3—the 100-day completion rates for the deadline treatments remain between 2.8 and 4.2 percentage points above the rate in the ND treatment ($p = 0.224$ [$p = 0.070$] when comparing D1 [D3] with ND, and $p = 0.082$ when comparing both deadline treatments jointly with ND; see Columns (5) and (7) in Table 4).

Note that these findings do not necessarily reject Hypothesis 4 from the baseline model: the cutoff date τ, after which the completion rate in the ND treatment starts to dominate, might simply be very large. However, even if we extend our observation period up to six months after the intervention—ignoring all potential problems that this may bring about[26]—we continue to observe lower completion rates in the ND treatment. In sum, our evidence strongly suggests that period-t completion rates in the ND treatment remain persistently below those in the deadline treatments, which is at odds with our baseline model but consistent with the cognitive-limitations framework (Hypothesis 4').

RESULT 4: *(i) The period-t completion rate in the D1 treatment is higher [lower] than in the D3 treatment, for any period $t <$ [$>$] 2.5 weeks. (ii) The period-t completion rates in the D1 and D3 treatment are persistently higher than in the ND treatment.*

4.2 Follow-up Experiment

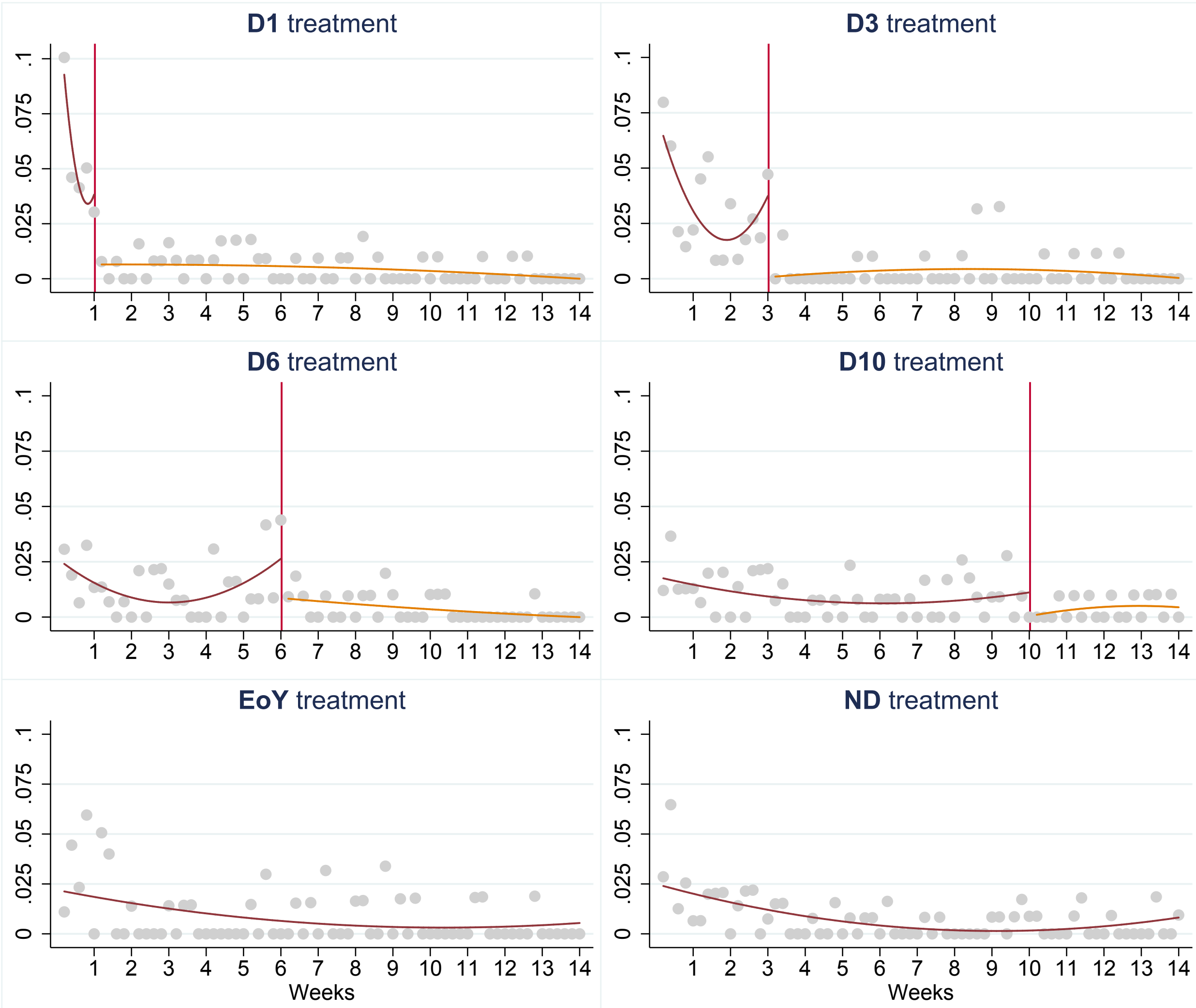
The findings from our main experiment are in conflict with a number of predictions from our baseline model. Instead, they support the notion that individuals exhibit limitations in memory and attention. In what follows, we study the robustness of these results and explore the impact of longer deadlines, based on data from our follow-up experiment.

Hazard Rates. Figure 4 depicts the raw hazard rates and quadratic fits for the pre- and post-deadline period in the six different treatments of the follow-up experiment. We observe a striking similarity between the shape of the hazard rates in the D1, D3, and ND treatment and the corresponding treatments from our main experiment (compare Figure 1).[27] For both the D1 and ND treatment, hazard rates are again decreasing over time. Furthermore, hazards in the D3 treatment exhibit the U-shaped pattern that we also observed for people facing the three-week deadline in our main experiment. A qualitatively similar picture emerges in the D6 and D10 treatment, though the U-shape is considerably flattened out under these longer deadlines. For people facing the ten-week deadline (D10), for example, we essentially observe no increase in hazard rates right before the deadline anymore. A further point worth noting in Figure 4 is the similarity of the declining hazard rates observed for the end-of-year deadline (which implied a deadline length of 5–7 months, see fn. 20) and the no-deadline treatment. This suggests that the no-deadline treatment can indeed be seen as a limit case of facing very long, explicitly specified deadlines.

[26]Recall from above that any analysis beyond the first 100 days is complicated by the fact that some participants are already due for the next check-up reminder again.

[27]The observed differences in response *levels* are hard to interpret, as the follow-up experiment differs from our main experiment in the reward level, sample, and implementation period.

Figure 4: Daily Hazard Rates (Follow-up Experiment)

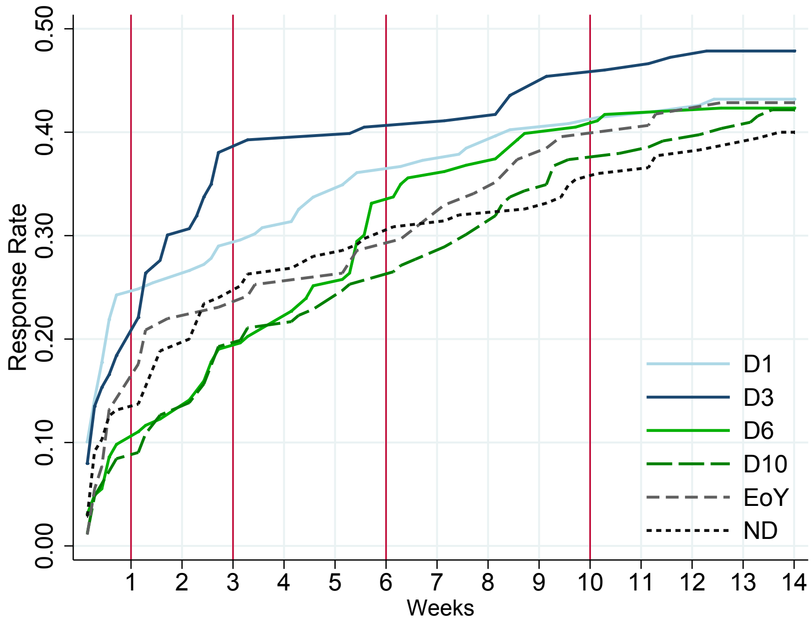


Notes: The figure displays daily hazard rates for all treatments of the follow-up experiment. The graph covers a period of 14 weeks after sending the mailings.

In sum, the evidence speaks against the baseline model's prediction of monotonically increasing hazard rates under finite deadlines and against constant hazard rates in environments with no (or extremely long) deadlines. These patterns, which corroborate Result 1 from above, are also confirmed by the results of a basic duration analysis (reported in Table B.3 in the Appendix). Most notably, the estimated shape parameter of the Weibull distribution is again significantly smaller than 1, reinforcing the notion of a declining baseline hazard in our follow-up experiment. In fact, the estimated coefficient is almost identical to the one from our main experiment ($\rho = 0.563$ in the follow-up and $\rho = 0.565$ in the main experiment; see Table 2). Further paralleling our results from Table 2 above, we also find that the estimated hazard ratios for the treatments with relatively short deadlines (D1 and D3 treatment) lie substantially above one. In contrast—and in line with the visual impression from Figure 4—hazards for the treatments with longer deadlines are rather similar to the baseline (i.e., no-deadline) case. We also estimated models that allow for time-varying treatment effects. The estimates for week-specific treatment differences indicate that hazard rates in the pre-deadline week are generally higher for shorter deadlines. A similar picture emerges when we compare hazards during the last pre-deadline days (see Table C.2 and Figure C.2 in the Online Appendix). Given the limited power of the follow-up experiment, however, these time-varying treatment effects (as well as some of the coefficients from the basic duration analysis) are estimated rather imprecisely. Yet, the data are qualitatively in line with Result 2 from above: hazard rates in a given pre-deadline period depend negatively on the length of the deadline.

Completion Rates. Figure 5 presents cumulative response rates across treatments in the follow-up experiment. The evolution of task-completion rates during the first three weeks follows a similar pattern as in the main experiment: under the D1 treatment, responses surge during the first week. In the following two weeks, the completion rate under D3 catches up and overtakes the D1 treatment before the end of the three-week deadline (see dark and light blue solid lines in Figure 5). After three weeks, task completion rates for the D3 and D1 treatment are substantially higher than the ones observed for treatments with longer (or no) deadlines (LPM estimates comparing period-t completion rates are provided in Table B.4 in the Appendix).

Figure 5: Cumulative Response Rates (Follow-up Experiment)



Notes: The figure presents the empirically observed cumulative response rates over 14 weeks.

During weeks 3–6, the completion rate in the D6 treatment (solid light green line) increases relative to the D1 and D3 treatment. However, this catch-up is incomplete such that response rates after week six are still higher in D1 and D3 as compared to the D6 treatment. A similar pattern is observed during weeks 7 to 10: the period-t completion rate in the D10 treatment (dashed dark green line) catches up but remains well below the levels observed for treatments with shorter deadlines (i.e., D1, D3 and D6). Both patterns are in conflict with Hypothesis 4 from the baseline model, lending further support to the relevance of cognitive limitations (Hypothesis 4'). After 100 days, completion rates in the D1, D6, and D10 treatment are virtually identical, but the completion rate in the D3 treatment remains about 6–9 percentage points above the values in the treatments involving longer or no deadlines.

Figure 5 also allows us to compare *within-deadline* completion rates across treatments. As in our main experiment (Result 3), we observe a higher fraction of pre-deadline responses in the D3 treatment than under the shorter deadline of the D1 treatment. However, the within-deadline completion rate in the D3 treatment also turns out to be higher than in the D6 treatment (38.0% vs 33.1%) and also marginally higher than the rate in D10 (37.3%). Put differently: the fraction of patients who respond within the three-week deadline of the D3 treatment is roughly 5 percentage point higher than the corresponding response rate for the twice as long time window in the D6 treatment. While the difference is not statistically significant, it suggests that longer deadlines may indeed diminish the likelihood of task completion within the deadline, as hypothesized in our cognitive-limitations model (Hypothesis 3').

In sum, the evidence from the follow-up experiment corroborates and further substantiates our findings from Section 4.1. Pre-deadline hazard rates tend to be U-shaped or decreasing, and do not only depend on how much time is left until the deadline, but also on how much time has passed since the beginning of the intervention. As a result, task completion rates can be higher under shorter deadlines, both in the short run as well as over longer time periods. Taken together, the evidence thus underlines the power of relatively short deadlines for accelerating and (potentially) boosting task completion.

4.3 Discussion

Deadlines and Incentives. Our analysis of the main experiment focused on treatment differences in the deadline dimension, while pooling across the different rewards employed in the experiment (see Section 3.1). To further explore if and how individuals' responses depend on the interplay between deadlines and incentives, we examine hazard and completion rates separately for all nine treatment cells in our 3×3-design (see Figures B.2 and B.3 in the Appendix).

The exercise reveals several points worth noting. First, the (conditional) response rates are highest in the large-reward treatments. At the same time, there are little systematic differences between the treatments with small and no explicit rewards. This order of treatments in the reward dimension—which details the estimated effects reported in Tables 2 and 4—holds across all deadline lengths. Second, the impact of deadlines as reflected in the discontinuity around the deadline is most pronounced for the high reward level. This is intuitive, as the individuals in this treatment face a greater 'loss' when missing the deadline, relative to the two other reward conditions. Third, there is no single treatment cell with monotonically increasing pre-deadline hazards rates: for all reward levels, hazard rates tend to be decreasing or exhibit a U-shaped pattern (the latter is most pronounced in the high-reward treatments). Forth, with small rewards, the period-t completion rates approach similar levels in the mid run. In the no- and high-reward treatments, however, we observe higher completion rates after 100 days under both deadlines as compared to the corresponding no-deadline condition. The fact that imposing a one- or three-week deadline triggers persistently higher response rates is particularly interesting for the large reward condition, as one would expect people to eventually make use of the high-value voucher in the no-deadline treatment—at least if they were still attentive.

Finally, the data suggest that deadlines affect behavior even if there is no explicit reward attached to them (see bottom panel of Figure B.3). From a practical perspective, the most simple and inexpensive treatment ('please call us within ... weeks') thus performs fairly well. Two possible explanations for this finding come to mind. First, deadlines in these treatments might be perceived as an informative signal of the (implicit) health rewards related to a check-up (e.g., signaling the dentist's perception of a patient's health status).[28] Second, deadlines in the no-reward treatments might function as a planning prompt (Rogers et al., 2015; Beshears et al., 2016), thereby amplifying the pure reminder effect from the postcards.

Survey Evidence. Complementing our evidence on the behavioral effects of deadlines, we conducted a large online survey experiment on individuals' perceptions of deadlines (N = 3,078) and a smaller post-experimental survey at the dental clinic (N = 273). For the latter survey (which was administered after the follow-up experiment), we invited patients in the waiting room of the clinic to participate in a short (approx. 5 min) structured face-to-face interview. Questions covered individuals' views on dental health prevention, appointment planning, and the perception of deadlines and check-up reminders. The online survey experiment was conducted in collaboration with a professional survey provider that maintains a sample that is representative for Germany's adult population. After a number of introductory questions, participants in the online survey were assigned to a vignette scenario in which they were shown one of the postcards from our experiment ('Imagine you receive the following postcard from your dentist...'). Across survey participants, we randomly varied the postcard texts, such that they corresponded to one of the treatment cells from the main or follow-up experiment. The survey then asked for individuals' perceptions of the respective postcard along various dimensions.

Tables C.3 and C.4 in the Online Appendix summarize participant characteristics and the main outcomes of the vignette experiment. Notably, participants in the online survey who were confronted with different deadline lengths do not differ systematically in how they perceive (i) the dentist's competence, (ii) his economic or (iii) benevolent intentions ('wants me not to postpone/forget'), or (iv) the potential presence of an acute health problem. Among respondents facing a deadline scenario, however, the fraction expressing that the deadline puts them under pressure when arranging a new appointment is higher, the shorter the deadline. We also asked survey participants about their views on deadlines in other domains of life. Almost all respondents (91%) classify deadlines as 'always' or 'mostly' helpful in avoiding problems related to 'postponing and forgetting'. A majority (75%) further indicate that they generally prefer relatively tight deadlines.

[28]The evidence from our vignette survey, which is discussed in more detail in the next paragraph, is inconclusive on this point. After being exposed to a no-reward scenario, survey participants who face a deadline are more likely to state that the postcard may indicate a potential health problem (as compared to those in the corresponding ND scenario). The difference is, however, not statistically significant.

The survey at the dental clinic yields a similar picture. Less than 5% of the surveyed patients think that a deadline indicates an acute dental health problem. The vast majority (95%) rather interprets the deadlines as a tool to prevent them from forgetting or postponing to arrange a check-up. We also asked participants which deadline length they would perceive as ideal in the check-up context. Interestingly, a majority of participants (78%) prefer a deadline of three weeks or less, whereas only 4% prefer deadlines of six weeks or more. These preferences happen to match the relative 'success' of treatments in terms of task completion rates in our follow-up experiment quite well (see Figure 5), suggesting that individuals may at least be partially sophisticated about their cognitive limitations.[29]

Type Heterogeneity. A potential concern with our analysis of hazard rate dynamics is that it does not account for possible heterogeneity in patient 'types' (see, e.g., Salant 1977 and, for a more recent discussion, Heffetz et al. 2016). If patients differ fundamentally—say, in their opportunity costs of arranging check-ups—our interpretation of decreasing hazards may be misleading, as it abstracts from differential sorting of patients who are more or less responsive at different points in time. To examine the relevance of type heterogeneity in our context, we first explore whether there are systematic differences in the shape of hazard rates for groups of patients who differ in terms of sociodemographics and other individual-level characteristics. Using these variables to split our sample in subgroups, we observe only small differences in the level and shape of hazard rates across groups (see Figures C.3–C.5 in the Online Appendix). In particular, none of the subgroups shows monotonically increasing pre-deadline hazards; instead, hazard rates are decreasing or U-shaped for all subgroups and all treatments.

In a second step, we exploit information on systematic differences in response propensities, using data on patients' reactions to reminder postcards in the period prior to our intervention. We draw on data from an earlier study which evaluated the basic impact of reminders at the same dental clinic (Altmann and Traxler 2014; this study involved random variation in the framing of reminder messages but no deadlines or economic incentives). To predict which 'types' have a particularly high propensity to quickly respond to our intervention, we first work with the sample from this earlier period. In doing so, we estimate the probability that a patient contacts the dentist within one week after receiving a reminder, using patients' observable characteristics as well as information about the precise timing of sending waves as explanatory variables. Based on this estimation, we then predict the propensity that a patient-treatment observation in our (main) experiment yields a 'quick' response.[30] Finally, we examine hazard rates separately for the quartile of patients

[29]For a more extensive discussion and empirical tests of the degree to which individuals correctly perceive their own limitations in memory and attention, see also Ericson (2011) and Haushofer (2015).

[30]More specifically, we fit a logit model using all background variables available for the full sample, indicators for sending weeks within a month and dummies for different sending quarters of the year on a sample of $N = 1,176$ patient-postcard observations. (We also include data from a short non-experimental phase, in which all patients received one postcard from Altmann and Traxler (2014).) To assess the robustness of our predictions, we also studied a LASSO logistic model (using the R-package *glmnet*). The shrunken model yields very similar predictions and split-sample results as those reported below. The same holds for using alternative estimation models (e.g., probit) or outcome variables (e.g., response within two weeks).

with the highest predicted probability of responding quickly and the remaining group of patients with lower immediate-response propensities.

Figure B.4 in the Appendix depicts the raw daily hazard rates and quadratic fits for the two groups of patients. In line with our prediction approach, hazard rates are higher for the high-propensity sample, in particular during the first week of the experiment. The level differences, however, tend to be relatively modest. Moreover, for both groups of patients, the pre-deadline hazards are again decreasing (D1 treatment) or U-shaped (D3 treatment).[31] Naturally, these findings do not preclude that there might still exist relevant residual heterogeneity within the subgroups. At the very least, however, they indicate that the observed shape of responses under different deadlines is remarkably similar for different strata of our sample.

The latter point also speaks to the question of who does (and who does not) 'benefit' from facing a deadline. To further assess this question, we examine cumulative response rates among different subgroups of our sample. The data indicate that imposing a deadline triggers a relatively *larger* increase in task completion for groups with *lower* response rates in the no-deadline environment. This point is illustrated in Figure B.5 in the Appendix, which again compares the subgroups with the highest vs. lower predicted propensities of responding quickly. For the former group, the completion rate after 100 days is 5.6% higher under the three-week deadline than in the no-deadline environment (with response rates of 48.9% vs. 46.3%, respectively). For the group with lower propensities to respond, the corresponding increase is 13.6% (response rates of 35.2% vs. 31.0%; see Figure B.5). Hence, deadlines seem particularly well suited for inducing task completion among subgroups with a lower baseline propensity to complete the task at hand.

5 Conclusions

In this paper, we studied the interplay of deadlines and cognitive limitations in intertemporal decisions. The results from a simple theoretical model, two field experiments, and a complementary survey study document an important role of cognitive limitations in individuals' reactions to deadlines. We find that imposing a deadline leads to persistently higher task completion rates than having no deadline at all. Furthermore, our evidence suggests that relatively tight deadlines can trigger both earlier as well as overall higher response rates than longer ones. In this sense, short deadlines can yield a 'double dividend'. On a more general account, our findings suggests that deadlines can be a powerful policy tool, but that regulators and policymakers need to be careful in assessing what constitutes 'good' deadlines. If, for instance, a company offers seemingly generous extensions of consumer rights to its customers—e.g., very long deadlines for cancellations or product returns—this offer might ultimately *not* be consumer friendly, at least for individuals with cognitive limitations.

[31] A similar picture emerges when we consider other subgroups of patients, e.g., the quartile with the lowest predicted propensity to quickly respond or the 'middle 50%'. Unsurprisingly, differences in hazard rates are least pronounced at the 'lower end' of the propensity distribution.

References

ALTMANN, S. AND C. TRAXLER (2014): "Nudges at the Dentist," *European Economic Review*, 72, 19–38.

APESTEGUIA, J., P. FUNK, AND N. IRIBERRI (2013): "Promoting Rule Compliance in Daily-Life: Evidence from a Randomized Field Experiment in the Public Libraries of Barcelona," *European Economic Review*, 64, 266–284.

ARIELY, D. AND K. WERTENBROCH (2002): "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 13, 219–224.

BERTRAND, M., D. KARLAN, S. MULLAINATHAN, E. SHAFIR, AND J. ZINMAN (2010): "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment," *Quarterly Journal of Economics*, 125, 263–306.

BESHEARS, J., K. L. MILKMAN, AND J. SCHWARTZSTEIN (2016): "Beyond Beta-Delta: The Emerging Economics of Personal Plans," *American Economic Review*, 106, 430–434.

BHARGAVA, S. AND D. MANOLI (2015): "Psychological Frictions and the Incomplete Take-up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review*, 105, 3489–3529.

BISIN, A. AND K. HYNDMAN (2014): "Present-bias, Procrastination and Deadlines in a Field Experiment," NBER Working Paper No. 19874.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2012): "Salience Theory of Choice under Risk," *Quarterly Journal of Economics*, 127, 1243–1285.

——— (2013): "Salience and Consumer Choice," *Journal of Political Economy*, 121, 803–843.

CALZOLARI, G. AND M. NARDOTTO (2017): "Effective Reminders," *Management Science*, forthcoming.

CAMPBELL, A., F. EDERER, AND J. SPINNEWIJN (2014): "Delay and Deadlines: Freeriding and Information Revelation in Partnerships," *American Economic Journal: Microeconomics*, 6, 163–204.

CARVALHO, L. S., S. MEIER, AND S. W. WANG (2016): "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday," *American Economic Review*, 106, 260–284.

CHETTY, R., J. N. FRIEDMAN, S. LETH-PETERSEN, T. H. NIELSEN, AND T. OLSEN (2014a): "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark," *Quarterly Journal of Economics*, 129, 1141–1219.

CHETTY, R., A. LOONEY, AND K. KROFT (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–1177.

CHETTY, R., E. SAEZ, AND L. SANDOR (2014b): "What Policies Increase Prosocial Behavior? An Experiment with Referees at the Journal of Public Economics," *Journal of Economic Perspectives*, 28, 169–188.

DAMGAARD, M. T. AND C. GRAVERT (2017): "Now or Never! The Effect of Deadlines on Charitable Giving: Evidence from Two Natural Field Experiments," *Journal of Behavioral and Experimental Economics*, 66, 78–87.

ENGLMAIER, F., A. SCHMÖLLER, AND T. STOWASSER (2017): "Price Discontinuities in an Online Market for Used Cars," *Management Science*, forthcoming.

ERICSON, K. M. (2011): "Forgetting we Forget: Overconfidence and Memory," *Journal of the European Economic Association*, 9, 43–60.

——— (2017): "On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation," *Journal of the European Economic Association*, 15, 692–719.

HAUSHOFER, J. (2015): "The Cost of Keeping Track," Working Paper, Princeton University.

HEFFETZ, O., T. O'DONOGHUE, AND H. S. SCHNEIDER (2016): "Forgetting and Heterogeneity in Task Delay: Evidence from New York City Parking-Ticket Recipients," NBER Working Paper No. 23012.

HEISS, F., D. MCFADDEN, J. WINTER, A. WUPPERMANN, AND B. ZHOU (2016): "Inattention and Switching Costs as Sources of Inertia in Medicare Part D," NBER Working Paper No. 22765.

HERWEG, F. AND D. MÜLLER (2011): "Performance of Procrastinators: On the Value of Deadlines," *Theory and Decision*, 70, 329–366.

HOLMAN, J. AND F. ZAIDI (2010): "The Economics of Prospective Memory," Working Paper, UC Berkeley.

HOSSAIN, T. AND J. MORGAN (2006): "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," *Advances in Economic Analysis & Policy*, 5.

KAHNEMAN, D. (2003): "Maps of Bounded Rationality: Psychology for Behavioral Economics," *American Economic Review*, 93, 1449–1475.

KARLAN, D., M. MCCONNELL, S. MULLAINATHAN, AND J. ZINMAN (2016): "Getting to the Top of Mind: How Reminders Increase Saving," *Management Science*, 62, 3393–3411.

KŐSZEGI, B. AND A. SZEIDL (2013): "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics*, 128, 53–107.

LACETERA, N., D. G. POPE, AND J. R. SYDNOR (2012): "Heuristic Thinking and Limited Attention in the Car Market," *American Economic Review*, 102, 2206–2236.

LAIBSON, D. I. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443–478.

LANG, W. P., M. M. FARGHALY, AND D. L. RONIS (1994): "The Relation of Preventive Dental Behaviors to Periodontal Health Status," *Journal of Clinical Periodontology*, 21, 194–198.

MANI, A., S. MULLAINATHAN, E. SHAFIR, AND J. ZHAO (2013): "Poverty Impedes Cognitive Function," *Science*, 341, 976–980.

McCall, J. J. (1970): "Economics of Information and Job Search," *Quarterly Journal of Economics*, 84, 113–126.

Mortensen, D. T. (1970): "Job Search, the Duration of Unemployment, and the Phillips Curve," *American Economic Review*, 60, 847–862.

Mullainathan, S. and E. Shafir (2013): *Scarcity: Why Having Too Little Means So Much*, New York, NY: Time Books, Henry Holt & Company LLC.

O'Donoghue, T. and M. Rabin (1999): "Incentives for Procrastinators," *Quarterly Journal of Economics*, 114, 769–816.

Rees-Jones, A. and D. Taubinsky (2016): "Tax Psychology and the Timing of Charitable-Giving Deadlines," The Urban Institute.

Rogers, T., K. L. Milkman, L. K. John, and M. I. Norton (2015): "Beyond Good Intentions: Prompting People to Make Plans Improves Follow-through on Important Tasks," *Behavioral Science and Policy*, 1, 33–41.

Saez-Marti, M. and A. Sjögren (2008): "Deadlines and Distractions," *Journal of Economic Theory*, 143, 153–176.

Salant, S. W. (1977): "Search Theory and Duration Data: A Theory of Sorts," *Quarterly Journal of Economics*, 91, 39–57.

Schilbach, F., H. Schofield, and S. Mullainathan (2016): "The Psychological Lives of the Poor," *American Economic Review*, 106, 435–440.

Shu, S. B. and A. Gneezy (2010): "Procrastination of Enjoyable Experiences," *Journal of Marketing Research*, 47, 933–944.

Simon, H. A. (1955): "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 69, 99–118.

Stigler, G. J. (1961): "The Economics of Information," *Journal of Political Economy*, 69, 213–225.

Sueyoshi, G. T. (1995): "A Class of Binary Response Models for Grouped Duration Data," *Journal of Applied Econometrics*, 10, 411–431.

Taubinsky, D. (2014): "From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice," Working Paper, Harvard University.

Taubinsky, D. and A. Rees-Jones (2017): "Attention Variation and Welfare: Theory and Ecidence from a Tax Salience Experiment," *Review of Economic Studies*, forthcoming.

Wälde, K. (2015): "Stress and Coping – An Economic Approach," Discussion Paper, University of Mainz.

Weinschenk, P. (2016): "Procrastination in Teams and Contract Design," *Games and Economic Behavior*, 98, 264–283.

A Appendix to Section 2

A.1 Option Values over Time: Inequality (3)

Combining (1) and (2) yields

$$V_t = \int_0^{\hat{c}_t} (\hat{c}_T - c) dF(c) + \delta V_{t+1}. \tag{23}$$

Since the game ends at the latest after period T, the option value of reaching a period after T is zero: $V_t = 0$ for all $t > T$. Hence, the cost threshold in period T is $\hat{c}_T = y > 0$ (see equation (1)). From (23) we thus see that $V_T > 0$, i.e., the option value decreases from period T to period $T+1$.

Next, consider the remaining periods. Because the agent chooses the cost threshold in every period t to maximize V_t, the Envelope Theorem together with (2) implies that the agent's current option value V_t benefits from a higher future option value V_{t+1}:

$$\frac{dV_t}{dV_{t+1}} = \frac{\partial V_t}{\partial \hat{c}_t} \cdot \frac{\partial \hat{c}_t}{\partial V_{t+1}} + \frac{\partial V_t}{\partial V_{t+1}} = \frac{\partial V_t}{\partial V_{t+1}} = (1 - F(\hat{c}_t))\delta > 0. \tag{24}$$

From equations (1) and (23) we see that if in period $T-1$ the option value would be the same as in period T, $V_T = V_{T+1}$, the current option values would be the same, too, $V_{T-1} = V_T$. But from above we know that $V_T > V_{T+1}$. Together with (24), this implies that $V_{T-1} > V_T$. Repeating these arguments for earlier periods yields that the option value decreases over time, as stated in inequality (3).

A.2 Upper Bound and Limit Behavior of Option Values: Equation (4)

We first prove a result concerning the cost thresholds: in all periods $t \leq T$, under the optimal decision rule, $F(\hat{c}_t) \in (0, 1)$. That is, neither $F(\hat{c}_t) = 0$, nor $F(\hat{c}_t) = 1$ can hold. To prove this, note that if $F(\hat{c}_t) = 0$, then $\hat{c}_t = 0$. Due to (2), it then holds that $V_t = \delta V_{t+1} \leq V_{t+1}$, which contradicts inequality (3). If $F(\hat{c}_t) = 1$, then $\hat{c}_t > y$. Due to (1), however, this requires that $V_{t+1} < 0$, which cannot hold since inequality (3) and $V_{T+1} = 0$ imply that all option values are non-negative.

Next we seek to find an upper bound for the option value. Consider some period $t \leq T$. By (23),

$$V_t - V_{t+1} = \int_0^{\hat{c}_t} (\hat{c}_T - c) dF(c) - (1 - \delta) V_{t+1}. \tag{25}$$

Using (1) and (25), let $\bar{V}$ solve

$$0 = \int_0^{y - \delta \bar{V}} (y - \delta \bar{V} - c) dF(c) - (1 - \delta) \bar{V}. \tag{26}$$

By the Intermediate Value Theorem and monotonicity, a unique solution $\bar{V}$ exists and $\bar{V} \in (0, y)$. The latter result implies that $F(\hat{c}_t) \in (0, 1)$ also holds in case of no deadline, $T \to \infty$. Recall that $V_{T+1} = 0$ and note that by (24)

$$\frac{d(V_t - V_{t+1})}{dV_{t+1}} = (1 - F(\hat{c}_t))\delta - 1, \tag{27}$$

which is, due to $F(\hat{c}_t) \in (0, 1)$, strictly between -1 and $\delta - 1$. Therefore, if we start at period $T+1$ and then go back in time, the option value V_t increases, approaches $\bar{V}$, but never exceeds $\bar{V}$. This property is summarized in equation (4).

A.3 Period-t Completion Rates without Cognitive Limitations: Equation (11)

Consider an agent without cognitive limitations and an arbitrary pair of deadlines T' and T'', with $T' < T''$. To see that the period-t completion rates are initially higher under a short deadline T' than under a long deadline T'', first note that for all periods $t \leq T'$, the hazard rate is higher with a short deadline T' than with the long deadline T''. Formally, by (6) and (8), it holds that $h_t|_{T=T'} < h_t|_{T=T''}$. Hence, according to definition (9),

$$p(t)|_{T=T'} > p(t)|_{T=T''} \quad \text{for all} \quad t \leq T'. \tag{28}$$

For periods $t \geq T''$, the period-t completion rates equal the within-deadline completion rates, i.e., $p(t)|_{T=T'} = p(T')$ and $p(t)|_{T=T''} = p(T'')$. From inequality (10), we know that $p(T') < p(T'')$ such that

$$p(t)|_{T=T'} < p(t)|_{T=T''} \quad \text{for all} \quad t \geq T''. \tag{29}$$

Consider finally the periods $t \in (T', T'')$. For the short deadline T', the period-t completion rate is equal to the within-deadline completion rate: $p(t)|_{T=T'} = p(T')$. The period-t completion rate $p(t)|_{T=T'}$ is hence the same for all $t \in (T', T'')$. In contrast, for the long deadline T'', the period-t completion rate $p(t)|_{T=T''} = 1 - \prod_{s=1}^{t}(1-h_s)$ is increasing over time $t \in (T', T'')$ since, see Appendix A.2, $h_s > 0$ for all $s \in [1, T'']$. Together with (28) and (29) this implies that, as stated in equation (11), there exists a threshold $\tau \in (T', T'')$ such that $p(t)|_{T=T'} > p(t)|_{T=T''}$ for all periods $t < \tau$ and $p(t)|_{T=T'} < p(t)|_{T=T''}$ for all periods $t > \tau$.

A.4 Hazard Rates over Time with Cognitive Limitations

The terms in equation (15) that capture the cognitive-limitation effect are

$$(1-\gamma)(1-F(\hat{c}_t))\frac{1-\sum_{s=1}^{t-1}F(\hat{c}_s)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r))}{1-\sum_{s=1}^{t}F(\hat{c}_s)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r))}. \tag{30}$$

Recognize that by (1) and (3) the fraction that captures the option-value effect, $F(\hat{c}_{t+1})/F(\hat{c}_t)$, is strictly greater than one and finite, such that the reciprocal, $F(\hat{c}_t)/F(\hat{c}_{t+1})$, is strictly between zero and one. Since the terms that capture the cognitive-limitation effect are decreasing in γ, approach one for $\gamma \to 0$ and zero for $\gamma \to 1$, by the Intermediate Value Theorem, there exists a unique threshold $\hat{\gamma}_t \in (0,1)$ such that the terms capturing the cognitive-limitation effect fall short of $F(\hat{c}_t)/F(\hat{c}_{t+1})$ if $\gamma > \hat{\gamma}_t$, are equal to $F(\hat{c}_t)/F(\hat{c}_{t+1})$ if $\gamma = \hat{\gamma}_t$, and exceed $F(\hat{c}_t)/F(\hat{c}_{t+1})$ if $\gamma < \hat{\gamma}_t$. Together with equation (15) this implies that

$$\begin{aligned} \tilde{h}_{t+1} &< \tilde{h}_t \ \text{ if } \ \gamma > \hat{\gamma}_t, \\ \tilde{h}_{t+1} &= \tilde{h}_t \ \text{ if } \ \gamma = \hat{\gamma}_t, \\ \tilde{h}_{t+1} &> \tilde{h}_t \ \text{ if } \ \gamma < \hat{\gamma}_t. \end{aligned} \tag{31}$$

A.5 Hazard Rates and Distance to the Deadline with Cognitive Limitations

Let us now compare the hazard rate of period t in case of some deadline T' to that of period $t+1$ in case of the deadline $T'+1$. From equation (13) we obtain

$$\tilde{h}_t\Big|_{T=T'} = \frac{F(\hat{c}_t|_{T=T'})(1-\gamma)^{t-1}\prod_{s=1}^{t-1}(1-F(\hat{c}_s|_{T=T'}))}{1-\sum_{s=1}^{t-1}F(\hat{c}_s|_{T=T'})(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r|_{T=T'}))}, \tag{32}$$

$$\tilde{h}_{t+1}\Big|_{T=T'+1} = \frac{F(\hat{c}_{t+1}|_{T=T'+1})(1-\gamma)^{t}\prod_{s=1}^{t}(1-F(\hat{c}_s|_{T=T'+1}))}{1-\sum_{s=1}^{t}F(\hat{c}_s|_{T=T'+1})(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r|_{T=T'+1}))}. \tag{33}$$

Using (1) and (7), it is readily verified that

$$\tilde{h}_t\Big|_{T=T'} > \tilde{h}_{t+1}\Big|_{T=T'+1} \quad \text{for all } \ T' \in \mathbb{N}. \tag{34}$$

Repeating these arguments leads to the result stated in inequality (16).

A.6 Sophistication about Cognitive Limitations

The hazard rate in period t is

$$\tilde{h}_t^S = \frac{F(\hat{c}_t^S)(1-\gamma)^{t-1}\prod_{s=1}^{t-1}(1-F(\hat{c}_s^S))}{1-\sum_{s=1}^{t-1}F(\hat{c}_s^S)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r^S))}. \tag{35}$$

Under the decision rule (17), the option value—given that the task is still on the top of the agent's mind—is

$$V_t^S = \int_0^{\hat{c}_t^S}(y-c)dF(c) + \left(1-F\left(\hat{c}_t^S\right)\right)(1-\gamma)\delta V_{t+1}^S. \tag{36}$$

To explore how the hazard rate evolves over time in case of a deadline, we need to adjust the hazard ratio from equation (14):

$$\frac{\tilde{h}_{t+1}^S}{\tilde{h}_t^S} = \frac{F(\hat{c}_{t+1}^S)}{F(\hat{c}_t^S)}(1-\gamma)(1-F(\hat{c}_t^S))\frac{1-\sum_{s=1}^{t-1}F(\hat{c}_s^S)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r^S))}{1-\sum_{s=1}^{t}F(\hat{c}_s^S)(1-\gamma)^{s-1}\prod_{r=1}^{s-1}(1-F(\hat{c}_r^S))}. \tag{37}$$

As with naivete, the hazard rate is higher in period t than in $t+1$ if the cognitive-limitation effect dominates the option-value effect, which holds true whenever the cognitive-limitation effect is sufficiently strong, i.e., when γ is sufficiently high.

Consider next the case without a deadline, $T \to \infty$. Since the environment is time-invariant, the option value is time-invariant, too. The option value solves

$$0 = \int_0^{y-\delta\bar{V}^S}(y-\delta\bar{V}^S-c)dF(c) - (1-\delta(1-\gamma))\bar{V}^S. \tag{38}$$

By the Intermediate Value Theorem and monotonicity, a unique solution $\bar{V}^S$ exists and $\bar{V}^S \in (0,y)$. The option-value effect is hence absent and the cognitive-limitation effect always dominates, which is why the hazard rate is decreasing over time. Note that this is qualitatively the same result as we obtained in case of naivete, but the option value in case of sophistication, $\bar{V}^S$, is different than the one with naivete, $\bar{V}$.

We now show that—in accordance with the case of naivete—the hazard rate depends negatively on the absolute value of time, holding the time distance to the deadline fixed. Note that by equations (35), (36), and

$V_{T+1} = 0$ it holds that

$$V_t^S\big|_{T=T'} = V_{t+s}^S\big|_{T=T'+s} \quad \text{for all} \;\; T' \in \mathbb{N} \;\; \text{and} \;\; s \in \mathbb{N}. \tag{39}$$

We can hence apply the formal arguments we used for the case with a naive agent to clarify that

$$\tilde{h}_t^S\Big|_{T=T'} > \tilde{h}_{t+s}^S\Big|_{T=T'+s} \quad \text{for all} \;\; T' \in \mathbb{N} \;\; \text{and} \;\; s \in \mathbb{N}. \tag{40}$$

We finally show that—as in case with naivete—the within-deadline completion rate can be decreasing in the length of the deadline. It is sufficient to consider the case where $y \to \bar{k}$ and the deadline is either $T' = 1$ or $T'' = 2$. With the deadline $T' = 1$, the within-deadline completion rate is

$$p(T') = \tilde{h}_1^S\Big|_{T'=1} = F(y) = 1. \tag{41}$$

In contrast, with the deadline $T'' = 2$ the within-deadline completion rate is

$$\begin{aligned} p(T'') &= 1 - \left(1 - \tilde{h}_1^S\Big|_{T''=2}\right)\left(1 - \tilde{h}_2^S\Big|_{T''=2}\right) \\ &= 1 - \left(1 - \tilde{h}_1^S\Big|_{T''=2}\right)\gamma. \end{aligned} \tag{42}$$

Note that since the option value $V_2^S > 0$ for all $\gamma \in (0, 1)$, it holds that $\hat{c}_1^S\big|_{T''=2} < y$ and thus that $\tilde{h}_1^S\Big|_{T''=2} = F\left(\hat{c}_1^S\big|_{T''=2}\right) < 1$ and $p(T'') < 1$. The within-deadline completion rate is hence higher with the short than with the long deadline, $p(T') > p(T'')$. As in case of a naivete, the long deadline then never catches up the short deadline in terms of task completion rates, i.e., the period-t completion rates are permanently lower with the long than with the short deadline.

A.7 Task Completion after the Deadline

We next show that our framework naturally extends to environments where the agent can still complete the task after the deadline, and where doing so yields a strictly positive, but smaller reward than completion before the deadline.[32] Formally, let the agent receive a reward of $\bar{y}$ if she completes the task before the deadline and $\underline{y}$ if she does so afterwards, where $\bar{y} \geq \underline{y} > 0$. In what follows we show that Hypotheses 1–4 are also valid for this alternative environment.

We are interested in the structure of the cost thresholds the agent uses, which in turn depend on the structure of the option values. Suppose initially that $\bar{y} = \underline{y}$, such that the agent uses a constant cost threshold for all periods, resulting in hazard rates that are also constant over time. Suppose now that $\bar{y}$ increases, whereas $\underline{y}$ remains the same as before. This leaves the cost thresholds and, consequently, the hazard rates after the deadline T constant. But how do the cost thresholds and hazard rates before the deadline change? From equation (2) and the Envelope Theorem it follows that

$$\frac{dV_T}{d\bar{y}} = F(\hat{c}_T), \tag{43}$$

[32] Note that environments where the reward is the same for completion before or after the 'deadline' are isomorphic to the case without a deadline, $T \to \infty$.

which is strictly between zero and one. Using this insight and again equation (2) and the Envelope Theorem, we get that

$$\frac{dV_{T-1}}{d\bar{y}} = F(\hat{c}_{T-1}) + (1 - F(\hat{c}_{T-1}))\delta \frac{dV_T}{d\bar{y}} = F(\hat{c}_{T-1}) + (1 - F(\hat{c}_{T-1}))\delta F(\hat{c}_T). \tag{44}$$

This is again strictly between zero and one. Repeating the arguments leads to the insight that

$$\frac{dV_t}{d\bar{y}} \in (0,1) \text{ for all } t \in \{1, ..., T\}. \tag{45}$$

By equation (1) it thus holds that

$$\frac{d\hat{c}_t}{d\bar{y}} > 0 \text{ for all } t \in \{1, ..., T\}. \tag{46}$$

Since by (5) the hazard rate is $h_t = F(\hat{c}_t)$, it directly follows that for $\bar{y} > \underline{y}$

$$h_t > h_s \text{ for all } t \in \{1, ..., T\} \text{ and } s \geq T + 1. \tag{47}$$

That is, if the agent is incentivized to complete the task before the deadline—such that the rewards for task completion before the deadline are strictly higher than after the deadline—the hazard rates after the deadline are lower than all hazard rates within the deadline.

We can derive further insights. In Appendix A.1, we explained that if $V_T > V_{T+1}$, then the option value decreases over time: $V_t > V_{t+1}$ for all $t \in \{1, ..., T\}$. By (43) this is also true in the environment we currently explore. Hence, it still holds that $V_t > V_{t+1}$ for all $t \in \{1, ..., T\}$. By (1) and (5) it thus follows that the hazard rate is increasing over time until period T,

$$h_t < h_{t+1} \text{ for all } t \in \{1, ..., T-1\}. \tag{48}$$

Moreover, by the same arguments as in the baseline model, the hazard rates depend only on the time left until the deadline, but not on the absolute value of time,

$$h_t|_{T=T'} = h_{t+s}|_{T=T'+s} \text{ for all } T' \in \mathbb{N} \text{ and } s \in \mathbb{N}. \tag{49}$$

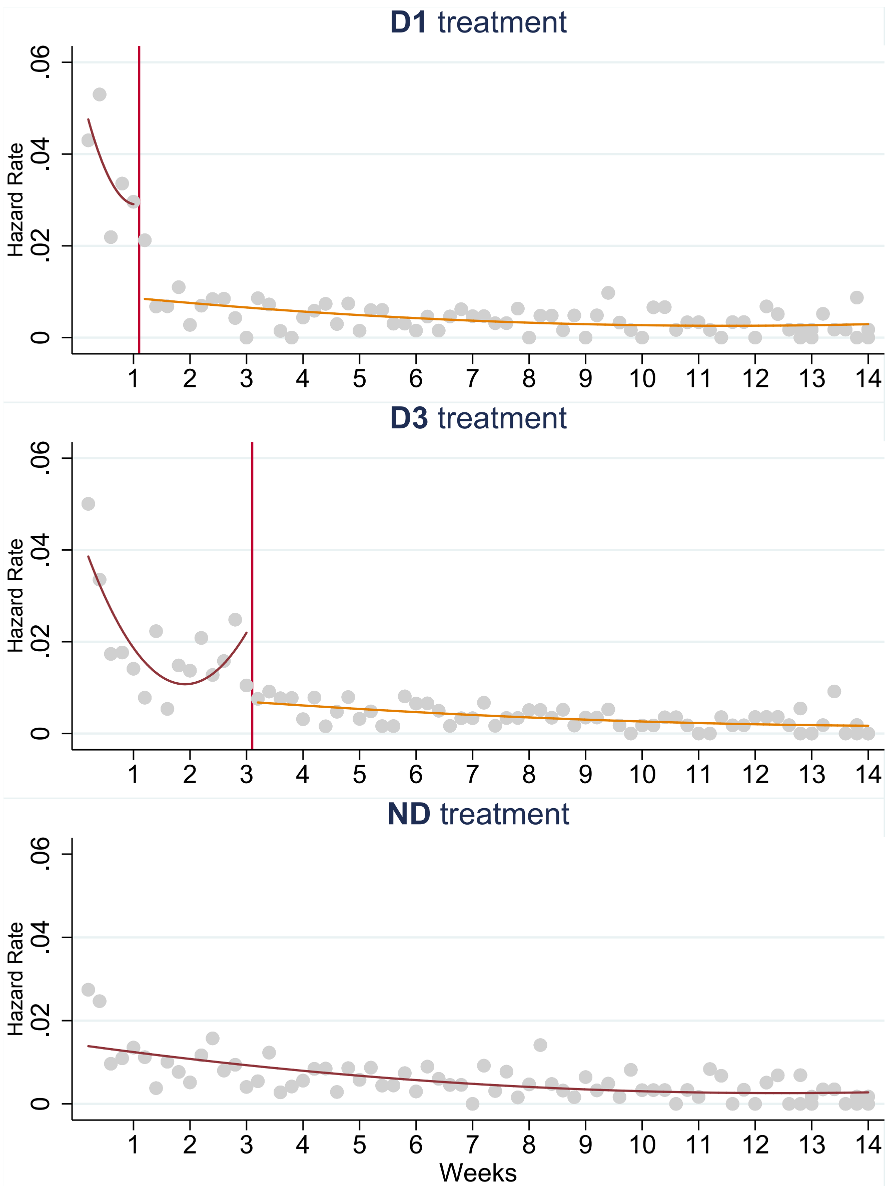
We can hence conclude that Hypotheses 1 and 2 still hold when task completion is still possible after the deadline.

It remains to be shown that Hypotheses 3 and 4 also hold in this environment. Since hazard rates are positive and only depend on the time left until the deadline, but not on the absolute value of time, it directly follows that the within-deadline completion rate is still increasing in the length of the deadline. In a final step, we show that our results for period-t completion rates (equation 11) also hold when task completion after the deadline is possible. Consider some arbitrary pair of deadlines T' and T'', with $T' < T''$. First, by (48) and (49), the hazard rates in all periods $t \leq T'$ are lower with the long deadline T'' than with the short deadline T', such that by (9) the period-t completion rates in all periods $t \leq T'$ satisfy $p(t)|_{T=T'} > p(t)_{T=T''}$. Second, since the hazard rates in the last T' periods before the deadline are the same for both deadlines T' and T'' (see condition (49)), and the hazard rates after the deadline T' are lower than the hazard rates in the first $T'' - T'$ periods in case of deadline T'' (see condition 47), definition (9) implies that the completion rate at period T'' is higher for the long than for the short deadline: $p(T'')|_{T=T'} < p(T'')_{T=T''}$. Finally, since the hazard rates under the long deadline T'' are higher for all periods between T' and T'' than those under the short deadline T' (see condition 47), there must exist a threshold period $\bar{\tau} \in (T', T'')$ such that $p(t)|_{T=T'} > p(t)_{T=T''}$ for all $t < \bar{\tau}$ and $p(t)|_{T=T'} < p(t)_{T=T''}$ for all $t > \bar{\tau}$.

B. Additional Figures and Tables (Appendix to Section 4)

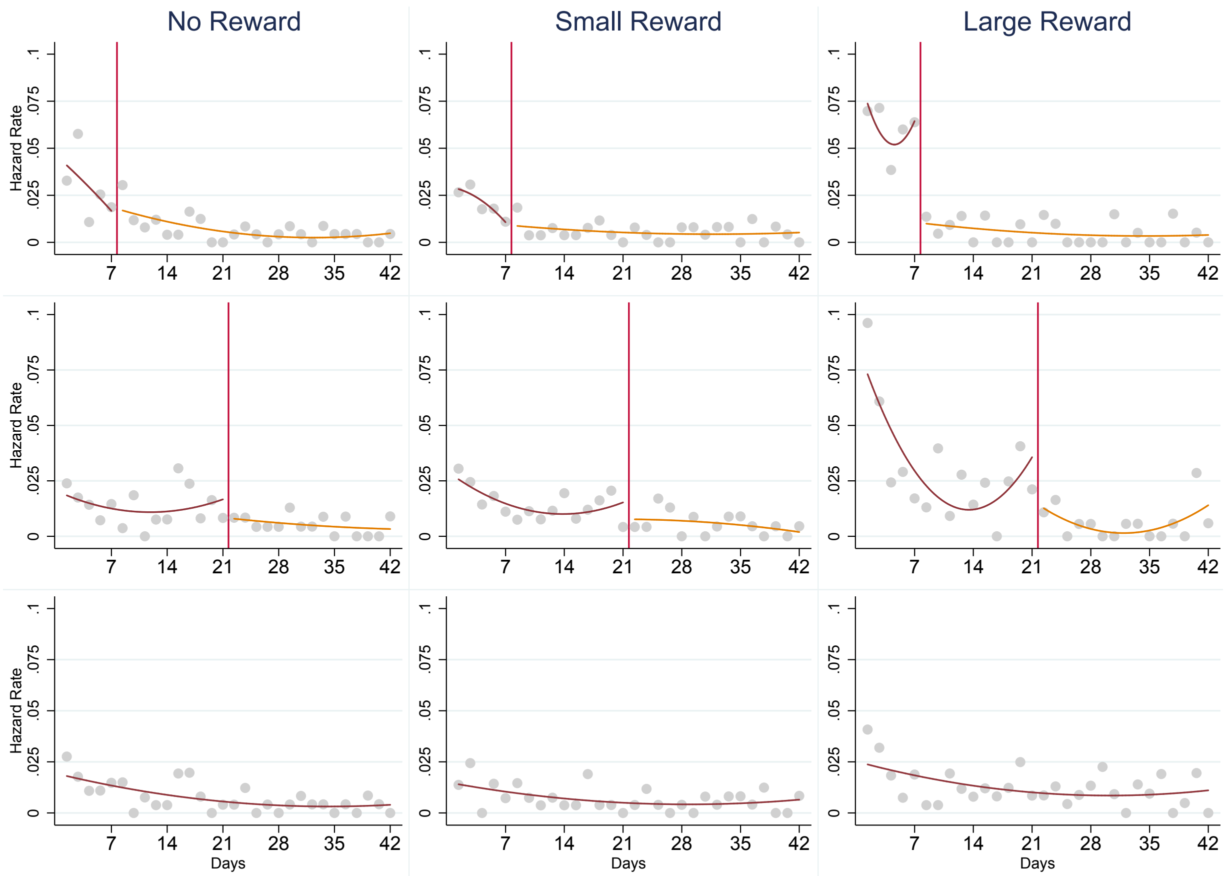
B.1 Figures

Figure B.1: Daily Hazard Rates over 14 weeks (Main Experiment)



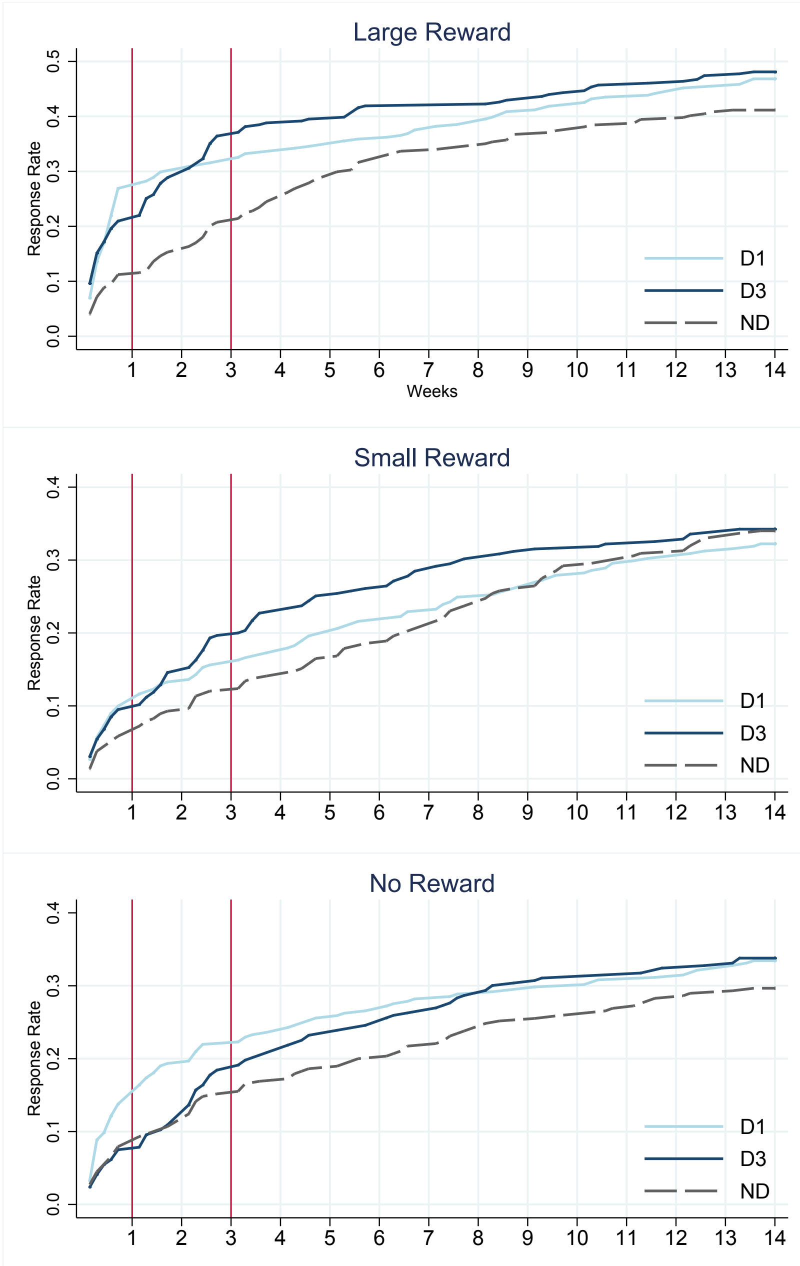
Notes: The figure displays daily hazard rates for the treatments with a one-week deadline (top panel), three-week deadline (middle panel) and no deadline (bottom panel) over a period of 14 weeks after sending the mailings.

Figure B.2: Daily Hazard Rates for all 3 × 3 Treatment Cells of the Main Experiment



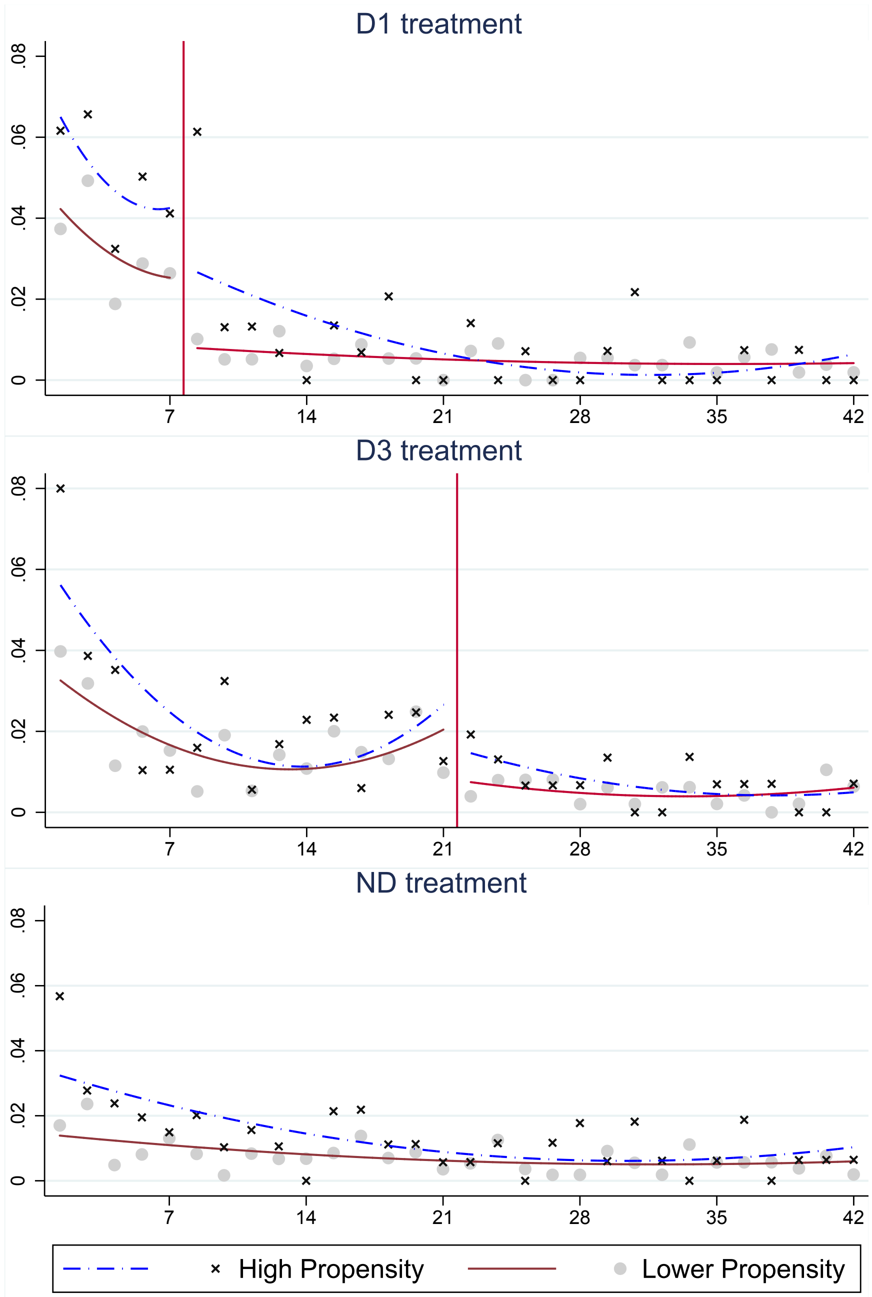
Notes: The figure displays daily hazard rates over a period of 6 weeks after sending the mailings, separately for each of the nine combinations of deadlines and rewards from our main experiment. The left/middle/right column presents hazard rates for the treatments with no/small/large-Rewards, and the first/second/third row contains D1/D3/ND treatments, respectively.

Figure B.3: Cumulative Response Rates for all 3×3 Treatment Cells of the Main Experiment



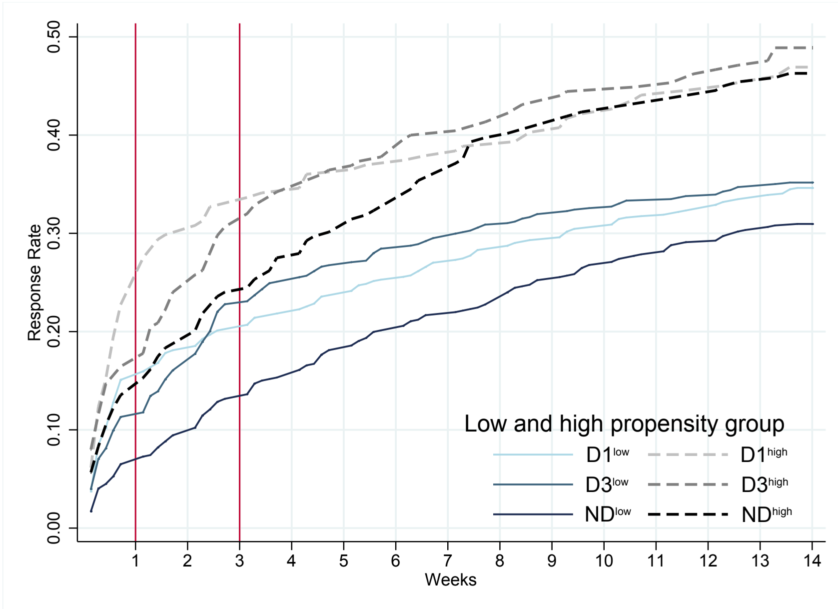
Notes: The figure presents the (empirically observed = raw) cumulative response rate over 14 weeks for each of the nine combinations of deadlines and rewards from our main experiment.

Figure B.4: Hazard Rates for Subgroups (Main Experiment)



Notes: The figure displays raw, daily hazard rates for the D1, D3 and ND treatment over a period of six weeks (42 days), splitting the sample among those with the highest (top 25%) and with lower predicted propensities to respond (see text for details).

Figure B.5: Cumulative Response Rates for Subgroups (Main Experiment)



Notes: The figure presents the empirically observed cumulative response rates over 14 weeks, splitting the sample among those with the highest (top 25%) and with lower predicted propensities to respond (see text for details).

B.2 Tables

Table B.1: Summary Statistics and Balancing Tests (Follow-up Experiment)

	D1	D3	D6	D10	ND	EoY	F-Stats
Age	38.65	36.56	35.96	34.94	37.08	38.55	1.127
	(15.92)	(17.12)	(16.59)	(14.92)	(16.21)	(17.81)	[0.345]
Female	0.53	0.58	0.58	0.48	0.62	0.48	1.949
	(0.50)	(0.49)	(0.49)	(0.50)	(0.49)	(0.50)	[0.084]
Distance	12.74	10.07	8.58	6.82	15.05	13.50	2.000
	(43.23)	(25.94)	(19.29)	(14.55)	(56.42)	(34.36)	[0.077]
Housing price	8.09	8.12	8.13	8.14	7.99	7.86	1.029
	(1.18)	(1.07)	(1.21)	(1.07)	(1.23)	(1.14)	[0.400]
Private HI	0.18	0.20	0.23	0.17	0.17	0.25	0.774
	(0.38)	(0.40)	(0.42)	(0.38)	(0.38)	(0.44)	[0.569]
Family	0.12	0.17	0.19	0.16	0.15	0.14	0.500
	(0.32)	(0.38)	(0.39)	(0.36)	(0.36)	(0.35)	[0.776]
At risk	0.12	0.15	0.06	0.15	0.10	0.11	1.688
	(0.32)	(0.36)	(0.24)	(0.36)	(0.30)	(0.31)	[0.135]
Patient retention	4.63	4.66	5.12	5.07	4.78	4.80	0.699
	(3.02)	(2.84)	(2.78)	(2.93)	(2.96)	(3.32)	[0.624]
Past showup$^\flat$	0.73	0.73	0.77	0.78	0.74	0.68	0.524
	(0.44)	(0.44)	(0.42)	(0.42)	(0.44)	(0.47)	[0.758]
Pain$^\flat$	0.24	0.20	0.20	0.27	0.24	0.27	0.576
	(0.43)	(0.40)	(0.40)	(0.44)	(0.43)	(0.45)	[0.719]
N	169	163	163	166	175	91	

Notes: The table presents summary statistics (mean and standard deviation in parenthesis) of participant characteristics for each treatment of the follow-up experiment. The last column reports F-statistic [and the corresponding p-values in brackets] from tests for treatment differences based on separate regressions of each of the characteristics on a full set of treatment dummies. Sample size is $N = 927$. The two variables indicated with $\flat$, *Pain* (records on painful dental treatments in the past) and *Past-Showup* (one show-up in the year before the intervention) are only available for 716 and 789 observations, respectively.

Table B.2: Duration Analysis with Time-Varying Treatment Effects (Main Experiment)

	(1)	(2)	(3)	(4)	(5)	(6)
	$Z = 1$ (last day)			$Z = 2$ (last two day)		
$\text{D1}^{\text{w1-w/o-last}Z\text{days}}$	3.493	3.567	3.509	3.363	3.446	3.695
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\text{D1}^{\text{w1-last}Z\text{days}}$	3.576	3.623	2.770	3.684	3.732	2.941
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
D1^{w2}	1.266	1.299	1.239	1.261	1.294	1.241
	[0.194]	[0.155]	[0.286]	[0.202]	[0.161]	[0.284]
D1^{w3}	0.819	0.835	0.612	0.818	0.833	0.610
	[0.375]	[0.426]	[0.036]	[0.371]	[0.421]	[0.035]
$\text{D1}^{\text{w4+}}$	0.672	0.675	0.723	0.675	0.677	0.721
	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.001]
D3^{w1}	2.587	2.633	2.444	2.555	2.603	2.470
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
D3^{w2}	1.667	1.717	1.696	1.660	1.710	1.698
	[0.003]	[0.002]	[0.005]	[0.003]	[0.002]	[0.005]
$\text{D3}^{\text{w3-w/o-last}Z\text{days}}$	2.631	2.709	1.946	2.319	2.408	1.722
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.005]
$\text{D3}^{\text{w3-last}Z\text{days}}$	1.600	1.608	1.140	2.641	2.652	1.900
	[0.216]	[0.213]	[0.734]	[0.000]	[0.000]	[0.004]
$\text{D3}^{\text{w4+}}$	0.678	0.699	0.749	0.680	0.701	0.748
	[0.000]	[0.001]	[0.010]	[0.001]	[0.002]	[0.009]
Controls	–	Yes^a	Yes^b	–	Yes^a	Yes^b
Post-Estimation Tests: Comparing last Z days						
$\text{D1}^{\text{w1-last}Z\text{days}} = \text{D3}^{\text{w3-last}Z\text{days}}$	0.063	0.061	0.040	0.198	0.185	0.096

Notes: The table presents hazard ratios estimated with Weibull proportional hazards model ($N = 2,661$). Specifications (2) and (4) includes dummies for different incentive levels, controls for individual characteristics, and dummies absorbing wave specific effects. Specifications (3) and (5) adds further dummies that account for week×year specific effects. p-values – based on robust standard errors, clustered at the household level – are reported in brackets.

Table B.3: Duration Analysis (Follow-up Experiment)

	(1)	(2)	(3)
D1	1.201	1.205	1.203
	[0.308]	[0.305]	[0.327]
D3	1.391	1.336	1.350
	[0.056]	[0.089]	[0.093]
D6	1.078	1.022	1.020
	[0.672]	[0.904]	[0.915]
D10	1.033	0.998	0.998
	[0.850]	[0.991]	[0.993]
EoY	1.079	1.205	1.215
	[0.702]	[0.383]	[0.382]
ρ (shape parameter)			0.563
			[0.000]
Controls	–	Yes	Yes
Post-Estimation Tests: (p-values)			
D1=D3	0.395	0.549	0.523
D1=D6	0.545	0.362	0.381
D1=D10	0.374	0.276	0.301
D1=EoY	0.603	0.999	0.965
D3=D6	0.134	0.116	0.118
D3=D10	0.072	0.075	0.079
D3=EoY	0.206	0.616	0.626
D6=D10	0.801	0.892	0.906
D6=EoY	0.998	0.437	0.431
D10=EoY	0.827	0.371	0.374

Notes: The table presents hazard ratios estimated with Cox (columns 1 and 2) and Weibull (column 3) proportional hazards model ($N = 927$). Specifications (2) and (3) include dummies for different incentive levels, controls for individual characteristics, and dummies absorbing wave specific effects. p-values—based on robust standard errors, clustered at the household level—are reported in brackets. ρ is the estimated shape parameter of the Weibull distribution, with $\rho < 1$ indicating a decreasing hazard rate. The lower part of the table reports p-values from Wald tests.

Table B.4: Treatment Effects on Period-t Completion Rates (Follow-up Experiment)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	7 Days		21 Days		42 Days		70 Days		100 Days	
D1	0.111**	0.107**	0.050	0.046	0.064	0.066	0.054	0.056	0.038	0.039
	(0.046)	(0.044)	(0.052)	(0.051)	(0.055)	(0.053)	(0.055)	(0.053)	(0.056)	(0.054)
D3	0.053	0.052	0.140***	0.137***	0.108*	0.100*	0.100*	0.087*	0.091	0.074
	(0.041)	(0.041)	(0.052)	(0.050)	(0.055)	(0.052)	(0.056)	(0.053)	(0.057)	(0.054)
D6	-0.033	-0.029	-0.050	-0.045	0.034	0.037	0.051	0.050	0.023	0.024
	(0.036)	(0.036)	(0.046)	(0.046)	(0.056)	(0.053)	(0.058)	(0.055)	(0.059)	(0.056)
D10	-0.047	-0.045	-0.047	-0.050	-0.038	-0.035	0.019	0.019	0.022	0.021
	(0.035)	(0.036)	(0.048)	(0.049)	(0.051)	(0.052)	(0.055)	(0.054)	(0.056)	(0.054)
EoY	0.000	0.027	-0.009	0.018	-0.011	0.026	0.041	0.073	0.029	0.055
	(0.046)	(0.046)	(0.059)	(0.058)	(0.061)	(0.061)	(0.063)	(0.063)	(0.064)	(0.065)
Constant	0.131***	-0.097	0.240***	-0.055	0.297***	0.073	0.354***	0.199	0.400***	0.197
	(0.027)	(0.113)	(0.035)	(0.174)	(0.038)	(0.186)	(0.038)	(0.188)	(0.040)	(0.194)
Post-Estimation Tests: (p-value)										
D1=D3	0.207	0.220	0.090	0.073	0.425	0.519	0.424	0.571	0.355	0.516
D1=D6	0.001	0.002	0.040	0.059	0.601	0.584	0.954	0.910	0.806	0.802
D1=D10	0.000	0.000	0.048	0.053	0.051	0.054	0.530	0.500	0.771	0.748
D1=EoY	0.035	0.108	0.327	0.625	0.231	0.514	0.846	0.791	0.889	0.803
D3=D6	0.025	0.037	0.000	0.000	0.196	0.240	0.409	0.510	0.258	0.382
D3=D10	0.009	0.011	0.000	0.000	0.009	0.014	0.154	0.219	0.225	0.328
D3=EoY	0.286	0.609	0.016	0.046	0.060	0.223	0.375	0.826	0.355	0.773
D6=D10	0.667	0.638	0.953	0.914	0.178	0.182	0.592	0.598	0.978	0.958
D6=EoY	0.447	0.201	0.476	0.266	0.476	0.870	0.890	0.722	0.939	0.649
Controls	–	Yes	–	Yes	–	Yes	–	Yes		

Notes: The table presents LPM estimates of the follow-up experiment treatment effects on the cumulative response rate (the probability of calling) within 7, 21, 42, 70 and 100 days, respectively ($N = 927$). The lower part reports the p-values from Wald tests. Every second specification includes individual control variables and dummies absorbing wave specific effects. Robust standard errors, clustered at the household level, are reported in parentheses. ***, **, * indicates significance at the 1%-, 5%-, 10%-level, respectively.