

Nudges at the Dentist*

Steffen Altmann[†] and Christian Traxler[‡]

August 2014

Forthcoming in the EUROPEAN ECONOMIC REVIEW

We implement a field experiment to study the impact of reminder messages on dental health prevention. Patients who are due to schedule a check-up receive no reminder, a neutral reminder postcard, or reminders including additional information on the benefits of prevention. Our results document a strong impact of reminders. Within one month after receiving a reminder, the fraction of patients who make a check-up appointment more than doubles. The effect declines slightly over time, but remains economically and statistically significant. Including additional information in the reminders does not significantly increase response rates relative to the neutral reminder. Finally, our data indicates that applying reminders repeatedly neither strengthens nor weakens their effects.

JEL classification: D03, I11, C93

Keywords: Field Experiment, Reminders, Nudges, Memory Limitations, Attention, Awareness, Prevention, Dental Health, Framing

* We thank the editor, Joerg Oechssler, an associate editor, and three anonymous referees for their thoughtful comments. The paper also benefited from discussions with Michele Belot, Christoph Engel, Michael Hallsworth, Marc Jekel, Andrew Oswald, Philipp Weinschenk, Matthias Wibrals, Joachim Winter, and from valuable comments by conference and seminar participants in Ann Arbor, Berlin, Bonn, Cologne, Munich, Rotterdam, and Vienna. We are indebted to Ulf Jüptner and Juliane Heyda for their invaluable support in conducting the experiment. Nicolas Kaufung and Tim Stahringer provided excellent research assistance. Financial support from the Institute for the Study of Labor, the Max Planck Institute for Research on Collective Goods, and Volkswagen Foundation is gratefully acknowledged.

[†] IZA and University of Bonn. Address: Institute for the Study of Labor, Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. Email: altmann@iza.org

[‡]Hertie School of Governance, Max Planck Institute for Research on Collective Goods, and CESifo. Address: Hertie School of Governance, Friedrichstrasse 180, 10117 Berlin, Germany. Email: traxler@hertie-school.org

1 Introduction

A vast body of evidence documents that limitations in self-control, memory, and attention may lead people to act against their long-run self interest (see Rabin 1999 and DellaVigna 2009 for comprehensive reviews of the literature). These findings have spurred the interest in how subtle changes in choice architectures can improve ‘Decisions About Health, Wealth, and Happiness’ (Thaler and Sunstein 2008). Choice-preserving changes in seemingly minor institutional details – nudges – hold the promise to help some people avoid making mistakes without distorting choices of others. Examples of nudges that have been shown to strongly affect decisions include the specification of default options, information disclosure policies, or the framing and labeling of policy interventions (e.g., Johnson and Goldstein 2003, Wisdom et al. 2010, Abeler and Marklein 2010).

In this paper, we study whether simple reminder messages can be used as a nudge to encourage dental health prevention. Preventive health care constitutes a potential hot spot for biases in people’s decisions. Activities such as screenings and routine check-ups should be conducted regularly but at a relatively low frequency. As a consequence, these key elements of effective health prevention are vulnerable to limitations in memory and attention: people may simply forget about scheduling the next check-up. Moreover, investments in prevention typically entail short-run costs but yield benefits only in the long-run. Present-biased preferences may thus lead people to procrastinate their next check-up. Finally, people might under-invest in preventive health care because they are not fully aware of its future benefits. Reminders about preventive activities represent a low-cost and unobtrusive intervention in that they let people schedule their check-ups whenever they want to. A person who rationally postpones a check-up can thus just ignore the reminder message. Someone who would otherwise delay a check-up due to memory limitations or other biases, however, might benefit from receiving a reminder.

In cooperation with a German dentist, we implemented a field experiment to examine the impact of reminders on the frequency of dental check-ups. Regular check-ups are associated with improved periodontal health and ensure that oral diseases are discovered early – which typically makes treatment less painful and less costly (e.g., Lang et al. 1994). In addition, healthy teeth yield economic returns (Glied and Neidell 2010). To identify the causal effect of reminders, we exogenously vary whether people are reminded to schedule a new check-up appointment. Within the group of people who receive a reminder, we additionally randomize

the content of the reminder messages. In particular, we vary whether the reminder includes a paragraph that informs patients about the benefits of prevention. We also randomize whether this information is framed positively or negatively, as it is predominantly the case with health warning messages for tobacco and other potentially harmful products.

We find that reminders cause a substantial increase in the number of check-ups. Within one month after receiving a reminder, the fraction of people who contact the dentist and schedule a check-up appointment is more than twice as high as in our control group (19.3% vs. 8.9%). Similarly, the fraction of patients who visit the dentist and have a check-up is more than 10 percentage points higher in the treated group. The magnitude of the reminder effect is remarkably persistent over time. Even 100 days after the intervention, the fraction of patients who scheduled a check-up is roughly 8 percentage points higher in the treated group (33.8% vs. 26.0%). Reminders thus encourage people to engage more strongly in dental health prevention.

In contrast to the strong overall impact of being reminded, the specific type and content of the reminder seems to matter relatively little. In particular, adding information about the benefits of prevention does not significantly increase patients' response rates relative to a neutral reminder. We also find no systematic differences in responses to a positive or negative framing of the reminders' content. These results hold for a broad range of patient subgroups and are robust, for instance, to controlling for individual differences in age, gender, and check-up history.

We also analyze how people behave when they are repeatedly exposed to reminders. Over the course of our experiment, several patients are treated twice as they are due for a further check-up. Our randomization procedure ensures that we obtain random treatment sequences, in terms of whether and how a person is reminded the first and second time she is up for a check-up. Our data show that patients who receive multiple reminders exhibit similar response patterns as patients who receive the first reminder in our experiment. At the same time, we find no detrimental effects if one stops sending a reminder at the second check-up date.

A number of previous studies used field experiments to analyze the effects of reminders in the context of health decisions (e.g., McDowell et al. 1989, Macharia et al. 1992, O'Keefe and Jensen 2009, Milkman et al. 2011, Calzolari and Nardotto 2014), as well as in other economically relevant settings such as loan repayment (e.g., Cadena and Schoar 2011, Karlan et al. 2012), saving (e.g., Kast et al. 2012, Karlan et al. 2013), or rule compliance (e.g.,

Apestequia et al. 2013). Table A.1 in the appendix provides an overview of the setup and key results of these studies. Our paper contributes to the existing literature in several dimensions. First, by varying the content and framing of reminder messages, we provide new insights on the mechanisms through which reminders affect behavior. Overall, our results indicate that reminders increase the number of check-ups primarily because they raise patients' attention and attenuate potential delays due to memory limitations. In contrast, limited awareness about the benefits of prevention seems to matter relatively little, as providing information on these benefits does not increase response rates relative to a neutral reminder. These results augment the evidence from other studies that have explored alternative channels through which reminders might affect behavior, such as habit formation (Calzolari and Nardotto 2014), monitoring (Kast et al. 2012, Karlan et al. 2012), or the provision of precisely defined goals, rules, and incentives (Karlan et al. 2013, Apestequia et al. 2013).

Second, for the domain of dental health, the existing randomized trials have exclusively focused on the question whether reminders influence patients' attendance at pre-arranged appointments for orthodontic surgery and related treatments (Reekie and Devlin 1998, Can et al. 2003, Bos et al. 2005). These studies have found mixed evidence on the effectiveness of reminders in reducing the frequency of broken appointments. In contrast, we show that reminders significantly increase patients' responses and check-up rates in a preventive-health setting where people are reminded about the need to schedule new dental check-ups.

Third, our extensive set of sociodemographic characteristics allows us to derive nuanced insights on whether certain groups of people are especially responsive to reminders, and whether particular types of reminders have stronger effects for some subgroups. This is of high practical relevance from a public policy perspective and might shed further light on why particular groups of people respond to reminders. Our analysis shows that simple nudges tend to work best for a wide range of people: the neutral reminder message triggers the highest response rate among almost all groups of patients. In fact, a uniform neutral reminder might be preferable to the costly targeting of specific groups with different reminder nudges. This is the case, as framed reminders sometimes backfire and erode the positive reminder effect in certain subgroups (e.g., for patients with painful treatment experiences in the past). At the same time, our evidence points to a potentially important role of personal costs in explaining differential responses to reminders. For instance, patients with private health insurance plans, who typically face higher costs of making check-ups, hardly respond to any of the reminders. This

complements findings from other papers that have compared the effectiveness of reminders to alternative interventions that directly vary the explicit or implicit incentives for carrying out a task (e.g., Cadena and Schoar 2011, Kast et al. 2012).

Finally, our paper provides a systematic evaluation of repeated treatment interventions in the context of reminders.¹ Other studies of recurrent reminder systems (e.g., Apestegua et al. 2013, Karlan et al. 2013, or Calzolari and Nardotto 2014) have either focused primarily on overall treatment effects or kept the treatment assignment constant for a given participant. In contrast, we study treatment sequences that randomly vary both the incidence of being reminded and the type of reminder received. The data from these treatment sequences indicate that applying reminders repeatedly neither strengthens nor weakens their effects. Importantly, we also observe no detrimental effect for patients who do not receive a reminder after having received one in the previous check-up interval. The response rate of such patients still lies above the one of a control group that remained untreated in both periods. This indicates that reminders do not crowd out people’s active choices once the nudge is taken away – a concern that has frequently been raised by critics of libertarian paternalistic interventions (e.g., Rizzo and Whitman 2009).

The remainder of the paper is organized as follows. In the following section, we first give an overview of the institutional context and the treatments of our experiment. We then introduce a simple theoretical framework to illustrate how different biases shape patients’ decisions to have check-ups, and how our reminders might affect these decisions. In Section 3, we describe the implementation of the experiment. Section 4 presents our empirical results and Section 5 concludes.

2 Dental health prevention, check-ups, and reminders

Health prevention is a non-trivial inter-temporal problem. In the case of dental health, brushing one’s teeth regularly, flossing, and avoiding sugary food is costly today – in terms of time, effort, and foregone pleasure. Such investments, however, potentially generate high returns through improved future health. In addition to health-related payoffs, healthy teeth may also yield higher wages (Glied and Neidell 2010) and other benefits associated with beauty (e.g., Mocan and Tekin 2010, Belot et al. 2012).

¹Further evidence on the long-run effects of repeated treatments in the context of energy-conservation policies and interventions to increase charitable giving is provided by Allcott and Rogers (2014) as well as Landry et al. (2010) and Adena et al. (2014), respectively.

A key factor in dental health prevention are routine check-ups. These ensure that oral diseases are discovered early, which typically makes treatment less painful and less costly (e.g., Lang et al. 1994). Recommendations for check-up intervals vary between three months and more than a year, depending on patients' risk factors and dental health status (see, e.g., National Institute for Health and Clinical Excellence 2004). In Germany, the recommended check-up interval for healthy patients is six months. At-risk groups (e.g., diabetes mellitus patients who face increased parodontitis risks) are recommended more frequent check-ups.² Acknowledging the importance of dental prevention, many health care providers use economic incentives to encourage regular check-ups. Patients covered by the German public health insurance (more than 80% of our sample), for instance, are eligible for free dental check-ups and receive a 20% [30%] allowance on treatment costs if they had at least one check-up per year for the last 5 [10] years.³

In our setup, there are two steps to a check-up. First, a patient has to contact her dentist to schedule an appointment. While this entails only trivial costs, a patient has to remember to contact her dentist when a new check-up is due.⁴ Second, once a date for the check-up is arranged, the patient has to keep the appointment and go to the dentist, which involves transaction and opportunity costs (e.g., from traveling to the dentist).

Fully rational patients who trade off their personal costs and benefits of prevention will attend check-ups at an individually optimal frequency. However, a growing body of research suggests that some people may be prone to make sub-optimal decisions in the context of health prevention (e.g., Loewenstein et al. 2007, Thaler and Sunstein 2008, Oullier et al. 2010). In the case of routine check-ups, there are three main factors that can cause delays in check-ups. First, people might not have the task of making a new check-up appointment on the 'top of their minds' because of limitations in memory and attention. Second, people might not be fully aware of the future benefits of check-ups, e.g., because they underestimate the risk and

²There is an ongoing debate about best practices for check-up intervals in dentistry. The debate reflects a lack of causal evidence on how different intervals affect long-run health outcomes (see, e.g., Mettes 2005). Motivated by the missing evidence, a large-scale randomized trial on different check-up intervals has recently been launched in the UK (see <https://viis.abdn.ac.uk/HSRU/Interval/>).

³ Private health insurance contracts are more heterogeneous. Generally, the incentives to regularly attend check-ups are typically weaker than in the public insurance system. In addition, the costs for check-ups tend to be higher for privately insured patients.

⁴Some dentists assign a date for the next check-up already at the previous appointment. With this procedure, it is less important to remember when to arrange a new appointment. It is crucial, however, to keep in mind the date of the future, pre-arranged appointment.

cost of future treatments. Finally, the short-run costs associated with check-ups might lead present-biased people to delay their check-ups.

Depending on their content, reminder messages may work against all of these biases. At the same time, reminders preserve patients' freedom to schedule check-ups whenever they want to. Those who rationally postpone check-ups (e.g., because of high opportunity costs) can thus simply ignore the reminders. In what follows, we first describe the reminder messages that we implemented in our experiment. We then present a simple theoretical framework to illustrate how memory limitations, limited awareness, and present biases affect people's propensity to schedule check-ups, and how reminders may influence their behavior.

2.1 Treatments

We examine a nested experimental design with two layers of randomization. First, we exogenously vary whether individuals who should make a new check-up according to their recommended check-up interval receive a reminder postcard. Second, within the treated group we randomize design, content, and framing of the reminder postcards. Overall, we consider six different treatments (an overview of the treatments is presented in Figure 1). The group of patients who receive no reminder (*No-Reminder*) serves as our main control group. Patients in treatment *Neutral* receive a reminder postcard that asks to make a new check-up appointment. In addition, the postcard provides the dentist's name, phone number, and address. This content is also included in all further reminder treatments, which contain additional information about the benefits of dental health prevention. Two treatments provide positively framed information that emphasizes the long-run health benefits as well as potential pecuniary advantages due to prevention. The final two treatments provide the same information, but the text stresses the negative consequences of not taking care of dental health (see Table A.2 in the appendix).

Together with the text manipulations we vary the cover design of the postcards (see Figure A.1 in the appendix). The neutral postcard displays a picture of a calendar, with the word 'Dentist' written on one of the dates. In the positively framed treatments, the postcard shows either a smiling female or male face with healthy teeth (treatments Pos^F and Pos^M , respectively). In the treatments with negative framing of information, the cover displays a female or male patient with toothache (Neg^F , Neg^M).

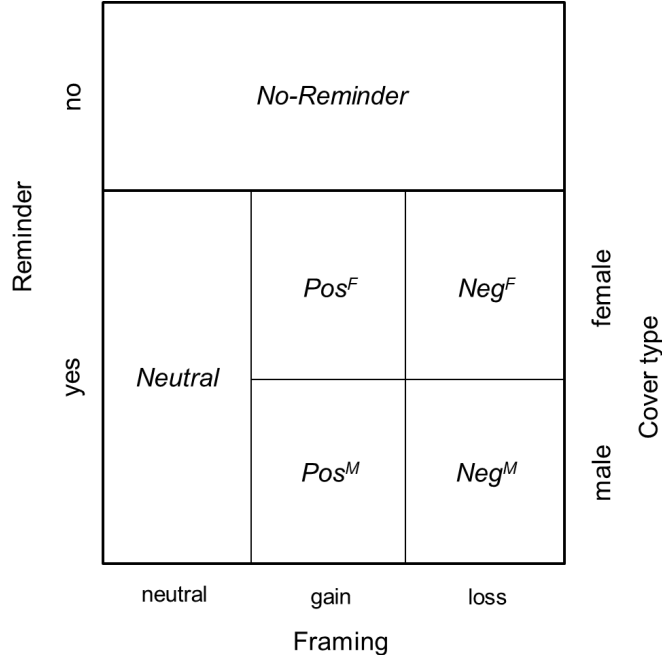


Figure 1: *Overview of treatments.*

2.2 Decision framework

Before discussing how our treatments might affect patients’ behavior, we first sketch a patient’s check-up problem in a simple three-period framework. In t_0 , the patient decides whether she calls the dentist to schedule a check-up for t_1 .⁵ Conditional on having arranged a check-up, she chooses whether or not to actually attend the appointment in t_1 . In the final period, t_2 , health outcomes are realized; the net benefits from having attended a check-up in t_1 are denoted by $H \geq 0$.⁶

Consider first the patient’s decision in t_1 . Conditional on having arranged an appointment, the patient attends the check-up in this period if the costs of doing so are smaller than the perceived net benefits of the check-up:

$$c_1 \leq \hat{c}_1 := \beta\delta\alpha H. \tag{1}$$

⁵In line with the institutional setting of the experiment, we do not allow for spontaneous check-ups.

⁶ H is a short-cut to capture the expected net present value (in t_2) of all future benefits that result from the check-up in t_1 . H might be specific to individuals and depends, among others, on baseline risk factors and check-up history. A more general analysis, which is beyond the scope of this paper, would also consider the optimal timing of check-ups and the impact of different check-up frequencies on health outcomes.

c_1 denotes the realization of the patient's (stochastic) attendance costs, which include the costs of traveling to the dentist, opportunity costs, fear, etc. The parameter α , with $0 \leq \alpha \leq 1$, captures the possibility that a patient may not be fully aware of the future benefits of the check-up. A patient who is partially unaware thus bases her attendance decision on the perceived benefits, αH , instead of the true benefits H . Furthermore, the patient discounts future benefits at rate $\beta\delta$, where $0 < \beta, \delta \leq 1$ measure the patient's time preferences and a potential present bias ($\beta < 1$), respectively.

Let c_1 be drawn from a c.d.f. $F(\cdot)$ that has full support over $[0; \bar{c}_1]$. The patient's attendance probability is then given by $F(\hat{c}_1)$. Conditional on having arranged an appointment for t_1 , a patient who is fully aware of the future health consequences ($\alpha = 1$) and exhibits no present bias ($\beta = 1$) will thus attend the check-up with probability $F(\delta H)$. Limited awareness and present bias have qualitatively similar effects on the patient's attendance probability. They both reduce the decision-relevant level of perceived benefits, resulting in a lower attendance probability of $F(\beta\delta\alpha H) \leq F(\delta H)$ (conditional on having arranged an appointment).

Now turn to t_0 where the patient decides whether or not to call her dentist for scheduling an appointment. At this stage, a patient might be affected by limitations in memory and attention. Given that check-ups have to be scheduled at a fairly low frequency, a patient might not remember to contact her dentist when it is time for a new check-up. We assume that this occurs with probability $1 - \pi$. Hence, in t_0 , a patient only enters the decision problem with probability $0 < \pi \leq 1$.⁷ Note that, in principle, it is also conceivable that a patient arranges an appointment in t_0 , but does not remember the appointment anymore when arriving at t_1 . Given that the average time gap between arranging and attending an appointment in our setup is only 14 days, we abstract from this possibility. Incorporating this additional facet of memory limitations would not qualitatively change the analysis.

If the patient processes the decision problem in t_0 she contacts the dentist to arrange a check-up if

$$c_0 \leq \hat{c}_0 := \beta\delta F(\hat{c}_1) [\delta\alpha H - \mathbb{E}(c_1 | c_1 \leq \hat{c}_1)], \quad (2)$$

⁷In this way, we capture a central theme of a broader class of models on limited attention, memory, and complexity (e.g., Eliaz and Spiegler 2011, Karlan et al. 2013, Calzolari and Nardotto 2014). In general, people have a very large set of alternative tasks and choices available at any point in time. To reduce complexity, a given task might thus not always enter the set of relevant problems that are processed by a decision maker.

i.e., whenever the costs of scheduling the check-up (c_0) are below the expected net benefits of the check-up.⁸ Assuming that c_0 is independently distributed according to the c.d.f. $G(\cdot)$ over the interval $[0, \bar{c}_0]$, the probability of arranging an appointment is given by $G(\hat{c}_0)$. It is straightforward to show that $G(\hat{c}_0)$ depends positively on both α and β . A lower degree of awareness about the benefits of prevention or a stronger present bias will therefore both reduce a patient’s likelihood of arranging an appointment. Independently of the rule described in (2), stronger memory limitations will further reduce the likelihood of scheduling an appointment, because the patient enters the decision problem in t_0 with a lower probability π .

2.3 Treatment effects

The framework from the previous section illustrates that memory limitations, limited awareness, and present biases all reduce a patient’s likelihood of having check-ups. How could reminders influence patients’ behavior in this setting? First, reminders bring the task of scheduling a new appointment to people’s minds. They might thus increase the level of π and encourage patients with memory limitations to arrange more check-ups. Second, some of our reminders provide patients with information about the benefits of prevention. These reminders might induce more check-ups as they lift the level of α for partially unaware patients. Third, all of our reminders include the dentist’s contact details. This might reduce the transaction costs of arranging an appointment (c_0) and lead to an increase in check-ups for all patients. Hence, even though our reminders do not specifically address present biases, they might raise the check-up rate of present-biased patients through the general reduction in the costs of responding.

By analyzing outcomes along our first treatment dimension – i.e., by comparing responses of patients who received a reminder to those who did not – we can examine whether reminders affect patients’ check-up behavior, regardless of the underlying mechanisms. From a public policy perspective, this is perhaps the most important question. In addition, the comparison of response rates between the different reminder treatments can shed some light on the relative importance of different underlying mechanisms. Even though our experiment is not designed to precisely pin down one specific channel through which reminders work, this allows us to narrow down the set of relevant motives for patients’ treatment responses.

⁸For simplicity we focus on the case that patients hold unbiased beliefs about α , β , δ , and $F(\cdot)$. An interesting deviation from this level of sophistication concerns (partially) naive patients, who anticipate a present bias $\hat{\beta}$ which is less severe than their true bias ($1 \geq \hat{\beta} > \beta$). An extended appendix, which is available from the authors upon request, provides a discussion of how this can shape patients’ check-up behavior.

To see this, consider first the hypothesis that patients experience limitations in memory and attention ($\pi < 1$), while limited awareness about the benefits of prevention plays no role ($\alpha = 1$). In this case, reminders induce $\pi^{\text{Rem}} > \pi^{\text{No-Rem}}$, thereby raising the check-up frequency of treated patients relative to the control group. Given that the memory-enhancing effect is similar for the different types of reminders (i.e., $\pi^{\text{Neut}} \simeq \pi^{\text{Pos}^F} \simeq \pi^{\text{Pos}^M} \simeq \pi^{\text{Neg}^F} \simeq \pi^{\text{Neg}^M}$), however, we should observe not behavioral differences across the reminder treatments.

In contrast, under the alternative hypothesis of limited awareness being crucial ($\alpha < 1$), our framework predicts differential effects for the different reminders treatments (for any given level of π). The framed treatments, which all stress the benefits [costs] of [not] making dental check-ups, should raise awareness relative to the *Neutral* and *No-Reminder* treatment, where no such information is provided (i.e., $\alpha^{\text{No-Rem}} \simeq \alpha^{\text{Neut}} < \alpha^{\text{Pos}^M} \simeq \alpha^{\text{Pos}^F} \simeq \alpha^{\text{Neg}^M} \simeq \alpha^{\text{Neg}^F}$). If limited awareness is the key mechanism behind patients' response to treatments, we should therefore observe that the framed reminders have a stronger effect than the neutral reminder. If, in addition, we observe *no* differences in the number of check-ups between the *Neutral* and the *No-Reminder* treatment, this would be strong evidence that neither a cost-reducing effect of reminders nor memory limitations are relevant (i.e., $\pi = 1$).

Further inference on the awareness channel can be drawn by analyzing differences in reactions to the positively and negatively framed reminders. Starting with Tversky and Kahneman (1981), a large body of literature has shown that the framing of messages can alter people's perception of the information that is provided. In our context, this could imply systematic differences between α^{Pos} and α^{Neg} with corresponding treatment differences in check-up frequencies.⁹ Similarly, the way people process the information on the postcards (i.e., the reminder's impact on α) might depend on the gender framing. We can examine such effects by comparing responses between Pos^F , Neg^F and Pos^M , Neg^M . This explorative analysis can be further motivated by evidence from other domains, such as consumers' responses to advertisement (Bertrand et al. 2010) or fundraising campaigns (Landry et al. 2006), which suggests that gender framing effects might be quantitatively important.

Since all postcards in our experiment remind patients about the need to schedule a new appointment and provide them with the dentist's contact details, it is difficult to differentiate a memory-enhancing effect (an increase in π) from a general reduction of transaction costs

⁹Earlier studies in health psychology have explored the effects of negative and positive frames in health communications. The evidence, however, has remained somewhat inconclusive regarding the size and direction of the framing effects (see O'Keefe and Jensen 2009 and Gallagher and Updegraff 2012 for recent reviews).

due to reminders (lower levels of c_0). Arguably, however, the pure transaction-cost effect is likely to be small, given that the dentist’s contact details are easily accessible via the internet or stored in the electronic directories of patients’ mobile phones. In fact, evidence from a randomized trial in the UK indicates that including contact details in patients’ reminders has no additional impact relative to a simple reminder that does not contain this information (Behavioural Insights Team 2014). In what follows, we therefore focus on the reminders’ role in tackling patients’ limitations in memory and awareness about the benefits of prevention.

3 Implementation of the experiment

We conducted the experiment in cooperation with a dentist in Bonn, Germany. At their first visit at the dentist, patients are classified depending on their dental health status as ‘regular’ or ‘at-risk’ patients, corresponding to recommended check-up frequencies of 6 or 4 months, respectively. We use this classification to determine who enters the experiment at a given point in time. In particular, patients with a recommended m -month frequency who have not been at the dentist since m months are marked with a reminder flag. On a monthly basis, we randomly assign all patients with an active flag to our treatments. When a patient is treated (or stays untreated in case she is assigned to the control group) or when a patient comes to a check-up, we switch off her reminder flag for another m months.¹⁰ Hence, a patient will receive reminders in intervals of at least m months.

Note that the procedure implies that we randomize within a selected sample: patients who schedule their check-ups in a particularly timely manner are less likely to be part of our sample, whereas patients who are not sticking to their recommended check-up schedule have a higher likelihood of participating in the experiment. From a public health perspective, our experimental sample seems to be interesting for implementing nudge-based policies, since it pools people who rationally postpone check-ups with others who might exhibit sub-optimally low check-up frequencies due to memory limitations, unawareness, or present biases.

Patients are assigned to each of our six treatments with equal probability. In each randomization wave, members of the same household are assigned to the same treatment. By randomizing at the household level we try to avoid possible treatment spillovers (Miguel and Kremer 2004) and minimize patients’ consciousness of an ongoing experiment. Note further that several patients are treated more than once during the 11 months in which the experiment

¹⁰If a patient visits the dentist after the flag is generated but before we randomize, the flag is switched off as well.

was conducted. This is the case since some patients are flagged for a check-up at two points in time during the experimental period. We independently randomize each time a patient is up for a check-up. This procedure generates random treatment sequences, which allows us to study the impact of repeatedly treating patients and the role of different treatment orders.

The experiment was conducted between April 2010 and March 2011. Overall, our sample covers 1227 instances in which a patient was up for a check-up appointment. These cases spread over 817 individuals from 713 households. For roughly half of the patients, we thus have data on repeated treatment sequences.¹¹ Randomization was implemented in 11 waves, with reminders being sent out on average every five weeks. All postcards were sent out by the dentist on Fridays, immediately after randomization took place. According to *Deutsche Post*, postcards should be delivered on Monday at the latest. We therefore measure the response time (in days), starting with the Monday after a randomization wave as day 1. For the response duration in the control group, we use the same starting day.

Our main variable of interest is whether or not a participant calls the dentist for scheduling an appointment. The date of the call measures the first point in time at which people respond to our intervention. In a second step, we study patients' actual show up at the check-up appointment (which is, on average, scheduled 14 days after a patient called). Considering the date of the first contact rather than the actual check-up has the advantage that it avoids potential congestion in check-up dates due to the dentist's capacity limits. More precisely, one might worry that finding longer delays in check-up dates for the *No-Reminder* condition might be the result of the response of patients in the reminder treatments, who 'block' dates for earlier appointments.

In addition to our main outcome variables, we compiled a rich set of patient background characteristics. Next to the patient's gender, age, and risk classification (i.e., the recommended check-up interval), we measure whether a patient is covered by private or public health insurance (*Private HI*).¹² We also observe whether a patient lives in a household in

¹¹Note that these figures reflect incidences of intention to treat: our data include about 5% of cases where reminders could not be delivered, e.g., because the patient had moved. The results reported below therefore slightly underestimate the impact of reminders in terms of treatment-on-treated effects.

¹² Controlling for insurance status is important since contractual terms might differ under private and public health insurance (see above). Moreover, in the German health system, the privately insured are a selected sample since only persons with a gross yearly income above 50K Euro, self-employed individuals, and civil servants are eligible for private health insurance. Due to Bonn's history as the former capital of West Germany, a large part of the privately health insured in our sample seem to be civil servants.

Treatment	<i>No-Rem</i>	<i>Neutral</i>	<i>Pos^F</i>	<i>Pos^M</i>	<i>Neg^F</i>	<i>Neg^M</i>	Total	F-test (<i>p</i> -values)	t-test
Age	37.23 (13.42)	38.62 (13.48)	40.11 (13.88)	36.64 (13.32)	37.28 (13.28)	38.40 (14.19)	38.05 (13.62)	0.128	0.364
Male	0.458 (0.500)	0.406 (0.492)	0.388 (0.488)	0.421 (0.495)	0.430 (0.496)	0.376 (0.485)	0.412 (0.492)	0.592	0.159
Family	0.151 (0.359)	0.129 (0.336)	0.092 (0.290)	0.134 (0.341)	0.087 (0.282)	0.108 (0.311)	0.117 (0.321)	0.288	0.105
Private HI	0.219 (0.414)	0.138 (0.346)	0.204 (0.404)	0.193 (0.396)	0.193 (0.396)	0.174 (0.380)	0.186 (0.389)	0.369	0.202
At-Risk	0.182 (0.387)	0.138 (0.346)	0.209 (0.408)	0.257 (0.438)	0.159 (0.367)	0.094 (0.292)	0.172 (0.378)	<0.001	0.680
Patient retention	3.717 (2.401)	3.627 (2.366)	3.830 (2.471)	3.410 (2.292)	3.546 (2.473)	3.549 (2.435)	3.611 (2.406)	0.593	0.505
Regular	0.469 (0.500)	0.482 (0.501)	0.447 (0.499)	0.405 (0.492)	0.470 (0.500)	0.466 (0.500)	0.457 (0.498)	0.727	0.731
Pain	0.249 (0.433)	0.196 (0.398)	0.172 (0.378)	0.237 (0.427)	0.204 (0.404)	0.252 (0.435)	0.218 (0.413)	0.292	0.272
Distance	0.125 (0.444)	0.217 (0.776)	0.125 (0.370)	0.138 (0.557)	0.117 (0.408)	0.116 (0.377)	0.141 (0.513)	0.296	0.646
Rental price	7.194 (1.181)	7.131 (1.183)	7.168 (1.224)	7.147 (1.182)	7.318 (1.225)	7.087 (1.264)	7.173 (1.210)	0.494	0.793
N	192	217	196	202	207	213	1227		

Table 1: *Background characteristics. Mean values (standard deviations in parentheses).*

Notes: ‘Family’ is a dummy, indicating if several household members took part in the study. ‘Private HI’ is equal to 1 if a patient is covered by private health insurance (0 for public health insurance). ‘At-Risk’ indicates if a patient is recommended a 4 months check-up interval. ‘Patient retention’ is the number of years since a patient first visited the dentist. ‘Distance’ is the great-circle distance between a patient’s home address and the dentist (scale: 1 = 100km). ‘Rental price’ is the average rent at the patient’s home address (Euros per square meter, excluding utilities and dues). The dummy ‘Regular’ indicates whether a patient made at least two check-ups in the year prior to a given randomization wave. ‘Pain’ indicates if a patient was exposed to a painful dental treatment in the past. Time-variant patient characteristics (e.g., age, patient retention, etc.) are computed at each randomization wave. The final two columns report (i) *p*-values of F-tests from regressions of the respective characteristics on individual treatment dummies and (ii) *p*-values of t-tests comparing the characteristics in the *No-Reminder* treatment vs. all patients who received a reminder.

which other household members took part in the experiment (*Family*).¹³ We further measure when a patient first visited the dentist, and derive information about individuals’ check-up and dental treatment history before the experiment. For people with a *Patient retention* of more than 12 months, we construct an indicator whether a patient attended at least two check-ups in the year prior to a randomization wave (*Regular*). Furthermore, we construct a binary variable, *Pain*, which captures whether a patient was exposed to a major root canal

¹³To avoid confounds from parents arranging joint appointments for themselves and their children, our analysis excludes children below the age of 15.

treatment or other painful dental treatments in the past.¹⁴ Beyond the information from the dentist’s database, we matched patients’ precise address with data obtained from *immobilien-scout24.de*, Germany’s largest real estate platform. In this way, we obtain the average rental price in the patients’ neighborhood (*Rental price*). From the address data, we also compute the spatial distance between each patient’s home and the dentist (*Distance*).

Table 1 presents the descriptive statistics of our data and summarizes the outcomes of randomization. Patients in our sample are on average 38 years old (with age varying between 15 and 92 years), 41% of them are male, and about 12% belong to households in which additional family members are patients at the dentist. As noted above, about 20% of patients are covered by private health insurance. On average, individuals in our sample have been patients at the dentist for 3.6 years, and 17% are classified as at-risk patients.

The final two columns of Table 1 report validation checks on the randomization. Comparing patients in the control treatment to all patients who received a reminder (pooled sample) yields no significant differences in patients’ background characteristics (see final column of Table 1). This indicates that randomization in patient characteristics was successful for our most important treatment comparison. F-tests from regressions of the individual patient characteristics on dummies for the different treatment conditions yield no significant difference except for the fraction of at-risk patients.

4 Results

This section presents our empirical results. Section 4.1 analyzes whether reminders increase the number of arranged and attended check-up appointments, and whether the observed effects are persistent over time. Based on our findings we also explore the underlying motives for patients’ responses. Section 4.2 discusses whether the impact of reminders differs for different groups of patients. Section 4.3 studies how people respond if they are nudged repeatedly.

¹⁴The results presented below are robust to using alternative definitions of patients’ check-up history and painful treatments. Note that one can not directly compare the figures on patients’ check-up and treatment history from Table 1 with the check-up frequencies reported below, as the numbers in Table 1 are based on ‘incidences of being treated’ (i.e., individual patients may enter the calculation more than once).

4.1 Do reminders influence behavior?

Descriptive analysis

Figure 2 provides a first impression of how reminders affect patient behavior. The figure depicts the fraction of patients who contact the dentist within one month after the treatment intervention. We observe a strong effect of reminders on patients' responses. On average, the fraction of patients who schedule a check-up in the reminder treatments is more than twice as high as in the group of patients that did not receive a reminder (19.3% vs. 8.9%). Compared to the strong overall impact of being reminded, the exact content of the reminder message seems to matter relatively little. Response rates for the different reminder treatments all lie in an interval between 16.3% (Pos^M) and 23.0% ($Neut$). In particular, Figure 2 indicates that the reminders that contain additional information on the benefits of prevention yield response rates that are similar to those in the *Neutral* treatment.

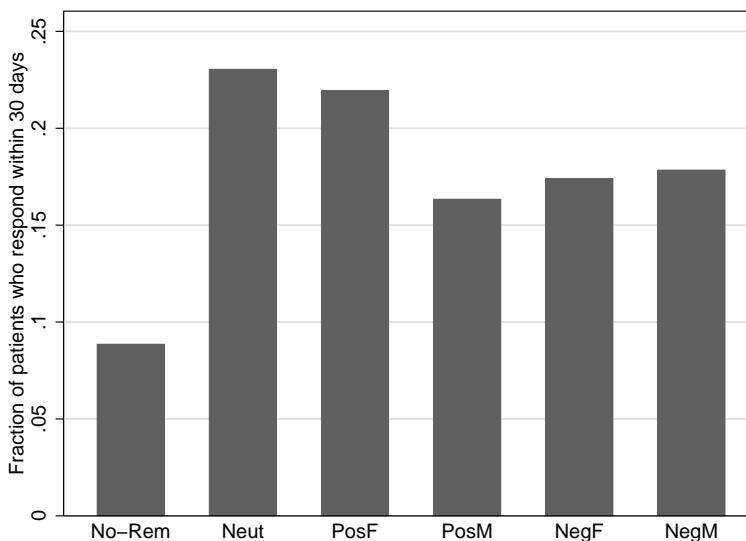


Figure 2: *Impact of reminders on response rates. Fraction of patients who contact the dentist within 30 days after treatment intervention.*

The strong direct impact of reminders on patients' response rates naturally raises the question whether these patterns also persist in the long run. A first answer to this question is given by Figure 3, which depicts the cumulative frequency of responses for patients in the reminder treatments (black line) and the control treatment (grey line) over the first 100

days after our treatment intervention.¹⁵ For ease of illustration, the figure pools the different reminder treatments. A more refined analysis of individual treatment differences is provided in the regression analysis below and in Figure A.2 in the appendix.

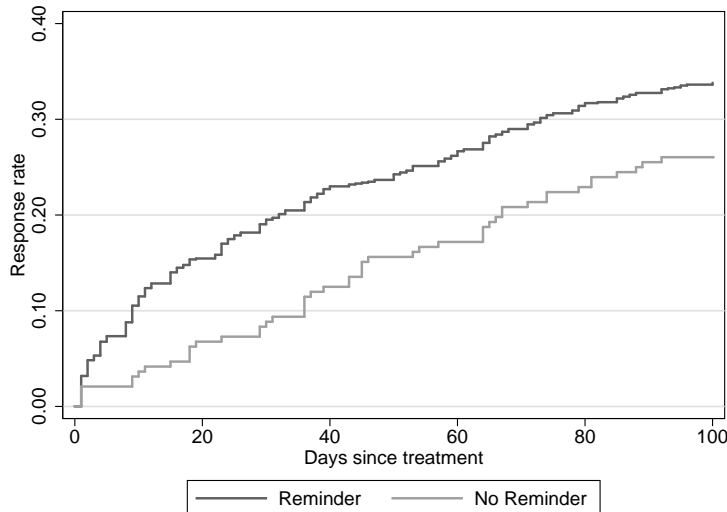


Figure 3: *Cumulated response in the reminder treatments (black line) and the control group (grey line).*

Substantiating what we have already seen in Figure 2, the graph demonstrates that already a few days after receiving a reminder, the response rate is strictly higher for treated patients. Over time, the gap in response rates declines slightly, but the treatment effect never vanishes. For days 51–100, the treatment difference remains almost constant at 8 percentage points. Overall, patients in the reminder conditions thus retain a persistently higher frequency of making check-up appointments.

Estimation analysis

In a next step, we parametrically estimate short-run treatment effects as well as the incidence and timing of patients’ responses in a duration framework. Throughout our analysis, we also examine whether the observed short- and long-run effects are robust to controlling for patient characteristics and potential seasonal differences. As our outcome variable, we first consider patient responses, i.e., whether a patient contacts the dentist to arrange a check-up appointment. Later we analyze whether patients actually show up at their appointments.

¹⁵Analyzing response rates for a longer time period is problematic because at-risk patients will already receive their next reminder.

The analysis of short-run responses builds on the model

$$\begin{aligned}
 Resp^{30} = & \alpha + \beta Rem + \gamma_1 Rem \times Pos^F + \gamma_2 Rem \times Pos^M \\
 & + \gamma_3 Rem \times Neg^F + \gamma_4 Rem \times Neg^M + X\delta + W\phi + \varepsilon.
 \end{aligned} \tag{3}$$

The dependent variable, $Resp^{30}$, measures whether a patient contacts the dentist within 30 days after the intervention and arranges a check-up appointment. X is a vector of patient characteristics and W includes dummies for the different randomization waves. Rem indicates whether a person has received any reminder postcard. The coefficient β thus identifies the main reminder effect. In the full specification of equation (3), β measures the impact of the neutral reminder relative to the control group. The effect of the framed reminders is given by $\beta + \gamma_k$, $k \in \{1, 2, 3, 4\}$. Each γ -coefficient thus captures the differences between the neutral and a framed reminder.

Columns (1)–(5) of Table 2 report estimation results from linear probability models.¹⁶ Specification (1) estimates the average impact of reminders on patients’ response rates. The estimate shows that the overall difference in response rates is sizeable and highly significant ($p < 0.001$). Specification (2) includes treatment dummies for the reminders that contain additional information, providing us with estimates on the short-run effects of the individual reminder treatments. This estimation yields two important basic results. First, compared to the control treatment, response rates are significantly higher in each of the reminder treatments (F -tests for significance of $\beta + \gamma_k$: $p < 0.001$ for *Neutral* and Pos^F ; $p = 0.031$ for Pos^M ; $p = 0.011$ for Neg^F ; $p = 0.009$ for Neg^M). Second, as already illustrated in Figure 2, all of the treatments with framed reminders display a slightly weaker increase in response rates than the neutral treatment. These differences to the neutral reminder, as well as all other pairwise differences between the individual reminder treatments turn out to be insignificant.

The two basic results from column (2) hardly change when we include dummies for the randomization waves (column 3) and controls for a basic set of background characteristics (column 4). Specification (5) includes the full set of controls (which slightly reduces the sample size due to missing observations). The point estimate in this specification indicates that the neutral reminder increases the response rate by 14.5 percentage points relative to

¹⁶To avoid complications with the interpretation of marginal effects from interaction terms in non-linear models (Ai and Norton 2003), we focus on LPM estimations. Results from probit estimations, which yield very similar results, are available from the authors. For all estimations, we report robust standard errors that account for potential clustering at the household level.

Dependent variable:	Response within 30 days					Actual check-ups (30 days)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reminder	0.105*** [0.024]	0.142*** [0.038]	0.133*** [0.038]	0.129*** [0.038]	0.145*** [0.040]	0.103*** [0.022]	0.139*** [0.036]	0.130*** [0.036]	0.126*** [0.036]	0.140*** [0.038]
Rem × Pos ^F		-0.011 [0.044]	-0.010 [0.043]	-0.017 [0.042]	-0.031 [0.044]		-0.003 [0.043]	-0.002 [0.042]	-0.007 [0.041]	-0.023 [0.043]
Rem × Pos ^M		-0.067 [0.041]	-0.065 [0.041]	-0.059 [0.041]	-0.074* [0.042]		-0.068* [0.040]	-0.066* [0.039]	-0.059 [0.040]	-0.075* [0.041]
Rem × Neg ^F		-0.057 [0.040]	-0.054 [0.040]	-0.051 [0.040]	-0.072* [0.043]		-0.048 [0.038]	-0.045 [0.038]	-0.042 [0.038]	-0.061 [0.041]
Rem × Neg ^M		-0.052 [0.043]	-0.051 [0.042]	-0.054 [0.042]	-0.064 [0.045]		-0.062 [0.040]	-0.060 [0.040]	-0.062 [0.040]	-0.074* [0.043]
Male				-0.021 [0.023]	-0.007 [0.024]				-0.010 [0.022]	0.005 [0.022]
Age				0.003*** [0.001]	0.001 [0.001]				0.003*** [0.001]	0.002* [0.001]
Family				0.004 [0.037]	0.008 [0.038]				0.018 [0.036]	0.021 [0.037]
Private HI				-0.028 [0.029]	-0.023 [0.028]				-0.016 [0.028]	-0.015 [0.027]
Distance				-0.041** [0.019]	-0.046*** [0.009]				-0.033* [0.019]	-0.042*** [0.008]
Patient retention				0.007 [0.005]	0.007 [0.005]				0.005 [0.005]	0.005 [0.005]
At-Risk				-0.001 [0.034]	-0.018 [0.035]				-0.011 [0.031]	-0.026 [0.032]
Regular					0.148*** [0.025]					0.118*** [0.024]
Pain					-0.025 [0.027]					-0.041 [0.025]
Rental price					0.030*** [0.009]					0.025*** [0.008]
Constant	0.089*** [0.020]	0.089*** [0.020]	0.126*** [0.039]	0.024 [0.050]	-0.220** [0.089]	0.073*** [0.019]	0.073*** [0.019]	0.123*** [0.038]	0.020 [0.047]	-0.168* [0.086]
Wave dummies	no	no	yes	yes	yes	no	no	yes	yes	yes
N	1,227	1,227	1,227	1,227	1,095	1,227	1,227	1,227	1,227	1,095

Table 2: *Short-run treatment effects*

All specifications are estimated with a linear probability model. In specifications (1)–(5), the dependent variable indicates whether a patient contacted the dentist within 30 days. In specification (6)–(10), the dependent variable indicates whether the arranged check-up was actually attended. Robust standard errors, clustered at the household-level, are reported in brackets. ***, ** and * indicates significance at a 1%-, 5%-, and 10%-level, respectively.

the control group. The increase in response rates is again somewhat less pronounced for all reminders with additional information. For the Pos^M and the Neg^F treatment the difference to the *Neutral* treatment is statistically significant ($p < 0.1$). Relative to the untreated control group, however, all framed reminders still display a significantly positive effect on the short-run response rate ($p = 0.003$ for Pos^F , $p = 0.053$ for Pos^M , $p = 0.051$ for Neg^F , $p = 0.033$ for Neg^M).¹⁷

¹⁷Consistent with earlier evidence on socio-demographic correlates of dental health prevention (e.g., Lang et al. 1994), specifications (4) and (5) of Table 2 also document that several patient characteristics influence response behavior.

We next turn to the duration analysis that more rigorously examines the data underlying Figure 3. Using a Cox proportional hazard model, we estimate the equation

$$h(t) = h_0(t) \exp(\beta_1 \textit{Neutral} + \beta_2 \textit{Pos}^F + \beta_3 \textit{Pos}^M + \beta_4 \textit{Neg}^F + \beta_4 \textit{Neg}^M + X\delta + W\phi). \quad (4)$$

Estimated hazard ratios from different specifications of the model are reported in columns (1)–(5) of Table 3. Column (1) shows that patients who received a reminder exhibit on average a 42% higher hazard rate than the control group. Starting with column (2), we estimate separate effects for the different reminder treatments. In line with the results for the short-run responses, the estimates document that all reminders increase the patients’ likelihood to arrange an appointment. The hazard ratio for *Neutral*, the treatment with the strongest effect, is roughly 1.5. A similar effect is found for the *Pos*^F treatment. For the other treatments, the coefficients imply that the chance of contacting the dentist (conditional on not having done so before) is increased by a factor of 1.1–1.4 relative to the control treatment. Except for the *Pos*^M treatment, the effects for the individual reminders reported in column (2) are statistically significant. When we include wave dummies and further controls (columns 3–5) the estimated hazard ratios slightly decrease, but the point estimates remain well above one in all cases. Comparing the estimates for the individual reminder treatments against each other, we again find only moderate differences between the different reminders. The only pairwise treatment comparison which indicates a significant difference in hazard ratios is for *Neutral* vs. *Pos*^M in specifications (2) and (3) of Table 3 ($p = 0.088$ and $p = 0.091$, respectively).

The design of our experiment also allows us to examine whether there are systematic differences in the responses to a positive or negative framing of the reminders’ content, or in the response to the ‘gender framing’ of the postcard (compare, e.g., Tversky and Kahneman 1981, Landry et al. 2006, Bertrand et al. 2010). To do so, we estimate two sets of alternative model specifications in which we compare the pooled treatment effects from both positively and both negatively framed postcards, as well as the postcards with female and male covers. The results, which are reported in Tables A.3 and A.4 in the appendix, indicate that applying gain vs. loss frames does not significantly affect short- or long-run responses in our setup. Similarly, while the reminders with female covers trigger somewhat higher response rates than those with male covers, the difference between the two types of reminders is insignificant in most specifications.

<i>Dependent variable:</i>	Duration to response					Duration to actual check-up				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reminder	1.418** (0.019)					1.473** (0.016)				
Neutral		1.581** (0.014)	1.506** (0.031)	1.488** (0.032)	1.542** (0.027)		1.629** (0.015)	1.527** (0.037)	1.505** (0.042)	1.528** (0.043)
Pos ^F		1.563** (0.014)	1.479** (0.035)	1.401* (0.068)	1.486** (0.041)		1.700*** (0.005)	1.589** (0.017)	1.503** (0.035)	1.544** (0.030)
Pos ^M		1.161 (0.449)	1.111 (0.597)	1.116 (0.577)	1.135 (0.537)		1.161 (0.484)	1.097 (0.665)	1.106 (0.633)	1.113 (0.624)
Neg ^F		1.428** (0.048)	1.365* (0.092)	1.351* (0.098)	1.360 (0.105)		1.534** (0.024)	1.447* (0.060)	1.432* (0.063)	1.409* (0.085)
Neg ^M		1.371* (0.086)	1.315 (0.141)	1.261 (0.208)	1.269 (0.208)		1.367 (0.114)	1.293 (0.201)	1.244 (0.269)	1.202 (0.364)
Wave dummies	no	no	yes	yes	yes	no	no	yes	yes	yes
Controls	no	no	no	yes ⁱ	yes ⁱⁱ	no	no	no	yes ⁱ	yes ⁱⁱ
N	1,227	1,227	1,227	1,227	1,095	1,227	1,227	1,227	1,227	1,095

Table 3: *Duration analysis (Cox proportional hazard model)*

The table reports hazard ratios from estimations based on Cox proportional hazards model. In specifications (1)–(5), the dependent variables is the ‘duration’ to a patient’s response (i.e., the time between the treatment date and the date at which a patient contacted the dentist, measured for up to 100 days after the treatment intervention). Specifications (6)–(10) use the ‘duration’ until the actual check-up (if it was attended). *P*-values based on robust standard errors, clustered at the household-level, are reported in parentheses. Specification (4) and (9) include the restricted, specifications (5) and (10) the full set of control variables as reported in column (4) and (5) of Table 2, respectively. ***, ** and * indicates significance at a 1%-, 5%-, and 10%-level, respectively.

To wrap up, our results so far demonstrate that reminder messages lead to a sizeable and robust increase in the number of patients who schedule check-ups. In contrast to the strong overall impact of being reminded, the specific content and framing of the reminder seems to be less relevant. Hence, being reminded *per se* is of first-order importance, whereas additional information on the benefits of prevention does not further increase patients’ response rates relative to a neutral reminder. This observation is important, as it indicates that raising patients’ awareness about the benefits of dental prevention – the α -channel in the framework from Section 2.2 – seems to play a minor role for the reminder effects in our setup. Instead, our evidence suggests that reminders work because they address limitations in patients’ memory and increase patients’ attention (π).

Arranging vs. attending a check-up

So far, we have focused on the impact of reminders on the propensity to arrange a check-up appointment. From a policy perspective, however, the crucial question is whether patients actually attend more check-ups. The positive impact of reminders could be partially or fully offset if patients in the reminder treatments are more likely to miss their appointments. In

a next step, we therefore analyze whether the increase in scheduled appointments due to reminders also translates into a higher number of realized check-ups.

The answer to this question is a clear ‘yes’. The frequency of missed appointments in the reminder treatments actually lies below the respective number for the *No-Reminder* treatment: conditional on having arranged an appointment within a month, 17.6% of patients in the control group, but only 10.0% of patients in the reminder treatments miss their check-up date.¹⁸ While the difference is not statistically significant ($p = 0.423$), the numbers indicate that the reminder effects documented above are not diluted by missed appointments. This is further corroborated in columns (6)–(10) of Table 2, which replicate the estimates of the short-run treatment effects using an indicator for actual check-up attendance as dependent variable. The estimated coefficients are essentially identical to those of the basic response estimations (compare columns 1–5). If anything, the relative impact of reminders is even stronger when we account for missed appointments: the estimates in column (6) document that the overall fraction of patients who contact the dentist within a month *and* actually show up at their check-up increases from 7.3% in the control to 17.6% for the treated group ($p < 0.001$). Controlling for wave fixed effects and patient characteristics does not alter this result.

The picture does not change if we look at a longer response period of 100 days. Missed appointments again occur less frequently in the reminder treatments than in the *No-Reminder* condition (9.7% vs. 14.0%, $p = 0.460$). Hence it is not surprising that the duration analysis for actual check-up attendance, reported in columns (6)–(10) of Table 3, yields qualitatively similar but slightly stronger results than the ones for arranged appointments. The specification in column (6) shows that the hazard ratio for actual check-ups in the treated group is now 1.47 (as compared to 1.42 in column 1). Comparing columns (7)–(10) with columns (2)–(5) of Table 3, we find that the three treatments with the strongest impact on patients’ behavior display marginally higher (*Neutral*) and more precisely estimated (Pos^F , Neg^F) hazard ratios. The estimates for the two treatments that exhibit weaker effects hardly change (Pos^M , Neg^M). Note further that all hazard ratios are still above one for the alternative dependent variable. Moreover, the comparison of check-up attendance between the different reminder treatments once more confirms our earlier results on differences in patients’ response rates: almost all

¹⁸A possible interpretation of this finding is that – in addition to pointing forgetful patients to arrange a new check-up – reminders might also serve as an ‘*aide memoire*’ that reminds these patients about the upcoming appointment.

of the pairwise differences between the individual reminder treatments are insignificant. The only exceptions are the *Neutral* and *Pos^F* treatment, which differ from the *Pos^M* treatment at the 10% level in specifications (7) and (8) of Table 3.

In sum, the results on check-up attendance corroborate our findings on how patients' respond to reminders. In particular, we observe that reminders do not come at the cost of increasing the rate of missed appointments. Our evidence thus shows that reminders indeed succeed in encouraging people to make more check-ups.

4.2 Does one reminder 'fit all'?

We next turn to the question whether certain groups of patients are especially responsive to reminders, and whether particular types of reminders have stronger or weaker effects in different subgroups of the population. From a policy perspective, this analysis is interesting as it provides insights on whether targeting different groups of patients with type-specific reminders can increase their overall effectiveness. To address this question, we estimate models that interact treatment effects with patients' background characteristics. We analyze treatment interactions with respect to each patient's gender, age, health insurance, and risk status. In addition, we consider the rental prices in a patient's neighborhood as well as a patient's past check-up attendance and exposure to painful dental treatments. To keep the exposition tractable, we focus on short-run responses and pool the observations from the treatments with positive and negative framing, respectively.¹⁹

Results from the interacted models are reported in Table 4. Column (1) presents the treatment effects from the non-interacted model. All further specifications display the interaction effects with a different background characteristic (D). The coefficients thus allow to compare the different treatment effects between subgroups of patients that do or do not exhibit the corresponding characteristic D . The estimates in column (2), for instance, indicate that the neutral reminder increases responses among female patients by 13.2 percentage points. For males the corresponding effect is insignificantly lower by 0.4 percentage points.

The results in Table 4 reinforce our findings on the effectiveness of the neutral reminder. For almost all subgroups, this reminder triggers qualitatively the highest response rate. There is not a single group for which the reminders with additional information significantly out-

¹⁹The corresponding estimations studying interactions with all individual treatments, actual check-up attendance, or hazard-rate models with a longer time horizon again yield qualitatively similar results. Results are available from the authors upon request.

<i>Dependent variable: Response within 30 days</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D =	–	Male	Age	Priv.HI	At-Risk	Pain	Regular	Rent
Reminder	0.129*** [0.038]	0.132*** [0.048]	0.130*** [0.038]	0.164*** [0.040]	0.125*** [0.041]	0.116** [0.046]	0.107** [0.048]	0.146*** [0.040]
Rem × D		-0.004 [0.073]	0.000 [0.003]	-0.182* [0.097]	0.005 [0.101]	0.138 [0.095]	0.079 [0.080]	0.097*** [0.029]
Rem × Pos	-0.038 [0.036]	-0.027 [0.046]	-0.039 [0.036]	-0.042 [0.040]	-0.058 [0.039]	-0.018 [0.042]	-0.076* [0.044]	-0.054 [0.038]
Rem × Pos × D		-0.027 [0.069]	0.002 [0.003]	0.036 [0.090]	0.113 [0.097]	-0.171* [0.095]	0.052 [0.079]	-0.039 [0.029]
Rem × Neg	-0.052 [0.036]	-0.033 [0.045]	-0.053 [0.036]	-0.058 [0.040]	-0.064 [0.039]	-0.032 [0.044]	-0.095** [0.045]	-0.070* [0.039]
Rem × Neg × D		-0.049 [0.068]	0.000 [0.003]	0.047 [0.089]	0.086 [0.099]	-0.171* [0.093]	0.056 [0.078]	-0.051* [0.028]
D		0.007 [0.043]	0.002 [0.002]	0.093 [0.062]	-0.078 [0.058]	-0.027 [0.048]	0.045 [0.046]	-0.022 [0.016]
Controls	yes ⁱ	yes ⁱ	yes ⁱ	yes ⁱ	yes ⁱ	yes ⁱⁱ	yes ⁱⁱ	yes ⁱⁱ
N	1,227	1,227	1,227	1,227	1,227	1,095	1,095	1,095

Table 4: *Short-run treatment effects: Interaction with patient background characteristics*

The table presents the interaction of treatment effects with different background characteristics (captured in variables D). All specifications include a full set of wave dummies and a constant; specifications (1)–(5) [(6)–(8)] include the restricted [full] set of control variables. In specifications (3) and (8), the interaction variables (age and rental prices) are centered around their respective mean values. Robust standard errors, clustered at the household-level, are reported in brackets. ***, ** and * indicates significance at a 1%-, 5%-, and 10%-level, respectively.

perform the neutral reminder. In some subgroups, response rates for the framed reminders even fall significantly below those of the *Neutral* treatment. For instance, for patients who had a painful treatment experience in the past (column 6) as well as for those who did not regularly attend past check-ups (column 7), the framed reminders erode the positive effects of the neutral treatment, bringing the response rate almost down to the level of the untreated control group.²⁰ Column (4) of Table 4 further reveals that there is one group of patients who do not react to reminders. Among patients covered by private health insurance, none of the reminders induces a significant increase in check-up appointments. The absence of a reminder effect for this subgroup might be attributable to differences in insurance conditions (e.g., the typically higher costs of check-ups in private health plans), which in turn could lead to more deliberate planning of check-up frequencies.

²⁰A possible interpretation of these findings is that reminders with more ‘personal’ cover illustrations might be more likely to bring back memories of painful experiences or fears.

Overall, our analysis reveals only a few significant interactions between patients’ background characteristics and the impact of different reminders.²¹ A unified reminder might therefore be preferable to (costly) targeting of specific subgroups with different types of reminders. Moreover, there are no subgroups where adding information increases the effectiveness of the reminders. This again indicates that mitigating unawareness about the benefits of prevention is of limited relevance in our context.

4.3 Being ‘nudged’ more than once

The final step of our analysis turns to the question whether reminder effects depend on a patient’s past exposure to the nudge. Over the course of our experiment, several people in our sample were treated more than once. For these patients, random treatment assignment generates different sequences of treatments. In particular, among the group of patients who are up for a check-up at two points in time during our experiment, at T1 and at T2, we can compare the following four treatment sequences: (1) Patients who neither received a reminder in T1 nor in T2 (i.e., patients who are in the control group twice), (2) patients who received a reminder in T1 and in T2 ($Rem_{T1} \times Rem_{T2}$), (3) patients who were untreated in T1, but were reminded in T2 ($NoRem_{T1} \times Rem_{T2}$), and finally (4) patients who received a reminder in T1, but not in T2 ($Rem_{T1} \times NoRem_{T2}$).

To analyze whether these different sequences generate differential response patterns after the *second* treatment intervention, we estimate the model

$$Resp_{T2}^{30} = \lambda_0 + \lambda_1 Rem_{T1} \times Rem_{T2} + \lambda_2 NoRem_{T1} \times Rem_{T2} + \lambda_3 Rem_{T1} \times NoRem_{T2} + \nu, \quad (5)$$

where $Resp_{T2}^{30}$ is an indicator for a response within 30 days after the second intervention (at T2). As an alternative outcome variable we also consider a dummy for actual attendance at a check-up. The coefficients λ_1 and λ_2 measure the reminder effects at T2 relative to the group which remains untreated twice, for patients who did (λ_1) or did not (λ_2) receive a reminder in T1. Comparing these coefficients allows us to assess whether reminders have a stronger effect when they are applied twice (which would imply $\lambda_1 > \lambda_2$) or whether repeatedly nudging

²¹This also holds for alternative specifications of the interacted models. For instance, we find little systematic differences in how the various subgroups react to reminders with female vs. male cover framings. A notable exception concerns the ‘gender match’ between the patient and the postcard cover. Complementing earlier evidence on ‘gender match’ effects in other domains (e.g., Landry et al. 2006, Bertrand et al. 2010), we observe that both genders react similarly to postcards with neutral or female covers, but that the reminder effect for male postcards is particularly weak (and turns insignificant) among male patients.

people has a weaker impact ($\lambda_1 < \lambda_2$). By comparing λ_1 and λ_2 , we can therefore assess whether reminders work mostly through a ‘surprise’ or a ‘habituation’ effect, or whether the effect is independent of a patient’s history of receiving reminders.

The coefficient λ_3 captures the behavior in the group that does not receive a reminder in T2, after having received one in T1. An estimate of $\lambda_3 < 0$ would indicate that ‘taking away’ the reminder has detrimental effects, in the sense that patients’ response rate falls below the one of the reference group that was neither reminded in T1 nor in T2. This could be the case, for instance, if people get used to being nudged and rely too heavily on receiving the next reminder. We thus explicitly address the concern that nudges can not easily be taken away – a point that has been raised by critics of libertarian paternalism (e.g., Rizzo and Whitman 2009).

The results from estimating equation (5) are reported in columns (1)–(4) of Table 5. Columns (5)–(8) consider the effect on actual check-ups.²² The estimates again document highly significant reminder effects. Column (1) shows that the reminders have a slightly weaker effect when applied repeatedly. Relative to the group that is untreated in both T1 and T2, reminders increase the response by 16 percentage points for those who are treated for the second time, while the increase is 21 percentage points for the group that receives their first reminder in T2. While the sample size is limited and the difference in coefficients is not statistically significant ($p = 0.573$), the results suggest that habituation seems of limited importance: if anything, the impact of receiving a second nudge in a row is slightly weaker than the impact of a first nudge after an untreated period.

The weakly positive coefficient for λ_3 reinforces this impression. It indicates that taking away reminders does *not* lead to a drop in the likelihood of making a check-up, relative to the group of people who had not been exposed to reminders in T1 or T2. Rather, patients who do not receive a reminder in T2 have very similar response rates, irrespective of whether they had or had not received a nudge in T1. This indicates that the increase in check-ups triggered by the reminders does not come at the potential cost of a decline in active patient responses, once the nudge is taken away.

²²Note that it would be misleading to directly compare the point estimates with those from Table 2, as the results from Table 5 are based on a different sample. Note further that, in T2, we again randomize within a selected sample since patients who arrange a check-up in time will not be part of the experiment, paralleling the situation in T1 (see Section 3).

<i>Dependent variable:</i>	Response after second treatment (T2)				Check-up after second treatment (T2)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rem _{T1} × Rem _{T2}	0.158***	0.144***	0.218***	0.227*	0.143***	0.135***	0.204**	0.201*
(λ ₁)	[0.050]	[0.047]	[0.082]	[0.119]	[0.050]	[0.048]	[0.081]	[0.117]
NoRem _{T1} × Rem _{T2}	0.207**	0.198**	0.249***	0.274**	0.208**	0.203**	0.253***	0.278**
(λ ₂)	[0.081]	[0.082]	[0.091]	[0.137]	[0.082]	[0.082]	[0.090]	[0.134]
Rem _{T1} × NoRem _{T2}	0.027	0.008	0.091	0.085	0.034	0.024	0.100	0.089
(λ ₃)	[0.053]	[0.051]	[0.090]	[0.126]	[0.055]	[0.052]	[0.090]	[0.124]
Response _{T1}		0.126**	0.128**	0.010		0.068	0.073	-0.043
		[0.060]	[0.058]	[0.064]		[0.056]	[0.055]	[0.059]
Controls	no	no	yes ⁱ	yes ⁱⁱ	no	no	yes ⁱ	yes ⁱⁱ
N	392	392	392	375	392	392	392	375

Table 5: *The effect of repeated treatment.*

All specifications are estimated with a linear probability model and include a full set of wave dummies. The dependent variable is an indicator for a response within 30 days after the second treatment (columns 1–4). Columns (5)–(8) consider actual check-ups. The sample is composed of patients that experienced a sequence of two treatments. Specification (3) and (7) [(4) and (8)] include the restricted [full] set of control variables. Robust standard errors, clustered at the household-level, are reported in parentheses. ***, ** and * indicates significance at a 1%-, 5%-, and 10%-level, respectively.

One might be concerned that the estimates are influenced by whether or not a patient made a check-up in T1, which in turn is affected by the treatment in T1. To account for this potential concern, column (2) includes an indicator for patients' behavior in T1 ($Response_{T1}$). The point estimate for $Response_{T1}$, however, should be treated with caution, as the dummy is likely correlated with the error term in equation (5). The coefficients for λ_1 , λ_2 , and λ_3 are hardly affected when we include this new variable. Qualitatively, this also holds when we add further control variables in Columns (3) and (4). The point estimates for λ_1 and λ_2 increase but their difference remains insignificant. Similarly, the point estimate for λ_3 remains positive in all specifications. Paralleling our earlier results, all effects remain qualitatively unchanged if we consider actual check-up attendance after the second treatment (columns 5–8 of Table 5).

Extending the analysis of repeated exposure to reminders, we also studied whether differences in the sequences of receiving (i) neutral or framed reminders, (ii) positively or negatively framed reminders, and (iii) male or female framings matter for the observed treatment effects. In addition, we tested whether there are 'alternation effects', in the sense that receiving different reminders in repeated treatment sequences strengthens or weakens the reminder effect. Our analysis did neither reveal any further sequence effects nor any evidence on an effect from alternating the nudges. One has to bear in mind, however, that the smaller sample size limits the power of this analysis.

5 Conclusions

Using random treatment assignment, we analyzed how reminders influence patients' likelihood of making dental check-ups. Our results document strong and persistent effects of reminders on patient behavior. Importantly, the effect of reminders does not come at the cost of a decline in active patient responses, once the nudge is taken away. From a public policy perspective, our results thus demonstrate that reminders can be used as an unobtrusive and low-cost instrument for encouraging health prevention. Our data further indicate that simple nudges seem to work best: being reminded *per se* seems to be of first-order importance, whereas including additional information and framing does not further increase the frequency of check-ups relative to a neutral reminder. Given that this observation generalizes to other domains, it casts doubt on the effectiveness of negatively framed information – such as warning messages on tobacco packages – in the regulation of (un)healthy behaviors.

The fact that providing additional information is ineffective also indicates that limited awareness about the benefits of prevention seems to play a minor role in our context. Instead, the data suggest that our reminders primarily work because they address limitations in patients' memory and attention. In order to precisely quantify the relative importance of the different mechanisms at work, a full-fledged structural assessment – as suggested in Card et al. (2011) – seems a promising avenue for future research.

References

- ABELER, J. AND F. MARKLEIN (2010): “Fungibility, Labels, and Consumption,” *CeDEx Discussion Paper No. 2010-13, University of Nottingham*.
- ADENA, M., S. HUCK, AND I. RASUL (2014): “Charitable Giving and Nonbinding Contribution-Level Suggestions - Evidence from a Field Experiment,” *Review of Behavioral Economics*, 1, 275–293.
- AI, C. AND E. C. NORTON (2003): “Interaction terms in logit and probit models,” *Economics Letters*, 80, 123–129.
- ALLCOTT, H. AND T. ROGERS (2014): “The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation,” *American Economic Review*, forthcoming.

- 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.
- BEHAVIOURAL INSIGHTS TEAM (2014): “Improving Attendance Reminder Systems: Preliminary Evidence from RCTs at London Hospitals,” *Presentation Slides, The Behavioural Insights Team*.
- BELOT, M., V. BHASKAR, AND J. VAN DE VEN (2012): “Beauty and the Sources of Discrimination,” *Journal of Human Resources*, 47, 851–872.
- 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.
- BOS, A., J. HOOGSTRATEN, AND B. PRAHL-ANDERSEN (2005): “Failed Appointments in an Othodontic Clinic,” *American Journal of Orthodontics and Dentofacial Orthopedics*, 127, 355–357.
- CADENA, X. AND A. SCHOAR (2011): “Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments,” *NBER Working Paper No. 17020*.
- CALZOLARI, G. AND M. NARDOTTO (2014): “Nudging with Information: A Randomized Field Experiment,” *Working Paper, University of Bologna*.
- CAN, S., T. MACFARLANE, AND K. D. O’BRIEN (2003): “The Use of Postal Reminders to Reduce Non-attendance at an Orthodontic Clinic: a Randomised Controlled Trial,” *British Dental Journal*, 195, 199–201.
- CARD, D., S. DELLA VIGNA, AND U. MALMENDIER (2011): “The Role of Theory in Field Experiments,” *Journal of Economic Perspectives*, 25, 39–62.
- DELLA VIGNA, S. (2009): “Psychology and Economics: Evidence from the Field,” *Journal of Economic Literature*, 47, 315–372.
- ELIAZ, K. AND R. SPIEGLER (2011): “Consideration Sets and Competitive Marketing,” *Review of Economic Studies*, 78, 235–262.
- GALLAGHER, K. M. AND J. A. UPDEGRAFF (2012): “Health Message Framing Effects on Attitudes, Intentions, and Behavior: A Meta-analytic Review,” *Annals of Behavioral Medicine*, 43, 101–116.
- GLIED, S. AND M. NEIDELL (2010): “The Economic Value of Teeth,” *Journal of Human Resources*, 45, 468–496.

- JOHNSON, E. J. AND D. GOLDSTEIN (2003): “Do Defaults Save Lives?” *Science*, 302, 1338–1339.
- KARLAN, D., M. MCCONNELL, S. MULLAINATHAN, AND J. ZINMAN (2013): “Getting to the Top of Mind: How Reminders Increase Saving,” *Working Paper, Harvard University*.
- KARLAN, D., M. MORTEN, AND J. ZINMAN (2012): “A Personal Touch: Text Messaging for Loan Repayment,” *Working Paper, Dartmouth College*.
- KAST, F., S. MEIER, AND D. POMERANZ (2012): “Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device,” *IZA Discussion Paper No. 6311*.
- LANDRY, C. E., A. LANGE, J. A. LIST, M. K. PRICE, AND N. G. RUPP (2006): “Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment,” *Quarterly Journal of Economics*, 121, 747–782.
- (2010): “Is a Donor in Hand Better Than Two in the Bush? Evidence from a Natural Field Experiment,” *American Economic Review*, 100, 958–83.
- 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.
- LOEWENSTEIN, G., T. BRENNAN, AND K. G. VOLPP (2007): “Asymmetric Paternalism to Improve Health Behaviors,” *Journal of the American Medical Association*, 298, 2415–2417.
- MACHARIA, W. M., G. LEON, B. H. ROWE, B. J. STEPHENSON, AND R. B. HAYNES (1992): “An Overview of Interventions to Improve Compliance With Appointment Keeping for Medical Services,” *JAMA*, 267, 1813–1817.
- MCDOWELL, I., C. NEWELL, AND W. ROSSER (1989): “A Randomized Trial of Computerized Reminders for Blood Pressure Screening in Primary Care,” *Medical Care*, 27, 297–305.
- METTES, D. (2005): “Insufficient Evidence to Support or Refute the Need for 6-monthly Dental Check-ups,” *Evidence-Based Dentistry*, 6, 62–63.
- MIGUEL, E. AND M. KREMER (2004): “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica*, 72, 159–217.
- MILKMAN, K., J. BESHEARS, J. CHOI, D. LAIBSON, AND B. MADRIAN (2011): “Using Implementation Intentions Prompts to Enhance Influenza Vaccination Rates,” *PNAS*, 108, 10415–10420.

- MOCAN, N. AND E. TEKIN (2010): “Ugly Criminals,” *Review of Economics and Statistics*, 92, 15–30.
- NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE (2004): “Dental Recall – Recall Interval Between Routine Dental Examinations,” *Clinical Guidelines*, CG19.
- O’KEEFE, D. J. AND J. D. JENSEN (2009): “The Relative Persuasiveness of Gain-Framed and Loss-Framed Messages for Encouraging Disease Detection Behaviors: A Meta-Analytic Review,” *Journal of Communication*, 59, 296–316.
- OULLIER, O., R. CIALDINI, R. H. THALER, AND S. MULLAINATHAN (2010): “Improving Public Health Prevention with a Nudge,” in *Improving Public Health Prevention with Behavioural, Cognitive and Neuroscience*, ed. by O. Oullier and S. Sauneron, Paris: Centre for Strategic Analysis, 38–46.
- RABIN, M. (1999): “Psychology and Economics,” *Journal of Economic Literature*, 36, 11–46.
- REEKIE, D. AND H. DEVLIN (1998): “Preventing Failed Appointments in General Dental Practice: A Comparison of Reminder Methods,” *British Dental Journal*, 185, 472–474.
- RIZZO, M. J. AND D. G. WHITMAN (2009): “Little Brother Is Watching You: New Paternalism on the Slippery Slopes,” *Arizona Law Review*, 51, 685–739.
- THALER, R. AND C. SUNSTEIN (2008): *Nudge – Improving Decisions About Health, Wealth, and Happiness*, New Haven, CT: Yale University Press.
- TVERSKY, A. AND D. KAHNEMAN (1981): “The Framing of Decisions and the Psychology of Choice,” *Science*, 211, 1124–1131.
- WISDOM, J., J. S. DOWNS, AND G. LOEWENSTEIN (2010): “Promoting Healthy Choices: Information versus Convenience,” *American Economic Journal: Applied Economics*, 2, 164–178.

A. Appendix



Figure A.1: *Covers of reminder postcards.* Top row (left to right): *Neutral*, Pos^F , and Pos^M treatment. Bottom row (left to right): Neg^M and Neg^F treatment.

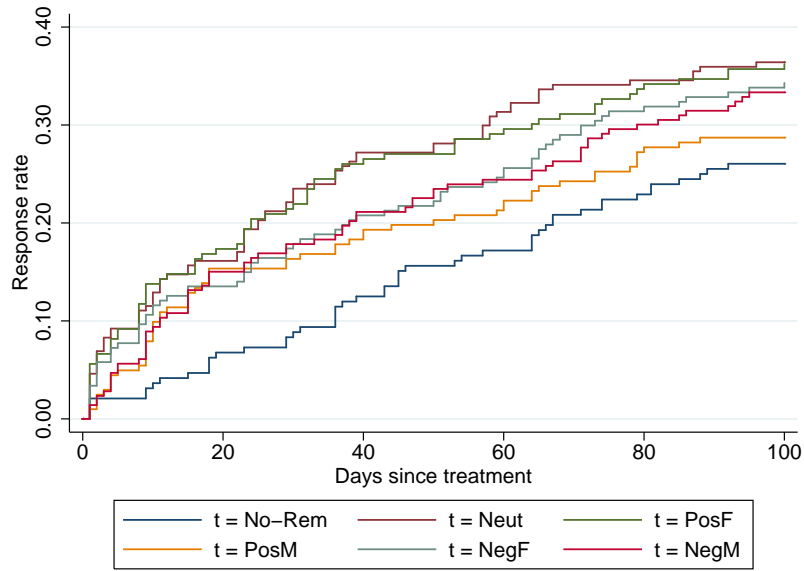


Figure A.2: *Cumulated response rates for individual treatments.*

Reference	Task	Sample	N	Treatments and procedures	Main findings
Apesteuguia et al. (2013)	Return items to library	Users of public libraries in Barcelona	50000	<ul style="list-style-type: none"> control email general reminder informing about rules rules + social appeal rules + information on past behavior rules + information on penalty Messages sent out twice via email (constant treatment assignment). Analyze responses up to 75 days after the intervention.	<ul style="list-style-type: none"> General reminder significantly reduces fraction of late returns. No additional effect of other treatments. Effect relatively short lived (2-4 weeks), but ATEs similar after second intervention. Study heterogeneous effects depending on past compliance, gender, age, and nationality.
Bos et al. (2005)	Attend appointment	Patients with appointments for surgery at an orthodontic clinic	343	<ul style="list-style-type: none"> control postal reminder telephone reminder SMS reminder One-time intervention, participants received reminder 1 day before appointment date.	<ul style="list-style-type: none"> Fraction of broken appointments and cancellation rates do not differ significantly across treatments.
Cadena and Schoar (2011)	Repay loan	Customers of microfinance lender in Uganda	1121	<ul style="list-style-type: none"> control reminder financial incentives (cash back vs. 25% interest rate reduction for next loan if all repayments are made in time) Reminder sent monthly via SMS for up to 12 months (constant treatment).	<ul style="list-style-type: none"> All treatments increase frequency of timely repayment and decrease average delay in repayments. No sig. differences across treatments. Cash back and reminders more effective for borrowers with smaller loans. Borrowers with larger loans respond more strongly to reductions in future interest rates.
Calzolari and Narotto (2014)	Visit health club	Members of university gym	247	<ul style="list-style-type: none"> control email reminder containing information and link to personalized web page of the gym Reminders sent weekly via email for up to 6 months (constant treatment). Measure responses up to 4 months after the intervention.	<ul style="list-style-type: none"> Reminder increases attendance rates. Effect turns insignificant in post-treatment period. Increase in attendance occurs immediately (within 24h) after receipt of reminder. Stronger effects for users with low baseline attendance rates in pre-treatment period.

Can et al. (2003)	Attend appointment	Patients with appointments at secondary care orthodontic clinic	231	<ul style="list-style-type: none"> • mail with appointment date (control) • mail with appointment date + postal reminder + confirmation slip <p>One-time intervention, reminder + confirmation were sent “at least 2 weeks” before scheduled appointment.</p>	<ul style="list-style-type: none"> • Fraction of broken appointments significantly lower for patients who received reminder <i>and</i> returned confirmation slip. • No significant differences between control group and patients who received reminder, but did not return confirmation.
Karlan et al. (2013)	Deposit money in savings account	Bank customers in Bolivia, Peru, and the Philippines	14017	<p>Control vs. reminders that vary</p> <ul style="list-style-type: none"> • timing (before and/or after scheduled deposit date) • framing (positive/negative) • mentioning of savings goal • mentioning of incentives for making regular deposits <p>Reminders sent monthly for up to 24 months via letter or SMS (constant treatment). Measure amount saved and achievement of savings goal at goal date.</p>	<ul style="list-style-type: none"> • Positive overall effect of reminders. • No significant influence of framing and timing. • Reminders that mention specific savings goal are more effective. • Reminders that mention financial incentives have somewhat stronger effects. • No interaction between reminder effects and survey-based measure of time inconsistency.
Karlan et al. (2012)	Repay loan	Customers of microfinance lender in the Philippines	943	<p>Control vs. reminders that vary</p> <ul style="list-style-type: none"> • framing (positive/negative), • timing (0/1/2 days before payment date), • mentioning of loan officer’s / client’s name <p>Reminders sent weekly via SMS for up to 15 weeks (constant treatment). Measure response up to 30 days past loan maturity.</p>	<ul style="list-style-type: none"> • No overall effect of reminders. • Positively framed reminder that mentions account officer’s name reduces frequency of late payments. No significant effects for other treatment cells. • No significant influence of timing.

Kast et al. (2012)	Deposit money in savings account	Micro-entrepreneurs in Chile	873	<p>Treatments (experiment 2):</p> <ul style="list-style-type: none"> • control • reminder about own achievement of savings goal and others' goal achievement • reminder about own goal achievement, sent to participant and to peer ("savings buddy") <p>Reminders sent weekly via SMS for up to 3 months (constant treatment). Measure responses up to 2 months after the intervention.</p>	<ul style="list-style-type: none"> • Reminders increase overall level and number of deposits during intervention period. • No significant differences between different reminder types. • No long-run impact of reminders after the end of the intervention.
Milkman et al. (2011)	Receive influenza vaccination	Employees at midwestern firm	3272	<ul style="list-style-type: none"> • mail with information about workplace vaccination clinic (control) • mail + prompt to plan vaccination date • mail + prompt to plan date and time <p>One-time mailing. Measure receipt of vaccination at workplace clinic (2-6 weeks after mailing). For subsample of participants, also measure responses up to 7 months after the experiment.</p>	<ul style="list-style-type: none"> • No significant difference between control and general planning prompt. • More specific prompt to plan date <i>and</i> time leads to significantly higher vaccination rates.
Reekie and Devlin (1998)	Attend appointment	Patients with appointments for surgery at a dental practice	2500	<ul style="list-style-type: none"> • control • reminder postcard • manual telephone reminder • automated telephone reminder • combination of all three reminders <p>One-time intervention. Postal reminders sent 3 days before scheduled appointment, calls 1 day before appointment.</p>	<ul style="list-style-type: none"> • All reminders lead to a significant reduction in failed appointments relative to control. • No significant differences in fraction of failed appointments for different types of reminders.

Table A.1: *Overview of field experiments on reminder effects.*

Framing	Text cover	Message text back
Neutral	time for dental prevention	Please make an appointment for your next check-up.
Positive	...keep your nice smile tomorrow	Investing some time in dental health prevention today decreases your risk of a painful dental disease in the future. In addition, you may avoid considerable costs of involved treatments. Please make an appointment for your next check-up.
Negative	...don't lose your nice smile tomorrow	Not investing some time in dental health prevention today increases your risk of a painful dental disease in the future. In addition, you may incur considerable costs of involved treatments. Please make an appointment for your next check-up.

Table A.2: *Text of reminder messages.*

Panel A						
<i>Dependent variable:</i>	Response within 30 days			Actual check-up (30 days)		
	(1)	(2)	(3)	(4)	(5)	(6)
Reminder	0.142*** [0.038]	0.129*** [0.038]	0.144*** [0.040]	0.139*** [0.036]	0.126*** [0.036]	0.140*** [0.038]
Rem × Pos-Frames	-0.039 [0.037]	-0.038 [0.036]	-0.053 [0.038]	-0.036 [0.036]	-0.034 [0.035]	-0.050 [0.037]
Rem × Neg-Frames	-0.054 [0.037]	-0.052 [0.036]	-0.068* [0.039]	-0.055 [0.035]	-0.052 [0.034]	-0.067* [0.037]
<i>Wald-Tests (p-Values):</i>						
Pos = Neg	0.593	0.601	0.588	0.482	0.489	0.529
Wave dummies	no	yes	yes	no	yes	yes
Controls	no	yes ⁱ	yes ⁱⁱ	no	yes ⁱ	yes ⁱⁱ
N	1,227	1,227	1,095	1,227	1,227	1,095

Panel B						
<i>Dependent variable:</i>	Response within 30 days			Actual check-up (30 days)		
	(1)	(2)	(3)	(4)	(5)	(6)
Reminder	0.142*** [0.038]	0.129*** [0.038]	0.144*** [0.040]	0.139*** [0.036]	0.126*** [0.036]	0.140*** [0.038]
Rem × F-Frames	-0.034 [0.037]	-0.034 [0.036]	-0.052 [0.038]	-0.026 [0.035]	-0.025 [0.034]	-0.042 [0.037]
Rem × M-Frames	-0.059 [0.037]	-0.056 [0.037]	-0.069* [0.039]	-0.065* [0.036]	-0.061* [0.036]	-0.074** [0.038]
<i>Wald-Tests (p-Values):</i>						
F = M	0.382	0.439	0.560	0.150	0.182	0.257
Wave dummies	no	yes	yes	no	yes	yes
Controls	no	yes ⁱ	yes ⁱⁱ	no	yes ⁱ	yes ⁱⁱ
N	1,227	1,227	1,095	1,227	1,227	1,095

Table A.3: *Short-run treatment effects – alternative specifications*

Panel A studies the treatment effects for positively vs. negatively framed reminders. Panel B compares responses to treatments with male vs. female postcard covers. All specifications are estimated with linear probability models and include a constant term. Column 2 and 5 include the restricted, column 3 and 6 the full set of control variables from Table 2. The table further includes the p -values from Wald tests of the hypothesis that the coefficients differ between the reminders with different types of framing. Robust standard errors, clustered at the household-level, are reported in brackets. ***, ** and * indicates significance at a 1%-, 5%-, and 10%-level, respectively.

Panel A						
<i>Dependent variable:</i>	Duration to response			Duration to check-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Neutral	1.581** (0.014)	1.489** (0.032)	1.543** (0.027)	1.629** (0.015)	1.506** (0.041)	1.530** (0.043)
Pos-Frames	1.353* (0.066)	1.257 (0.165)	1.306 (0.122)	1.418** (0.047)	1.304 (0.131)	1.326 (0.123)
Neg-Frames	1.399** (0.037)	1.305* (0.099)	1.316* (0.100)	1.448** (0.031)	1.336* (0.093)	1.307 (0.129)
<i>Wald-Tests (p-Values):</i>						
Neutral = Pos	0.303	0.257	0.298	0.385	0.363	0.398
Neutral = Neg	0.419	0.376	0.324	0.448	0.438	0.344
Pos = Neg	0.781	0.760	0.955	0.870	0.848	0.915
Wave dummies	no	yes	yes	no	yes	yes
Controls	no	yes ⁱ	yes ⁱⁱ	no	yes ⁱ	yes ⁱⁱ
N	1,227	1,227	1,095	1,227	1,227	1,095

Panel B						
<i>Dependent variable:</i>	Duration to response			Duration to check-up		
	(1)	(2)	(3)	(4)	(5)	(6)
Neutral	1.581** (0.014)	1.487** (0.032)	1.542** (0.027)	1.629** (0.015)	1.504** (0.042)	1.528** (0.043)
F-Frames	1.492** (0.012)	1.374** (0.049)	1.419** (0.039)	1.613*** (0.005)	1.466** (0.026)	1.473** (0.029)
M-Frames	1.268 (0.154)	1.190 (0.291)	1.205 (0.277)	1.266 (0.190)	1.178 (0.359)	1.160 (0.420)
<i>Wald-Tests (p-Values):</i>						
Neutral = F	0.699	0.590	0.594	0.949	0.866	0.818
Neutral = M	0.155	0.146	0.136	0.125	0.136	0.118
F = M	0.186	0.246	0.205	0.064	0.094	0.082
Wave dummies	no	yes	yes	no	yes	yes
Controls	no	yes ⁱ	yes ⁱⁱ	no	yes ⁱ	yes ⁱⁱ
N	1,227	1,227	1,095	1,227	1,227	1,095

Table A.4: *Duration analysis (Cox proportional hazard model) – alternative specifications*

Panel A studies the treatment effects for positively vs. negatively framed reminders. Panel B compares responses to treatments with male vs. female postcard covers. The table reports hazard ratios from estimations based on Cox proportional hazards model of responses and check-up attendance up to 100 days after the treatment intervention. p -values based on robust standard errors, clustered at the household-level, are reported in parentheses. The table further includes the p -values from Wald tests of the hypothesis that the hazard ratios differ between the different types of reminders. Specification (4) and (5) include the (i) restricted and (ii) the full set of control variables, respectively. ***, ** and * indicates significance at a 1%-, 5%- and 10%-level, respectively.