World Development 203 (2026) 107369

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Regular research article

Can reminders promote regular pro-environmental behavior? Experimental evidence from Peru

Hanna Fuhrmann-Riebel [a,*], Ben D'Exelle [b], Kristian López Vargas [c,d], Sebastian Tonke [e,f], Arjan Verschoor [g]

[a] German Institute of Development and Sustainability (IDOS), Tulpenfeld 6, Bonn, 53113, Germany
[b] University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK
[c] University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[d] Universidad del Pacífico, Jr. General, Jirón Luis Sánchez Cerro 2141, Lima, Jesús María, 15072, Peru
[e] Technical University of Munich, TUM School of Management and TUM School of Life Sciences, Alte Akademie 16, Freising, 85354, Germany
[f] Max Planck Institute for Behavioral Economics, Kurt-Schumacher-Straße 10, Bonn, 53113, Germany
[g] University of East Anglia & Norwich Institute for Sustainable Development, Norwich Research Park, Norwich, NR4 7TJ, UK

HIGHLIGHTS

- Mobile text reminders increase household recycling behavior in urban Peru.
- The effects are stronger when reminders are repeated over a longer period.
- Reminder effects emerge earlier among households that already recycled previously.
- New recyclers require prolonged reminder exposure to begin recycling.
- Reminders promote habits in pro-environmental behavior by mitigating limited attention.

ARTICLE INFO

JEL classification:
C93
D83
D90
D91
Q53
Q58

Keywords:
Reminders
Limited attention
Habit formation
Recycling
Peru

ABSTRACT

Pro-environmental behavior, such as recycling, often needs to be regular to be effective, and interventions to encourage behavioral change may therefore need to be repeated; yet, little evidence exists on the optimal time pattern and frequency of such repeated interventions. To fill this gap, we investigate the impact of mobile text reminders on households' recycling behavior in urban Peru by randomly varying the exposure length and continuity of reminders. We find that reminders increase both the likelihood that households start to recycle and the frequency of recycling among households that already did so before the intervention. The effects are stronger when reminders are repeated over a longer period. Our findings suggest that both limited attention and habit formation matter for recycling behavior, and that low-cost mobile text reminders can effectively support regular pro-environmental behavior.

1. Introduction

Increasing global waste accumulation poses a serious threat to ecosystems, biodiversity, and human health, and contributes significantly to environmental pollution and climate change (UNEP, 2024). While the waste crisis requires a collective response from both the public and the private sector, it also demands a permanent shift in households' regular behavior; household recycling behavior is a paradigmatic example of this much-needed behavioral change.[1] Weekly

* Corresponding author.
Email addresses: hanna.fuhrmann-riebel@idos-research.de (H. Fuhrmann-Riebel), B.DExelle@uea.ac.uk (B. D'Exelle), klopezva@ucsc.edu (K. López Vargas), sebastian.tonke@tum.de (S. Tonke), A.Verschoor@uea.ac.uk (A. Verschoor).

[1] As a result, identifying ways to encourage household recycling behavior has become a core interest of both researchers and policymakers over the last years (see, e.g., Kirakozian, 2016; Knickmeyer, 2020; Miafodzyeva & Brandt, 2013; Thomas & Sharp, 2013; Varotto & Spagnolli, 2017, for literature reviews and meta-analyses).

https://doi.org/10.1016/j.worlddev.2026.107369
Accepted 21 February 2026
Available online 5 March 2026

kerbside collection of recyclables, typical in many places around the world, requires people to organize their waste in the household in advance, collect materials throughout the week, and finally put the bags outside their house on the same day each week. Recycling remains limited in many low- and middle-income countries: while in Western Europe around 55% of municipal solid waste is recycled, in South America only around 5% is recycled (ibid.). Consequently, large amounts of waste end up in dumpsites, affecting people's health and the environment (Kaza et al., 2018).

Following the existing literature, in this paper we distinguish two critical behavioral aspects to understand what can be done to increase recycling behavior. First, people may fail to implement their recycling intentions because of *limited attention* (Datta & Mullainathan, 2014; Karlan et al., 2016; Tiefenbeck et al., 2018). Humans' limited cognitive capacity means that they can only pay limited attention to recycling, which requires substantial organization (such as separating recyclables, putting them in a bag, placing the bag on the street on the correct day and time, etc.). Second, *habit formation* may play a role: once people have recycled a few times, they may form a routine/habit in recycling practices, and the barrier of limited attention may weaken or disappear.

The issue of limited attention might be addressed by sending timely reminders that bring the specific behavior "to the top of mind" (Karlan et al., 2016).[2] However, it is unclear how effective reminders are when aiming to change *regular* behavior (as opposed to one-time or irregular actions). We may need to repeat reminders to overcome limited attention repeatedly. Yet, the effectiveness of reminders may also decrease with repetition due to habituation (Rankin et al., 2009), i.e., we may experience a weakened behavioral response due to repeated treatment. If so, interrupting the series of reminders (and possibly restarting it after a pause) may recover the behavioral response to the treatment. In addition, reminders may need to be repeated for habits to be formed, which, once established, will reduce the need for further reminders. However, here again it is *a priori* not clear how long people need to be treated for habits to be formed (Allcott & Rogers, 2014; Byrne et al., 2024; Charness & Gneezy, 2009; Royer et al., 2015).

This leads to two critical questions: 1) whether reminders can increase recycling, and 2) what time pattern of reminders can promote repeated recycling most effectively. Regarding the second question, we are especially interested in understanding i) whether a longer series of reminders is more effective than a shorter series, and ii) whether a continuous series of reminders is more effective than an interrupted and restarted series to encourage regular recycling.

To answer these questions, we conducted a field experiment with households that are registered in a recycling program in urban Peru. Many households do not recycle regularly, even though they voluntarily signed up for the recycling program. This suggests that there may be a gap between their intention to recycle and actual recycling behavior, and possibly also a lack of routine and habit formation. Assuming both could be addressed with the help of reminders, we sent mobile text reminders to households with varying exposure length and continuity to test their effect on recycling.

We collected data on households' recycling behavior over 12 weeks by accompanying the recycling trucks collecting the bags on their daily routes. Having the addresses of registered households allowed us to identify which households put a recycling bag on the street in which week, without needing to contact them. This approach has key advantages, as it avoids interference with the intervention (people were not made aware of our study) as well as the limitations of self-reported data.

While the first three weeks of our study period serve as a baseline measure (weeks 1–3), the subsequent nine weeks (weeks 4–12) constitute our intervention period where households are randomly assigned to one of four groups: i) a control group that does not receive any reminders, ii) a group that receives weekly reminders continuously during the whole intervention period, weeks 4–12 (continuous treatment), iii) a group that receives reminders only during the first three weeks of the intervention period, weeks 4–6 (interrupted treatment), iv) or a group that receives reminders for the first three weeks, weeks 4–6, and for the last three weeks, weeks 10–12, of the intervention period, with a three weeks' pause in between during which no reminders are received (restarted treatment).

Our main findings can be summarized as follows. First, we show that reminders are generally effective at increasing recycling behavior: weekly reminders in the first three weeks of the intervention increase the likelihood of recycling in all treatment groups compared to the control group. This initial effect seems to be mainly driven by households that already recycled at baseline. Second, we find that the effect becomes stronger if reminders are continued for another three weeks. Third, we observe that households that did not yet recycle at baseline might need a longer series of reminders to start recycling than households that already recycled before. Fourth, we find suggestive evidence that both limited attention and habit formation matter for recycling behavior and the effectiveness of reminders.

We contribute to the existing literature in three ways. First, there is a growing yet still limited literature on the effectiveness of repeated reminders in the context of regular behaviors, for example, concerning gym attendance (Calzolari & Nardotto, 2017; Muller & Habla, 2018), regular savings (Milkman et al., 2025; Rodríguez & Saavedra, 2019) or the reuse of plastic bags (Essl et al., 2021) and recycling (Chong et al., 2015).[3] These regular behaviors stand in contrast to infrequent or one-time actions, such as attending a doctor's appointment.[4] As outlined above, making this distinction is important, as beyond limited attention, habit formation can matter as well for repeated behaviors and the responsiveness to reminders. We contribute to this literature by generating new evidence on the role of the exposure length and continuity with which reminders are sent, since we deploy treatments with long (continuous), short (interrupted), and paused (restarted) reminder exposure. In contrast to previous studies, our design allows us to compare the

[2] The effectiveness of reminders has been demonstrated in a variety of settings, such as personal savings (Karlan et al., 2016; Milkman et al., 2025; Rodríguez & Saavedra, 2019), electricity consumption (Allcott & Rogers, 2014; Gilbert & Zivin, 2014), loan repayment (Cadena & Schoar, 2011), civil servant compliance (Dustan et al., 2023), gym attendance (Calzolari & Nardotto, 2017; Milkman et al., 2021a; Muller & Habla, 2018), participation in land conservation programs (Wallander et al., 2017), the adoption of agricultural technologies (Larochelle et al., 2019), donations to charity (Damgaard & Gravert, 2018; Huck & Rasul, 2010; Sonntag & Zizz, 2015), political participation (Grácio & Vicente, 2021), returning books on time to a library (Apesteguia et al., 2013), dental health prevention (Altmann & Traxler, 2014) or vaccination uptake (Milkman et al., 2021b), such as in the context of COVID-19 vaccinations (Dai et al., 2021).

[3] There are several other field experiments that use repeated reminders in combination with other intervention features to encourage regular pro-environmental behavior, such as the "home energy reports" encouraging energy conservation (Allcott, 2011; Allcott & Rogers, 2014; Andor et al., 2020; Bonan et al., 2020), "home waste reports" to reduce unsorted waste volumes (Bonan et al., 2025), or appeals to save or pay for piped water (Ferraro & Price, 2013; Rockenbach et al., 2025; Tonke, 2024, 2025). A key distinction from this literature is that these studies do not isolate the effect of reminders per se, but rather estimate the combined impact of reminders and complementary elements such as social comparisons. Consequently, it is typically not possible to disentangle whether behavioral changes arise from the reminder itself, from historical feedback on one's own behavior, from social norm comparisons, or from other components of the treatment. Similarly, recent papers that study the impact of real-time feedback on resource conservation present repeated reminders for regular behavior as well – albeit in a special way as the reminder is transmitted while people engage in the behavior, including feedback about current (and sometimes also past) consumption levels (Byrne et al., 2024; Fang et al., 2023; Tiefenbeck et al., 2018, 2019).

[4] Studies that generated evidence on the effectiveness of repeated reminders on one-time or infrequent decisions are Altmann and Traxler (2014) on dental health check-ups, Antinyan et al. (2021) on overdue tax payment or Sonntag and Zizz (2015) and Damgaard and Gravert (2018) on donations to charity.

effects of interrupted reminders (i.e., post intervention effects) to the effects of reminders that are continued during the same time period, or restarted, which extends evidence of looking at post-intervention persistence effects only, without having a comparison group.[5] A recent paper by Byrne et al. (2024) takes a similar approach by varying the treatment exposure duration in the context of real-time in-the-shower feedback on water usage in Australia. With our novel design, we can generate important new insights on the role of limited attention and habituation as well as their potential interplay for regular (recycling) behavior, as outlined in more detail in our conceptual framework.[6]

Second, we contribute to the so far limited evidence on the effectiveness of reminders on household recycling behavior in particular.[7] While recycling behavior constitutes a behavior that requires regular, repeated action, it is also characterized by the fact that it mirrors a collective action problem that requires individual action for collective benefit (see, e.g., Fuhrmann-Riebel et al. (2024)). This distinguishes the behavior from other regular behaviors studied in the reminder literature, which generate direct individual benefits, such as gym attendance or personal savings. Moreover, its measurement is considerably more labor-intensive than that of other behaviors studied in the literature so far (e.g., gym attendance, savings, or energy consumption), which can be measured digitally. The "hard to measure" nature of recycling behavior is a main reason why most studies on recycling behavior rely on self-reported data, which has several critical limitations (Huffman et al., 2014). There are few studies that test the impact of reminders on actual recycling behavior. Essl et al. (2021) found a positive effect of weekly reminders in the form of stickers and flyers on the reuse of plastic bags (i.e., plastic waste reduction) in the context of weekly food delivery boxes in Switzerland. Chong et al. (2015), which comes closest to our study, also focused on the effect of reminders on household recycling behavior in a recycling program in Peru. The authors did not find any significant impact of mobile text reminders on households' recycling activity. The ineffectiveness of their intervention was probably due to the lack of recycling equipment within the household (concretely, the absence of recycling bins), which, according to the authors, was the main obstacle for households to recycle. With an observation period of six weeks, their study period was also considerably shorter than ours. Our main contribution to this literature is that we do find that reminders can significantly increase households' recycling activity, supporting the authors' assumption that lacking recycling equipment was the main reason why their intervention failed. Also, with a considerably longer study period, we are able to provide novel evidence on the role of habit formation in recycling practices and how this might affect the effectiveness of reminders, which previous studies could not.

Third, we contribute to the growing body of research on behavioral interventions in developing countries where institutional and infrastructure constraints as well as different attitudes and responses to such interventions can play essential roles.[8] Recent meta-analyses on behavioral interventions find that only a small share of the current evidence comes from developing countries and is concentrated in a few policy areas (Hummel & Maedche, 2019; Mertens et al., 2022). Other reviews in behavioral development economics emphasize that experiments in developing countries often find strong patterns of attentional and self-control frictions, and that the underlying institutional and physical infrastructure as well as market failures might amplify or dampen the impact of nudges (Kremer et al., 2021). This is especially crucial for recycling behavior as very little academic work has been conducted on the topic in developing countries even though these are the regions that suffer most from inefficient waste management and the resulting environmental consequences (Chong et al., 2015). Our paper thus adds important new insights on the effectiveness of behavioral interventions in developing countries for a highly relevant behavior on which the evidence so far is very limited. A great benefit of our intervention is that sending SMS reminders has low implementation costs, which is especially important for studies conducted outside a high-income country context where financial resources are often limited.

The rest of the paper proceeds as follows. Section 2 presents the design of the experiment, including the conceptual framework, treatments, hypotheses, and implementation. In Section 3 we report the results, while Section 4 concludes the paper.

2. Design

2.1. Conceptual framework

Before presenting our experiment, it is important to conceptualize how we understand the role of reminders for recycling behavior. Recycling imposes an immediate, tangible hassle (rinsing containers, sorting waste, dragging a bin to the curb), but its environmental benefits arrive slowly in the future and favor a diffuse collective. Households that voluntarily enroll in a municipal program display an underlying pro-social cooperative motive. Yet, they remain exposed to cognitive frictions and biases that often derail low-stakes, cooperative tasks, such as recycling. They might be present-focused and might devote their scarce attention to whatever is most salient at the moment. With present bias, each collection cycle resembles a soft deadline that can always be met "later tonight or tomorrow"; with limited attention, the critical moment can pass unnoticed. The joint presence of procrastination and forgetting is damaging because these frictions might enhance one another.

If these negative psychological forces are overcome and the individual is effectively reminded to recycle (for the first time, or in a higher frequency and volume), a new habit could be formed, and the need for the reminder might disappear. Psychology offers a view of how behavior repetition might translate into durable change. Habits, in the sense as articulated by Wood and Neal (2007), are automatic responses to familiar cues that bypass deliberation once a context–response link is replicated often enough. Lally et al. (2010) observe that automaticity grows quickly at first and then plateaus, reaching 95 percent of its long-run level after two months, on average. However, there is a substantial heterogeneity across individuals and behaviors, and in most cases, habits are difficult to form.

A series of reminders delivered early in the morning of the pickup day, the intervention we study in this paper, addresses the negative behavioral elements described above and is a potential contributor to habit formation. However, the length of the exposure is likely to play a crucial role. A burst of reminders may trigger meaningful gains in the first weeks, yet might not generate enough exposure for a new habit to be established. Also, the reminder's efficacy is unlikely to be uniform over time. Early in an intervention, households might suffer from behavioral inertia or sit below a salience threshold required for action so

[5] Allcott and Rogers (2014) compare persistence effects of sending reminders in the form of "home energy reports" to a group that continues to receive reminders during that same period as well, but without having a comparison group that was first interrupted and then restarted (the restarted treatment).

[6] In this sense, our paper also adds to the literature that uses behavioral experiments to generate evidence on the role of attention and habit formation in behavior change, such as Charness and Gneezy (2009); Gravert and Olsson Collentine (2021); Tiefenbeck et al. (2018, 2019) or Byrne et al. (2024), as we discuss more in our conceptual framework in Section 2.1.

[7] By doing so, we also add to the growing literature on "green nudges" (Carlsson et al., 2021; Grelle et al., 2024; He et al., 2023).

[8] Other behavioral interventions in developing countries include, for example, the commitment-savings product evaluated by Ashraf et al. (2006) in the Philippines; the earmarked savings boxes studied by Dupas and Robinson (2013) in Kenya; small in-kind incentives and reminders that boosted childhood immunization in rural India (Banerjee et al., 2010); SMS reminders that raised savings in Bolivia, Peru and the Philippines (Karlan et al., 2016); social-norm feedback that reduced household electricity use in Mexico (Allcott & Rogers, 2014); and, more recently, feedback on (dynamic) social norms that increased sign-ups to a recycling program in Peru (Fuhrmann-Riebel et al., 2024).

that additional prompts may show increasing returns toward the propensity of the desired behavior. Over time, however, a countervailing force may emerge. Repeated messaging might create "habituation," a well-documented decrement in responsiveness to unchanging stimuli (Rankin et al., 2009); people could become desensitized to the reminders as they receive more of them.

It is essential, however, to recognize that reminders might sustain recycling even without forging true habits. By making recycling top-of-mind exactly when the weekly deadline looms, messages can keep attention consistently above the threshold needed for action; Byrne et al. (2024) shows that such an "attention stock" accumulates gradually, dissipates slowly, and can be re-activated after a pause. This perspective highlights several potential choke points–perception of the cue, valuation of the trade-off, execution of the task–at which interventions may fail: messages may arrive too early and invite procrastination, repeat so often that they fade into background noise, or leave the perceived physical effort of filling the recycling bag untouched.

Our experimental design takes these insights to the data by asking three sequential questions. First, do reminders delivered early on the day of collection increase the probability that a household sets out its recycling at all? A null answer would indicate that attention is not the binding constraint, or that present bias overwhelms any gains from salience. Second, conditional on some positive effect, does the cumulative duration of exposure matter? Comparing a three-week to a nine-week sequence allows us to test whether continued salience translates into higher compliance or slower decay. Third, does inserting a pause midway through a nine-week reminder schedule raise effectiveness relative to sending the messages continuously? Evidence that this paused arm outperforms the continuous long exposure would validate the role of desensitization and the possibility of attention-stock reactivation.

Important recent papers study interventions in other domains that share our theoretical underpinnings, underscoring that the temporal profile of a reminding cue shapes both its immediate impact and its persistence. Byrne et al. (2024) shows in residential water use that longer salience campaigns flatten the decay of conservation once nudges stop, consistent with an attention stock that accumulates gradually and dissipates slowly. Gravert and Olsson Collentine (2021) find that doubling a free-trial period for public transport, rather than issuing a one-off social-norm message, leads to lasting increases in ridership, signaling that sheer exposure length can seed new routines when intrinsic incentives are weak. Complementing these duration effects, Tiefenbeck et al. (2018, 2019) demonstrate that real-time shower feedback achieves sizable and immediate resource savings by overcoming salience bias without attenuation of the treatment. Together, these studies highlight multiple pathways through which repeated or well-timed interventions achieve persistence, and they help motivate our experiment's focus on disentangling the roles of cumulative exposure, temporal spacing, and desensitization.

2.2. Treatments

To address our research questions, we randomly assign individuals to one of four treatments in a between-subjects design. The four groups consist of one control group and three treatment groups (see Table 1). None of the groups receives any reminders during the baseline period. From period 2 onward, we vary the time pattern of reminders, creating four groups as follows:

- Control (T0): no reminders
- Continuous treatment (T1): reminders over all nine weeks; long exposure
- Interrupted treatment (T2): reminders only for the first three weeks; short exposure
- Restarted treatment (T3): reminders for the first three weeks and for the last three weeks with a three-week pause in between; paused exposure

Table 1
Study design overview.

	Baseline			Intervention								
	Period 1			Period 2			Period 3			Period 4		
	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12
Control (T0)												
Continuous (T1)				x	x	x	x	x	x	x	x	x
Interrupted (T2)				x	x	x						
Restarted (T3)				x	x	x				x	x	x

Notes: The table provides an overview of the study design, including three weeks of baseline period followed by nine weeks of intervention period. "x" indicates the weeks in which households received a reminder in the different treatment groups. The 12 weeks of the study ran from mid-June to mid-September 2021.

All treatment groups receive the same treatment during the first three weeks of the intervention period, period 2. In the subsequent three weeks, period 3, only the continuous treatment group (T1) continues to receive reminders, while both T2 and T3 are interrupted and thus not distinguishable from each other up to this point. Only during the last three weeks, period 4, the restarted treatment group (T3) receives reminders again while the interrupted treatment group (T2) remains discontinued, so that all three treatment groups differ.

The treatment variation based on blocks of three weeks was chosen, given the irregular nature of households' recycling behavior. From pre-COVID data (until February 2020) from the municipality, we know that only a few households recycle regularly every week (or even more than once a week, if the bags are collected twice). Most households recycle irregularly, and rather every second or third week, on average. Some households that are registered do not recycle at all. We do not have data on households' recycling behavior afterwards, as the municipality had to stop all measurements due to the pandemic. However, during a one-week pilot that we conducted at the beginning of May 2021, only 13% of all enrolled households recycled in that week, confirming the irregularity in recycling behavior from pre-COVID times.[9]

The treatment message of the mobile text reminder contained a friendly greeting from the municipality and a reminder for people to put their recycling bags outside on that day. We also provided people with a number they could call in case of doubts and referred them to the social media channels of the municipality. The formulation of the reminder message was chosen carefully based on joint discussions with the municipality. All households that received a reminder received the same treatment message.[10] We present the message here in English, while the original message was formulated in Spanish:

IT'S RECYCLING DAY: The Municipality of Miraflores reminds you that the recycling truck will pass by your house today. Don't forget to take out your recycling material! Please find more information on the municipality's social media or by contacting us at [number].

2.3. Treatment comparisons and hypotheses

In this section, we explain how we will compare recycling behavior across the treatments, and present the hypotheses that we pre-registered.[11] We start with a comparison of recycling behavior in period

[9] This percentage is based on the number of addresses that recycled as a fraction of all enrolled addresses.

[10] Research has shown that varying the content of the reminder messages tends to have little or no effect (Altmann & Traxler, 2014; Apesteguia et al., 2013; Larochelle et al., 2019; Wallander et al., 2017). Based on a review of studies that vary the content of reminders, Gravert (2022) therefore concludes that using pure reminder messages directed at the action of interest is often the most effective, and also the most cost-effective approach, especially from a policy perspective.

[11] Note that the formulation of the pre-registered hypotheses was different, yet the content has not changed. The pre-analysis plan can be accessed at the AEA RCT Registry (AEARCTR-0007780).

2. In this period, all treated households receive a weekly reminder. We expect that this will address people's limited attention to the recycling program, while most will not have formed any recycling habits yet. Therefore, we expect the reminders in this period to increase recycling activity, as outlined by Hypothesis 1.

Hypothesis 1 (Period 2). Recycling in the treatment groups (T1, T2 and T3) is higher than in the control group (T0).

In period 3, two treatment groups (T2 and T3) stop receiving reminders. We expect that some households might develop a habit relatively fast and keep recycling in period 3. As a result, we expect that recycling will remain higher in each of the treatment groups, relative to the control group. At the same time, we expect recycling to be smaller in the interrupted treatment groups T2 and T3 relative to the continuous treatment group T1. This might be driven by households that would benefit from a longer series of reminders to start recycling, and households that formed some recycling habits in the first period, whose habits weaken without further reminders. This is summarized in Hypothesis 2.

Hypothesis 2 (Period 3). i) Recycling in the continuous treatment (T1) and the interrupted treatments (T2 and T3) is higher than in the control group (T0); ii) Recycling is higher in the continuous treatment group (T1) than in the interrupted treatments (T2 and T3).

In period 4, the treatment groups are separated into a continuous, interrupted and restarted treatment group. We expect that recycling remains higher in all three treatment groups T1, T2 and T3 (thanks to recycling habits in period 2, supported by the reinforcement of another series of reminders in period 3 and period 4) relative to the control group. As recycling habits might weaken over time, recycling will be higher in the continuous treatment (T1) than in the interrupted treatment (T2). In addition, under the conjecture of desensitization/habituation playing a role, we expect recycling to be higher in treatment T3 than in treatment T1. The restarted reminders in treatment T3 will capture people's attention better than the continuous reminders in T1, since a pause in the reminders will make their resumption stand out, so that the reminders will be more effective when restarted. We summarize these predicted differences in Hypothesis 3.

Hypothesis 3 (Period 4). i) Recycling in each treatment group (T1, T2 and T3) is higher than in the control group (T0); ii) Recycling is higher in the continuous group (T1) than in the interrupted group (T2); iii) Recycling is higher in the restarted treatment group (T3) than in the continuous treatment group (T1).

2.4. Implementation

2.4.1. Setting

For this study, we worked together with a municipal recycling program in the capital city of Peru, Lima. Managing solid waste in Lima is a huge challenge. Around 10 million people live in the country's largest city. It is estimated that around 8468 tons of garbage are generated daily in Lima, of which only 4% is recycled (WWF, 2018). Municipalities are responsible for coordinating recycling activities at the household level in Peru.[12] Some municipalities in Lima have established their own recycling programs, in which households can participate voluntarily and free of charge. This is also the case for the municipality of Miraflores, an upper-middle-to high-income neighborhood in Lima, with which we collaborated for this study. Households participating in the recycling program need to separate their recyclable materials at home, collect them in a separate bag, and deposit this bag on the street on a specific day of the week, so that it can be collected by a recycling truck.

The district of Miraflores consists of 14 zones, based on which the recycling program is organized. Recycling bags are collected on weekdays in the mornings and afternoons, and on Saturdays in the mornings. There are always two recycling trucks operating in two different zones at the same time. In some zones (six out of 14), the bags are collected once a week; in the other zones (eight out of 14), they are collected twice a week, always on the same day(s) at approximately the same time. The municipality informed residents about the collection days across different zones when households registered for the program. Thus, paying attention to this specific day and time to place the recycling bags on the street is crucial for people to participate successfully.

Sample

There are 7183 households officially registered in the recycling program (end of March 2021).[13] The district of Miraflores contains single-family houses as well as apartment buildings and other multiple dwelling units. It is therefore possible that two households are registered with the same address in the program if they live in the same building.

In our study, we only include households that are registered with a unique address. We do so, as we want to link household recycling behavior to specific individual households. If there is more than one household registered in an apartment building, it is impossible to identify to which household the collected recycling bags belong, as recycling bags are usually deposited in a shared space and then taken outside on the street by the caretaker or doorman of the building. Of the 7183 households registered in the recycling program, 3480 had a unique address.

When households register for the recycling program, they are asked to provide a phone number. Households can decide whether to register with a landline or a cell phone number. Since sending mobile text reminders requires access to mobile phone numbers, we focus on those households for which mobile phone numbers are available. This is the case for 1392 households out of the 3480 households that are registered with a unique address. This is the sample that we will use in our study.[14]

Data collection

To keep track of households' recycling behavior, we accompanied the recycling trucks that are responsible for collecting the recycling bags on their daily routes over a total time period of 12 weeks, from mid-June to mid-September 2021. Enumerators followed the recycling trucks by bike.[15] The recording of households' recycling behavior was done through audio recordings via headsets.[16] Enumerators were instructed to record the following details: street name, house number, house type (single family house or apartment building/other multiple dwelling unit), and number of bags. The audio recordings were transcribed to an Excel sheet afterwards. Regular quality checks were applied to both

[12] For a more detailed mapping of different actors involved in the Peruvian recycling sector, see Borasino and Fuhrmann-Riebel (2022).

[13] Studied households were not recruited specifically for our study. They are the actual population of enrolled households in the program as of the beginning of our study. Regarding the enrollment process prior to 2021, the Municipality of Miraflores brought households into the "Basura que no es Basura" program through door-to-door promoters, sign-up booths at public fairs (known as Reciclatones), and other channels such as social media. Residents could also register by phone or by email to the municipality. When COVID-19 hit, face-to-face visits paused and the city pivoted to contact-free enrollment: it launched a WhatsApp line and an online form, and emphasized social-media pushes. Our study uses the full roster of households that had completed one of these municipal enrollment steps before the baseline date.

[14] As per pre-analysis plan, our design yields a minimum detectable effect (MDE, 80% power, $\alpha = 0.05$) of roughly 7 percentage points (henceforth, p.p.) when the baseline recycling rate is 5%, 8 points when it is 10%.

[15] All enumerators were provided a cyclist's insurance for the period of data collection and were experienced in riding a bicycle. They were further instructed to wear a helmet as well as face masks at all times (which were provided to them).

[16] Our initial plan was to have the enumerators sit in the recycling trucks together with the driver and record the data directly in their notebooks. Yet, due to the COVID-19 pandemic, this was not possible at the time given the risk of contagion, so that we had to revert to this alternative way of data collection.

data collection in the field and transcriptions of the audio recordings. Enumerators were not aware of the treatment assignment, i.e., they did not know which households had received reminders and which had not.

The reminders were sent to households in the early mornings of the collection days via mobile text through a Peruvian provider. This was adjusted based on the different zones and respective collection days (so that, for example, households from a zone where the bags are collected on Wednesdays received the reminders on Wednesdays in the early morning). We sent reminders to all households only once a week, regardless of whether the bags were collected once or twice a week in the respective zones.[17] In zones with two collection days per week, the reminders were sent in the morning of the first collection day of the week.

3. Results

3.1. Descriptives

Table A.1 in Appendix A provides a detailed overview of household characteristics by treatment group, including the zone, sign-up year, and house type. Zones range from 1 to 14, sign-up years range from 2015 to 2021. We distinguish the different house types between single-family houses, apartment buildings or other multiple dwelling units, and unknown house types. The table further shows whether households recycled during the baseline period. Balance tests confirm that these household characteristics are balanced across treatment groups.[18]

Since sign-up years go back to 2015, we decided to focus our analysis primarily on households with a recent sign-up year (2018 or later), as there is a risk that some of the respondents who registered a long time ago might not respond to our treatments, mainly for two reasons. First, people might have moved from the district and are thus unable to participate in the recycling program. Second, people might have changed their phone numbers and be unable to receive our mobile text reminders. Ideally, we would have been able to remove those respondents from our database. However, as the municipality does not update their database with contact details on a regular basis, it was not possible for us to identify these respondents. A second-best approach is therefore to focus on households with a more recent sign-up year, as the risk of reaching an invalid household is smaller among recently registered households. This assumption is supported by Table A.2 in Appendix A, which confirms that the percentage of households that recycled at least once during the whole intervention period and during the baseline period is higher among those that signed up more recently.

In our main analysis, we will focus on households that signed up in 2018 or later, which reduces our total sample size from 1392 to 898 households. We provide a detailed overview of household characteristics by treatment group as well as balance tests for those households with a recent sign-up year separately in Appendix A (Table A.3). We can confirm that household characteristics are also balanced across treatment groups in the sub-sample of recently registered households. To test the robustness of our results, we will also report all results using the full sample (see Appendix E).

It should be noted that there was a week of public holiday after the first three weeks of the intervention period, i.e., between weeks 6 and 7 of our study period, during which no recycling service was in place and thus no reminders were sent and no data could be collected. Therefore, the entire implementation period was extended to 13 weeks, while the analysis focuses on the 12 weeks of our study period.

The same was the case for the Monday of week 11, where again no recycling service was in place due to a public holiday, and thus no reminders were sent and no data could be collected for that day. In this case, no additional week was added, so that week 11 contains one day less in our analysis.

3.2. Treatment effects by period

The main outcome variable of our analysis is a binary variable equal to one if a household recycled at least once per period of three weeks, zero otherwise. We focus on recycling behavior per period rather than per week because of the irregular nature of households' recycling activity, as explained in Section 2.2. Histograms on households' recycling frequency in Appendix B confirm that hardly any household recycled every week during the baseline period.

We start with a visual presentation of the recycling activity over the individual periods 1 to 4. Fig. 1 traces period-by-period the share of households that recycled at least once. In Fig. 1(a), we pool the three treatment groups together to compare them with the control group. While the recycling rates are similar during the baseline period (at around 19%), we observe a clear treatment–control difference once the intervention begins: in periods 2 and 3, the percentage of treated households that recycle is about 5 p.p. higher than in the control group, with largely non-overlapping confidence intervals (CIs). Although the recycling activity in all groups dips in period 4, the treatment CI band remains above the control recycling rate.

In Fig. 1(b), we pool groups that have been treated identically in the respective periods (that is, groups that share the same reminder exposure profile). Because all treatment groups are identical in period 2, in this period, the red CI is the same as in the previous panel. From period 3 onward, the designs diverge. Households that keep receiving weekly reminders (T1, orange) raise their recycling rate to 24% (compared to roughly 18% in the control group), whereas those whose reminders are paused (purple, T2&T3) still show a higher recycling rate at 21%, yet now exhibit a CI that mostly overlaps with the control group. In period 4, when the restarted treatment group (grey, T3) resumes the messages, all treatments are different. While the recycling activity remains higher in all treatment groups, their CIs now mostly overlap with the control group, although less so for the interrupted treatment.

In Fig. 1(c), we show the three treatment arms separately for each period. We see that in period 2, all treatment groups exhibit a higher recycling rate than the control group (between 21–22% versus ∼16%). In period 3, recycling rates remain higher in all treatment groups compared to the control group, although the CI of the restarted group mostly overlaps with that of the control group. In period 4, as shown in Fig. 1(b), all treated groups still have recycling rates above the control group, but all CIs largely overlap. We conduct the formal statistical tests of treatment effects within the regression analysis.

The analyses presented below in Table 2 are the regressions associated with Fig. 1(b). We focus on these regressions in our main analysis as they most accurately mirror the treatment exposure of our intervention. As an additional robustness check, we present panel regressions associated with all sub-figures of Fig. 1 in Table F.2 of the Appendix F. For our main analysis, we estimate a regression of the following form:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + u_i \tag{1}$$

with Y_i being a binary variable equal to one if a household recycled at least once in the respective period, zero otherwise. T_i is a set of indicator variables for the treatment groups, using the control group as reference. X_i includes controls for households' baseline recycling activity (i.e., whether they recycled at least once in period 1), the house type (whether it is a single-family house as opposed to a multiple dwelling unit or an unknown house type), and indicator variables for the different enumerators. We estimate regressions with OLS.[19] To adjust for

[17] We did so to maintain treatment uniformity within and across treatment groups.

[18] The p-values of these tests are reported in Table A.1 of Appendix A.

[19] A robustness check using logistic regressions is presented in Appendix D.1. All results are robust to using logit specifications.

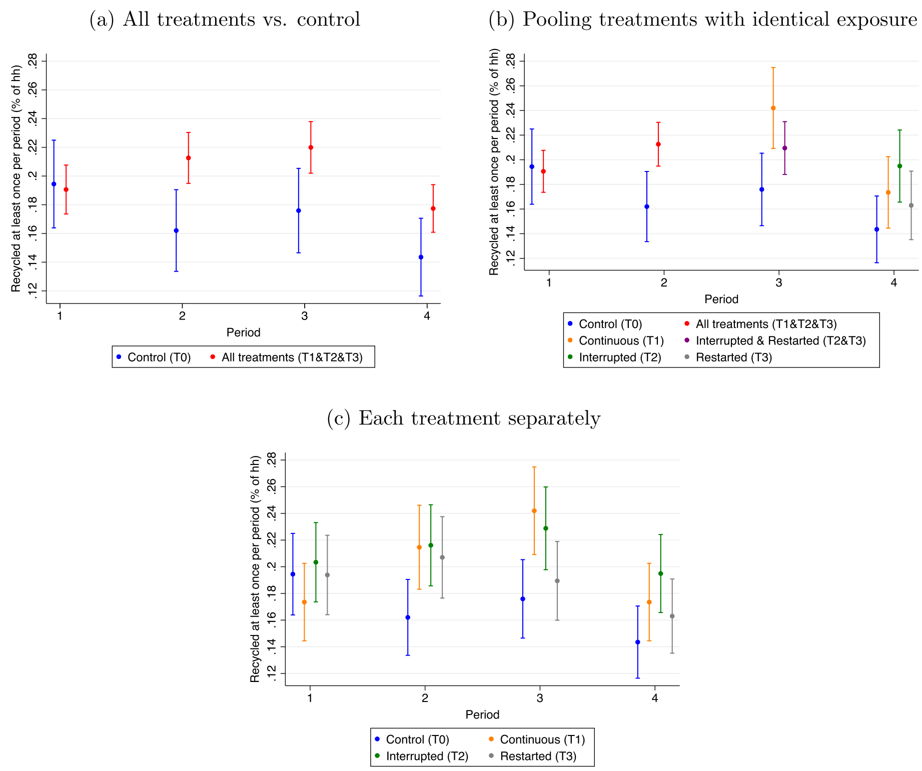

Fig. 1. Proportion of households that recycled at least once per period, by treatment. *Notes:* The figure shows the proportion of households that recycled at least once per period in the control group and in the different treatment groups over the whole study period, focusing on households with a recent sign-up year (≥ 2018). (a) compares all treatment groups pooled together with the control group; (b) pools treatment groups that have identical exposure as of each period in the horizontal axis; and (c) compares the control group with each treatment group separately. Vertical lines and whiskers denote the 95% CI. N(T0) 216, N(T1&T2&T3) = 682, N(T1) = 219, N(T2&T3) = 463, N(T2) = 236, N(T3) = 227.

Table 2
Treatment effects on whether households recycled at least once, by period.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.047** (0.019)	0.054** (0.026)		
Continuous (T1)			0.083*** (0.030)	0.042 (0.028)
Interrupted & Restarted (T2&T3)			0.034 (0.025)	
Interrupted (T2)				0.049* (0.028)
Restarted (T3)				0.019 (0.027)
Predicted Probability Control (T0)	0.158*** (0.017)	0.160*** (0.023)	0.172*** (0.020)	0.141*** (0.018)
Adjusted R^2	0.355	0.341	0.393	0.332
Observations	2694	898	898	898

Notes: The table reports OLS regressions with dependent variable equal to one if a household recycled at least once during the respective period. All regressions include a control for households' binary baseline recycling activity and the house type, and use fixed effects for enumerators. Only households included that signed up in 2018 or later. Robust standard errors are in parentheses; standard errors are clustered at the household level in column (1). ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Exact p-values as well as q-values resulting from multiple hypotheses corrections are presented in the text. Logit regressions are reported in Table D.1. Panel regression analog is reported in Table F.2. A panel logit regression analog is reported in Table F.4.

multiple hypothesis testing, we calculate sharpened q-values for the seven estimates presented in Table 2 (Anderson, 2008; Benjamini et al., 2006).

We first look at the effect of all treatment groups pooled together over the entire intervention period. The regression results reported in column (1) of Table 2 confirm that the treatment-control difference in recycling rates is 4.7 p.p. and statistically significant (p-value = 0.015; q-value = 0.048).

We next look at each of the intervention periods separately according to our treatment exposure. Column (2) of Table 2 shows that in period 2, the difference in recycling levels between the treatment groups and the control group is 5.4 p.p. and statistically significant (p-value = 0.039; q-value = 0.070), supporting our Hypothesis 1.

We now turn to period 3, where the treatment groups are split between households that continued receiving weekly reminders (continuous treatment, T1) and those that did not (interrupted treatment groups, T2 and T3). As explained above, we continue to pool treatments T2 and T3 together for this analysis since both did not receive any reminders during period 3 and are thus technically not distinguishable from each other up to this point. Column (3) in Table 2 confirms that in period 3, the percentage of households that recycled is 8.3 p.p. higher in the continuous group than in the control group (p-value = 0.006; q-value = 0.041). While the coefficient of T2&T3 is positive in column (3) as well, it is not statistically significant. Moreover, when comparing treatment coefficients, we find weak evidence that the coefficient of T1 is significantly larger than the coefficient of T2&T3 (p-value of the two-sided Wald test = 0.065). This indicates that when only looking at period 3, the recycling activity of households in the continuous group is significantly higher than that of households in the interrupted groups. We can thus partly confirm our Hypothesis 2, although the positive difference in recycling activity that we observe in the interrupted treatment groups compared to the control group in period 3 is not statistically significant.

Table 3
Treatment effects on whether households recycled at least once, by period–baseline recycling yes vs. no.

	Baseline recycling = yes				Baseline recycling = no			
	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)	Period 2-4 (5)	Period 2 (6)	Period 3 (7)	Period 4 (8)
All treatments (T1&T2&T3)	0.107	0.220**			0.028*	0.010		
	(0.069)	(0.087)			(0.016)	(0.024)		
Continuous (T1)			0.191**	0.000			0.058**	0.050**
			(0.095)	(0.109)			(0.029)	(0.024)
Interrupted & Restarted (T2&T3)			0.051				0.024	
			(0.089)				(0.023)	
Interrupted (T2)				0.036				0.048**
				(0.104)				(0.024)
Restarted (T3)				−0.022				0.023
				(0.108)				(0.022)
Predicted Probability Control (T0)	0.590***	0.514***	0.652***	0.606***	0.058***	0.080***	0.061***	0.034**
	(0.061)	(0.077)	(0.074)	(0.075)	(0.013)	(0.021)	(0.017)	(0.014)
Adjusted R^2	0.018	0.036	0.008	−0.020	0.028	0.010	0.055	0.022
Observations	516	172	172	172	2178	726	726	726

Notes: The table reports OLS regressions with dependent variable equal to one if a household recycled at least once during the respective period. Only households included that signed up in 2018 or later. Columns (1)–(4) only include households that recycled during the baseline period; columns (5)–(8) only include households that did not recycle during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in columns (1) and (5). Using two-sided Wald tests to compare treatment coefficients gives the following p-values: column (3) p = 0.073 for T1 = T2&T3; column (4) p = 0.740 for T1 = T2, p = 0.839 for T1 = T3, p = 0.580 for T3 = T2; column (7) p = 0.211 for T1 = T2&T3; column (8) p = 0.938 for T1 = T2, p = 0.310 for T1 = T3, p = 0.338 for T3 = T2. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. OLS regression with baseline recycling as a dependent variable is reported in Table D.2. Panel regression analog is reported in Table F.3.

Finally, we turn to period 4, where the treatment groups are separated into the continuous, interrupted, and restarted group. As seen in Fig. 1, the recycling activity in period 4 remains higher in the treatment groups than in the control group, yet drops in all groups compared to previous periods, and even falls below levels in the baseline period in the case of the control group.[20] The regressions reported in column (4) of Table 2 confirm that the treatment coefficients are still positive, but not statistically significant. The only exception is the interrupted treatment group, which is statistically significant at the 10% level, yet only before adjusting for multiple hypotheses testing (p-value = 0.081; q-value = 0.101). Using a two-sided test that compares the treatment coefficients, none of the differences are statistically significant (p-values are: 0.827 for T1 = T2, 0.438 for T1 = T3, and 0.313 for T3 = T2). In sum, we do not find evidence in support of Hypothesis 3.

In Appendix F, Table F.2, we report the results of the panel regression (see specification Eq. (2)). Results are largely consistent with those in Table 2. In these fully disaggregated exercises, T1 has a significant positive effect in periods 2 and 3 (0.083*** in OLS vs. 0.087*** in the panel regression), and none of the other treatments have any significant impact on recycling behavior in any period. In particular, the weakly significant effect of T2 in period 4 does not show up in this panel analysis (0.049* in OLS vs. 0.042 in the panel regression). The combined effect of all treatments compared to the control group remains significant at the 10% level (periods 2–4: 0.047** in OLS vs. 0.047* in the panel regression, period 2: 0.054** in OLS vs. 0.054* in the panel regression).

3.3. Heterogeneous effects and mechanisms

3.3.1. Heterogeneity by recycling at baseline

While the previous section presented our main analysis that informed our pre-registered hypotheses, this section provides additional insights regarding heterogeneous effects and the mechanisms behind our effects. The histogram presented in Fig. B.1 in Appendix B shows that around 20% of the households in our sample did recycle at least once in the baseline period, with only around 3% recycling every week. Given this heterogeneity, it is interesting to test whether the reminder effects vary between households that already recycled at baseline and those that did not. It is plausible that both types of households would benefit from reminders, but in a different way. For example, households that did not yet recycle at baseline might need more reminders to start recycling than households that already recycled recently.

Table 3 separates the main regressions by whether households recycled at least once during the baseline period (period 1). We observe that the predicted probabilities are substantially higher in the left panel compared to the right panel. This indicates that respondents who recycled at baseline are much more likely to recycle in each of the subsequent periods, even without being treated. Looking at the size of the treatment coefficients, we observe that the reminder effects in periods 2 and 3 are stronger among households that already recycled at baseline. Reminder effects among households that did not recycle at baseline are absent in period 2, but do appear in period 3 with the continuous treatment. These findings suggest that households that did not yet recycle at baseline may need a longer series of reminders to start recycling, while the immediate effects that we notice in period 2 seem to be mainly driven by those households that already recycled before. Comparing the effects of treatment T1 and the interrupted treatments (T2 & T3) in period 3, we observe that their difference is substantial and (weakly) significant in the left panel, but not in the right panel.

In period 4, we find significant positive effects of the continuous and the interrupted treatment among households that did not recycle at baseline, but not among households that already recycled before. In fact, the effect for baseline recyclers goes from around 20 p.p. to zero. The lack of significant effects among the respondents who recycled in the baseline period may be due to the lower sample size. In Appendix D.2, Table D.2, we report the single specification version of the same table, where we interact the baseline recycling dummy with the treatment indicators. In Appendix F, Table F.3, we report the panel regression analog of the analysis of this section. In both cases, the results closely track those reported above.

[20] As already indicated in Section 3.1, there was a public holiday on the Monday of week 11 where no recycling service was in place. As a result, one day of data is missing in week 11. Moreover, from discussions with the municipality staff, we know that many people used this public holiday to travel, which might have led to even lower recycling rates during the whole week. Fig. C.1 in Appendix C confirms that the low recycling rates in period 4 are mainly driven by low recycling levels in week 11.

Table 4
Treatment effects on recycling in periods 2 and 3 combined.

Recycled in ...	(1) neither p2 nor p3	(2) only p3	(3) only p2	(4) p2 and p3
Panel A: Pooled sample (N = 898)				
Continuous (T1)	−0.152***	0.059**	0.040	0.053*
	(0.045)	(0.026)	(0.025)	(0.030)
Interrupted & Restarted (T2&T3)	−0.068**	0.011	0.020	0.036
	(0.034)	(0.017)	(0.019)	(0.023)
Predicted Probability Control (T0)	0.859***	0.039***	0.041***	0.062***
	(0.026)	(0.014)	(0.014)	(0.018)
Panel B: Recycled at baseline (N = 172)				
Continuous (T1)	−0.233***	0.034	0.049	0.150
	(0.084)	(0.090)	(0.050)	(0.111)
Interrupted & Restarted (T2&T3)	−0.153*	−0.076	0.110**	0.120
	(0.083)	(0.069)	(0.044)	(0.095)
Predicted Probability Control (T0)	0.308***	0.175***	0.024	0.494***
	(0.073)	(0.061)	(0.025)	(0.079)
Panel C: No recycling at baseline (N = 726)				
Continuous (T1)	−0.076**	0.038**	0.023	0.016
	(0.035)	(0.019)	(0.024)	(0.020)
Interrupted & Restarted (T2&T3)	−0.016	0.016	−0.008	0.008
	(0.026)	(0.012)	(0.018)	(0.016)
Predicted Probability Control (T0)	0.913***	0.015*	0.041***	0.031**
	(0.021)	(0.009)	(0.015)	(0.013)

Notes: Marginal probabilities of a multinomial probit regressions on (1) whether households recycled neither in period 2 nor in period 3, (2) whether households recycled only in period 3, (3) whether households recycled only in period 2, (4) whether households recycled at least once in both period 2 and period 3. All regressions include a control for the house type and use fixed effects for enumerators. The regressions for the pooled sample also include a control for households' binary baseline recycling activity. Only households included that signed up in 2018 or later. Standard errors are in parentheses. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. We compare the effect of T1 and T2&T3 using two-sided Wald tests, which give the following p-values. Panel A: column (1) p = 0.046; column (2) p = 0.053; column (3) p = 0.403; column (4) p = 0.562. Panel B: column (1) p = 0.168; column (2) p = 0.136; column (3) p = 0.279; column (4) p = 0.750. Panel C: column (1) p = 0.060; column (2) p = 0.252; column (3) p = 0.136; column (4) p = 0.669.

3.3.2. Mechanisms

Our conceptual framework argues that reminders might boost recycling through two main psychological mechanisms. The first is *limited attention*: households that intend to recycle often fail to do so because the task never becomes sufficiently salient at the right moment. Recurring reminder messages might help with this. The second is *habit formation*: repeated recycling might build a routine that could persist even when reminders cease. Several subtleties could accompany these forces. The effect of reminders could be immediate for some individuals, while for others, we might need a longer treatment exposure for the nudges to start drawing their attention. To probe the evidence on these possible behavioral mechanisms, we focus on behavior in period 2, where every treated household received weekly reminder messages, and period 3, where only the continuous treatment (T1) received reminders while the interrupted treatments (T2 & T3) did not. As discussed above, a public holiday might have affected behavior and collection in period 4, so we focus on the first two post-baseline periods (2 and 3) in this analysis.[21]

[21] In this analysis, as stated above, we do not consider period 4, because that period appears different from the rest, even for the control group. We acknowledge that this may stem from the holiday-related drop in overall activity and an increase in zero observations, but also the possibility that some other factors we do not consider are at work in that period.

Table 4 reports on the effects from a multinomial model that classifies every household into one of four mutually exclusive states: recycling in neither period (never), only in period 3 (only later), only in period 2 (only early), or in both periods (always). The results of this exercise reveal that the continuous treatment (T1) generates a sizable reduction in the probability of *never* recycling across periods 2 and 3 (−15 p.p., in Panel A, Column 1), mainly driven by baseline recyclers who show a strong reaction to the messages (−23 p.p. in Panel B, Column 1). T1, in particular, impacts the probability of recycling only late (+5.9 p.p. in Panel A, Column 2). Interestingly, this improvement is largely driven by households that were baseline non-recyclers (Panel C, Column 2). In other words, the behavioral response to T1 crosses the salience or attention threshold and becomes action only after an extended exposure of six consecutive weeks. There is also weak evidence (at 10% level) that T1 raises the probability of recycling in *both* periods, hinting that some households react to the treatment faster. The fact that T1 does not impact recycling only in period 2 suggests that backlash or habituation to treatment are not strong forces in our setting.

The interrupted treatments (T2 & T3), which received reminders in period 2 and not in period 3, exhibit a different profile of effects. Table 4 shows that they, too, boost the likelihood of recycling at least once (lowering the chance of never recycling by 6 p.p.), and this effect is concentrated among baseline recyclers who react to the messages contemporaneously (Panel B, Column 3, +11 p.p.). When the nudges stop in period 3, the gain vanishes. The probabilities of recycling only in period 3 or in both periods remain statistically unchanged. In other words, reminder messages in period 2 create salience, eliciting an immediate response from baseline recyclers, yet fail to induce persistent behavior.

In sum, the evidence indicates that a lengthy exposure to reminders first works by overcoming limited attention immediately among baseline recyclers, and slowly among non-recyclers who only start acting after being given enough exposure. Briefer exposure triggers only a short-lived contemporaneous response among households that were already recycling in the baseline period and fails to establish lasting routines. It does not change the behavior of baseline non-recyclers. These patterns associated with the effects of the continuous (T1) and interrupted treatments (T2&T3) are mostly consistent with our findings in Tables 2 and 3, as well as in the panel regression (Tables F.2 and F.3 in Appendix F).

3.4. Other analyses

3.4.1. Treatment effects over the entire intervention period

Next, we look at treatment effects on households' recycling behavior aggregated over the entire intervention period (periods 2-4).[22] This analysis is potentially interesting for policy-makers who want to know how many reminders they need to send out and in what sequence to maximize their effect on recycling behavior.

Fig. 2 shows the distribution of the total recycling frequency over the whole intervention period, by treatment. Specifically, for each period, we count whether a household recycled at least once, and we take the sum of this count across the three intervention periods (periods 2, 3, and 4). This means that this variable has a value of 3 if a household recycled at least once in each of the three periods. Comparing the three treatments with the control group in Fig. 2, we observe a slight shift to the right (i.e., toward higher recycling frequency) in the treatment groups. This suggests that in all treatment groups households recycled in more periods, on average, than in the control group over the whole intervention period. Comparing Fig. 2(b) and (c), we observe that this slight shift seems to hold irrespective of whether respondents already recycled at baseline.

To test whether treatment differences are statistically significant, we use an ordered probit regression, combining the four frequency categories as the dependent variable. On the right-hand side, we use the same specifications as in our main regression analysis. Table 5 presents

[22] Note that this analysis is exploratory and was not pre-registered.

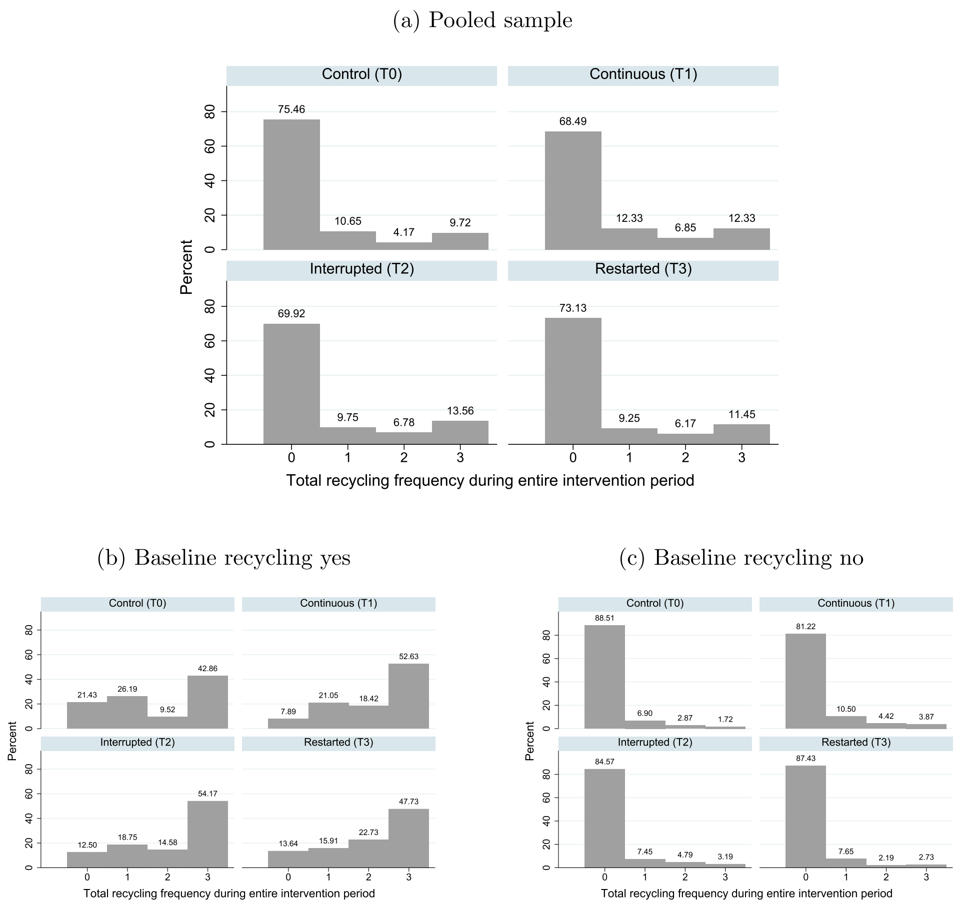

Fig. 2. Total recycling frequency during entire intervention period, by treatment. *Notes:* The figure shows histograms for the total recycling frequency during the entire intervention period (period 2-4). Concretely, it shows in how many of the different periods of the intervention period households recycled at least once (from zero to three periods). Only households included with a recent sign-up year (≥ 2018). (a) for pooled sample: N(T0) = 216, N(T1) = 219, N(T2) = 236, N(T3) = 227; (b) for households that recycled during the baseline period: N(T0) = 42, N(T1) = 38, N(T2) = 48, N(T3) = 44; (c) for households that did not recycle during the baseline period: N(T0) = 174, N(T1) = 181, N(T2) = 188, N(T3) = 183.

the results for the pooled sample and the two sub-samples. In Panel A, we observe that treatments T1 and T2 significantly reduce the likelihood of no recycling and increase the likelihood that respondents recycle at any of the three frequencies, relative to the control group. Using the predicted probabilities in the control group as reference, these treatment effects are sizable. Comparing the treatment effects, we observe that the coefficients of T1 are larger in absolute terms relative to treatments T2 and T3, yet the differences are not statistically significant.

Panels B and C disaggregate the results by baseline recycling. All coefficients go in the expected direction, and suggest sizable effects. The null of the coefficients in Panel B, however, cannot be rejected, likely due to the small sample size. In Panel C, the effects of the continuous treatment remain highly significant, and are now also statistically different from the effects of the restarted treatment T3 at the 10% significance level.

3.4.2. Intensive margin

So far, we focused on the extensive margin of whether households recycled at least once in a certain period. Yet, there might also be an effect of reminders on the intensive margin of households' recycling activity. As a first measure of recycling intensity, we look at how often (i.e., the number of weeks) a household recycled per period. Given that each period consists of three weeks, this variable can take the values zero to three. Table D.3 in Appendix D.3 presents regression results with this measure as dependent variable. The analysis shows that the earlier identified treatment effects are still present, yet more difficult to identify with this new specification, in terms of statistical significance. These findings suggest that the treatments are less effective at increasing recycling frequencies beyond once per period. This can also be observed in the histograms presented in Appendix B, where we can see that very few households recycle more than once per period, even after receiving reminders.

Splitting up this analysis by households that did or did not recycle during the baseline period, we find that among households that did recycle at baseline, none of the coefficients are significant (see Table D.4). While this might be due to the lower sample size of this sub-group of households, the coefficients have the expected positive sign. Among households that did not recycle at baseline (see Table D.5), we find support for the previous treatment effects: reminders increase the frequency of recycling, and this effect is driven by the continuation of reminders in period 3. Again, we observe positive effects among this sub-group in both the continuous and the interrupted treatment in period 4.

As a second measure of recycling intensity, we focus on the number of bags a household recycled per period. Table D.6 in Appendix D.4 presents regressions results using the number of bags recycled as dependent variable. Our analysis shows that none of the treatment effects are statistically significant.

Disaggregating the analysis by baseline recycling, we find again that among households that did recycle at baseline, none of the coefficients are significant (see Table D.7). While this might again be partly due to the lower sample size, it is noteworthy that all coefficients have negative signs. If anything, it suggests that reminders make recycling somewhat

Table 5
Treatment effects on recycling over entire intervention period.

Recycled in …	no period (1)	one period (2)	two periods (3)	three periods (4)
Panel A: Pooled sample (N = 898)				
Continuous (T1)	−0.112***	0.067***	0.030**	0.038**
	(0.043)	(0.018)	(0.012)	(0.015)
Interrupted (T2)	−0.085**	0.036**	0.022**	0.027**
	(0.041)	(0.018)	(0.011)	(0.013)
Restarted (T3)	−0.044	0.019	0.011	0.013
	(0.040)	(0.018)	(0.011)	(0.012)
Predicted Probability Control (T0)	0.812***	0.111***	0.043***	0.033***
	(0.028)	(0.016)	(0.009)	(0.009)
Panel B: Recycled at baseline (N = 172)				
Continuous (T1)	−0.084	−0.056	−0.010	0.150
	(0.058)	(0.039)	(0.011)	(0.100)
Interrupted (T2)	−0.071	−0.044	−0.006	0.121
	(0.057)	(0.035)	(0.008)	(0.095)
Restarted (T3)	−0.051	−0.029	−0.003	0.083
	(0.060)	(0.035)	(0.005)	(0.097)
Predicted Probability Control (T0)	0.187***	0.240***	0.170***	0.403***
	(0.051)	(0.039)	(0.029)	(0.071)
Panel C: No recycling at baseline (N = 726)				
Continuous (T1)	−0.077**	0.037**	0.020**	0.019**
	(0.035)	(0.018)	(0.010)	(0.010)
Interrupted (T2)	−0.049	0.025	0.013	0.012
	(0.034)	(0.017)	(0.009)	(0.008)
Restarted (T3)	−0.013	0.007	0.003	0.003
	(0.031)	(0.017)	(0.008)	(0.007)
Predicted Probability Control (T0)	0.901**	0.062***	0.023***	0.015*
	(0.022)	(0.013)	(0.007)	(0.006)

Notes: Marginal probabilities of ordered probit regressions on whether households recycled in (1) no period, (2) one period, (3) two periods, or (4) three periods out of the three periods of the intervention. All regressions include a control for the house type, and use fixed effects for enumerators. The regressions for the pooled sample also include an indicator variable equal to one, if they recycled at baseline. Only households included that signed up in 2018 or later. Standard errors are in parentheses. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Comparing the marginal probabilities of T1, T2 and T3, respectively, using two-sided Wald tests gives the following p-values:
Panel A: column (1) T1 = T2: p = 0.508, T1 = T3: p = 0.103, T2 = T3: p = 0.321; column (2) T1 = T2: p = 0.509, T1 = T3: p = 0.106, T2 = T3: p = 0.324; column (3) T1 = T2: p = 0.509, T1 = T3: p = 0.108, T2 = T3: p = 0.325; column (4) T1 = T2: p = 0.510, T1 = T3: p = 0.111, T2 = T3: p = 0.325;
Panel B: column (1) T1 = T2: p = 0.781, T1 = T3: p = 0.510, T2 = T3: p = 0.686; column (2) T1 = T2: p = 0.782, T1 = T3: p = 0.512, T2 = T3: p = 0.685; column (3) T1 = T2: p = 0.785, T1 = T3: p = 0.538, T2 = T3: p = 0.690; column (4) T1 = T2: p = 0.781, T1 = T3: p = 0.509, T2 = T3: p = 0.685;
Panel C: column (1) T1 = T2: p = 0.465, T1 = T3: p = 0.076, T2 = T3: p = 0.294; column (2) T1 = T2: p = 0.466, T1 = T3: p = 0.080, T2 = T3: p = 0.297; column (3) T1 = T2: p = 0.468, T1 = T3: p = 0.089, T2 = T3: p = 0.089; column (4) T1 = T2: p = 0.470, T1 = T3: p = 0.094, T2 = T3: p = 0.304.

more efficient in terms of fewer bags used. Among households that did not recycle at baseline (see Table D.8), we find again that reminders increase recycling as measured by the number of bags recycled, and that this effect is driven by the continuation of reminders in period 3. Also for this variable, we observe significant positive effects among this sub-group in period 4 in both the continuous and the interrupted treatment.

4. Conclusion

In this paper, we test whether simple reminder messages can increase the recycling activity of households registered in a recycling program in urban Peru. Even though households signed up for the recycling program voluntarily, many do not recycle regularly. This suggests that there may be a gap between their intention to recycle and actual recycling behavior, as well as a lack of routine and habit formation in recycling practices. The reminders aim to bring the recycling program to people's attention on the day when the recycling bags are collected. This could address the problem of limited attention, which often hampers people from following through with good intentions and forming habits in desirable behaviors. We study how varying the length of exposure to reminders and their continuity impacts recycling behavior over time and whether it supports habit formation in recycling practices.

To identify the causal effect of reminders, we collect data on households' weekly recycling behavior over 12 weeks and compare it across experimental treatments that vary the exposure length and continuity of the reminders. We find that reminders are generally effective at increasing recycling. In the first intervention period where all treatment groups receive weekly reminders (period 2), recycling increases relative to the control group by 5.4 percentage points. This effect is mainly driven by households that already recycled at baseline, where the series of reminders increases the likelihood of recycling by 22 percentage points.

In period 3, the continuous treatment increases recycling rates even further to 8.3 percentage points above levels in the control group. It also performs (weakly) significantly better than the interrupted treatments. The continuation of reminders maintains recycling among households that recycled at baseline, while recycling drops in this subgroup if reminders are not continued. In this period, we also find that households that did not recycle at baseline start recycling if reminders are continued, which suggests that they need a longer series of reminders to change their recycling behavior.

When looking at households' recycling activity over the whole intervention period, treatment differences become less evident. The continuous and interrupted treatments show a positive effect relative to the control group, and the continuous treatment performs better than the restarted treatment. The difference between the continuous and interrupted treatments, however, becomes difficult to distinguish statistically. It might be due to a lack of statistical power in period 4, where we split up the treatments, and households' recycling behavior was reduced by a public holiday and the related traveling of many people. This is supported by the observation that where we only include periods 2 and 3 (Table 4), we do find strong differences between the continuous and interrupted treatments.

Our results provide evidence that both limited attention and habit formation matter for recycling behavior. Limited attention is supported by the fact that several behavioral responses are contemporaneous with the start of the messaging. It is further supported by the finding that the continuation of reminders increases the likelihood that respondents start to recycle in period 3, relative to the interrupted treatments. The finding that the interrupted treatments increase the likelihood that respondents only recycle in period 2 (and not in period 3, especially among recyclers) suggests that although their attention is activated, the short length of exposure did not establish a firm recycling habit by the time we stop sending reminders. At the same time, we find that respondents who recycled at baseline have a 50% probability to recycle in two consecutive periods, irrespective of whether we send reminders, which suggests that these respondents have already established firm recycling habits.

Our results have the following implications for policy. First, we show that a simple, low-cost tool in the form of sending weekly SMS reminders can effectively increase households' recycling activity. Given the low implementation costs and the widespread use of mobile phones, the intervention can be scaled up easily.[23] Second, people who did not yet recycle at baseline might need a longer series of reminders to start

[23] Back of envelope calculations of the net benefit are available upon request. We find that for a significant range of SMS costs and recycled material prices (as lower-bound proxies of the social benefit of a unit of recycled material), the

recycling. Third, as for people who already recycled at baseline, some of them have already established recycling habits and do not need any reminders, while others benefit from reminders and increase the frequency with which they recycle. Finally, we do not find evidence for a backlash in recycling or habituation as a result of our intervention.

We end with two final reflections. First, the estimated effects of our intervention most likely present a lower bound. As we were unable to check whether the phone numbers that we received from the municipality were still active (or still belonged to the initial owner), it is likely that a proportion of participants never received our reminder messages.[24] Second, we cannot assert with confidence that our intervention period was long enough for the development of permanent recycling habits. It is possible that a longer study period might have led to the development of stronger recycling habits. It may also be that, if recycling habits were formed, they would unravel over time, and a new set of reminders might be needed again. Moreover, habituation to the treatment may eventually occur, which would support the stopping and restarting of reminders. Research investigating recycling behavior and the importance of reminders in combination with habit formation over a longer time frame would be a promising line of future research.

CRediT authorship contribution statement

Hanna Fuhrmann-Riebel: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Ben D'Exelle:** Writing – review & editing, Visualization, Validation, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Kristian López Vargas :** Writing – review & editing, Visualization, Resources, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Sebastian Tonke:** Writing – review & editing, Visualization, Validation, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Arjan Verschoor:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Erwin Bulte, Ximeng Fang, Bereket Kebede, as well as participants of the Behavioral Economics Exchange Group of IDOS and the University of Bonn, the IDOS PhD Colloquium 2022, a CSERGE (UEA) Workshop 2023, a Busara Center for Behavioral Economics Learning Event on Tackling Waste 2023, and the BEHAVE Conference 2023 for valuable comments and suggestions. We are grateful to the Municipality of Miraflores in Lima, Peru, for their cooperation on this project, to Ricardo Arce, to Micaela Venancio and her team from B-Green for outstanding operational assistance, and to all enumerators who were involved in the data collection. We further thank Paula Armas, Valerie Dube, and Martín Sánchez for their great research assistance. The study was approved by the International Development Research Ethics Committee of the University of East Anglia on March 3rd, 2020. We preregistered our experiment at the AEA RCT Registry (AEARCTR-0007780: https://www.socialscienceregistry.org/trials/7780). This research was funded by the German Federal Ministry of Education and Research (BMBF), the German Institute of Development and Sustainability (IDOS), the University of East Anglia (UEA GCRF Rapid Response Fund), the University of California, Santa Cruz, and the Max Planck Institute for Behavioral Economics.

net effect is positive. We do not report them in the paper, as these calculations depend heavily on numerous assumptions and contextual variables.

[24] While we cannot provide evidence on the proportion of text messages that reached their intended recipients in our study, it would be in line with the findings of other research that made use of SMS reminders, such as Chong et al. (2015) or Aker et al. (2020).

Appendix A. Balance tests and descriptive overview

See Tables A.1–A.3

Table A.1
Household characteristics by group–full sample.

		Control (T0)	Continuous (T1)	Interrupted (T2)	Restarted (T3)	Total
Zone	1	29	31	24	21	105
	2	16	19	25	23	83
	3*	21	31	29	19	100
	4*	23	20	25	30	98
	5*	15	12	20	12	59
	6*	14	22	19	16	71
	7	9	9	18	8	44
	8*	13	13	12	10	48
	9	28	25	15	36	104
	10	26	21	25	23	95
	11*	77	58	57	71	263
	12	13	17	15	16	61
	13*	18	25	12	26	81
	14*	46	45	52	37	180
Sign-up year	2015	2	2	4	0	8
	2016	13	12	8	12	45
	2017	117	115	100	109	441
	2018	16	12	14	15	57
	2019	90	82	85	86	343
	2020	78	90	105	94	367
	2021	32	35	32	32	131
House type	Single family house	157	152	140	153	602
	Multiple dwelling unit	117	124	135	122	498
	n/a	74	72	73	73	292
Baseline recycling	yes	63	59	65	61	248
	no	285	289	283	287	1144
	N	348	348	348	348	1392

Notes: The table shows the number of households per group for different household characteristics. A * behind the zone number means that bags are collected on two days per week in the respective zone, without a star means that they are collected once a week. Balance tests using chi2 tests show that household characteristics are balanced across groups: p = 0.213 for zone, p = 0.839 for sign-up year, p = 0.598 for house type, and p = 0.942 for baseline recycling.

Table A.2
Households that recycled at least once: old vs. recent sign-up years.

	Old sign-up years	Recent sign-up years	Full sample
Recycled never (whole study period)	73.89%	69.04%	70.76%
Recycled at least once (whole study period)	26.11%	30.96%	29.24%
Recycled never (baseline period)	84.62%	80.85%	82.18%
Recycled at least once (baseline period)	15.38%	19.15%	17.82%
Recycled never (intervention period)	77.53%	71.71%	73.78%
Recycled at least once (intervention period)	22.47%	28.29%	26.22%
N	494	898	1392

Notes: The table shows the percentages of households that recycled at least once during the whole study period, during the baseline period, and during the intervention period, separated for old and recent sign-up years. Old sign-up years include years 2015–2017 (N= 494); recent sign-up years include years 2018–2021 (N = 898); the full sample includes all sign-up years from 2015–2021 (N= 1392).

Table A.3
Household characteristics by group–recent sign-up years.

		Control (T0)	Continuous (T1)	Interrupted (T2)	Restarted (T3)	Total
Zone	1	25	20	16	15	76
	2	9	13	16	14	52
	3*	12	22	19	13	66
	4*	18	9	14	12	53
	5*	9	11	13	7	40
	6*	12	12	15	8	47
	7	5	6	10	5	26
	8*	7	7	8	6	28
	9	10	14	9	20	53
	10	14	12	17	16	59
	11*	57	47	52	62	218
	12	6	11	10	11	38
	13*	10	7	7	14	38
	14*	22	28	30	24	104
Sign-up year	2018	16	12	14	15	57
	2019	90	82	85	86	343
	2020	78	90	105	94	367
	2021	32	35	32	32	131
House type	Single family house	63	55	56	65	239
	Multiple dwelling unit	79	92	107	89	367
	n/a	74	72	73	73	292
Baseline recycling	yes	42	38	48	44	172
	no	174	181	188	183	726
	N	216	219	236	227	898

Notes: The table shows the number of households per group for different household characteristics. A * behind the zone number means that bags are collected on two days per week in the respective zone, without a star means that they are collected once a week. Only households included that signed up in 2018 or later. Balance tests using chi2 tests show that household characteristics are balanced across groups: p = 0.682 for zone, p = 0.896 for sign-up year, p = 0.094 for house type, p = 0.875 for baseline recycling.

Appendix B. Histograms of recycling frequency per period

See Figs. B.1–B.4

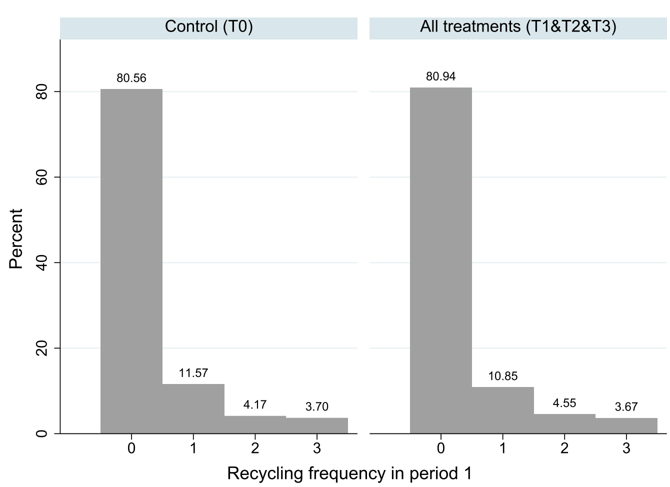

Fig. B.1. Histograms of recycling frequency during the baseline period (week 1–3), comparing all treatment groups together with the control group. *Notes:* The figure shows histograms for the recycling frequency during the baseline period (from zero to three times) in the control group and in the aggregated treatment groups, focusing on households with a recent sign-up year (≥ 2018). N in the control group (T0) = 216, the treatment groups together add up to N = 682.

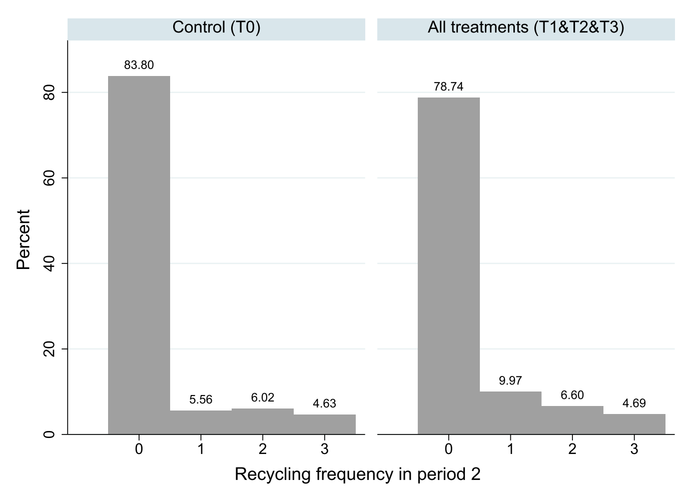

Fig. B.2. Histograms of recycling frequency during period 2 (week 4–6), comparing all treatment groups together with the control group. *Notes:* The figure shows histograms for the recycling frequency during period 2 (from zero to three times) in the control group and in the aggregated treatment groups, focusing on households with a recent sign-up year (≥ 2018). N in the control group (T0) = 216, the treatment groups together add up to N = 682.

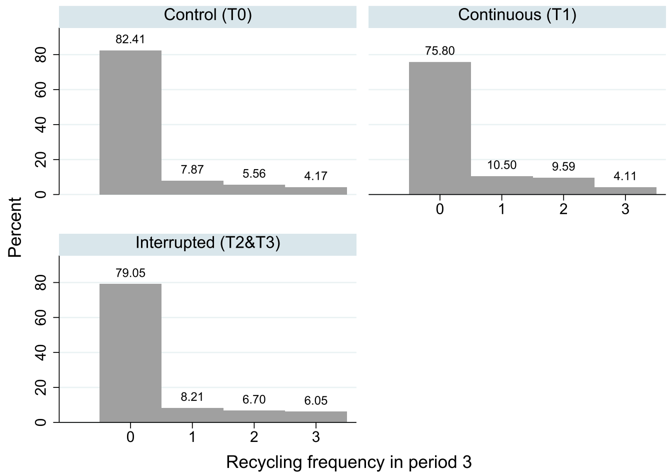

Fig. B.3. Histograms of recycling frequency during period 3 (week 7–9), comparing the continuous and the interrupted treatment groups with the control group. *Notes:* The figure shows histograms for the recycling frequency during period 3 (from zero to three times) in the control group and in the continuous and interrupted treatment groups, focusing on households with a recent sign-up year (≥ 2018). N(T0) = 216, N(T1) = 219, N(T2&T3) = 463.

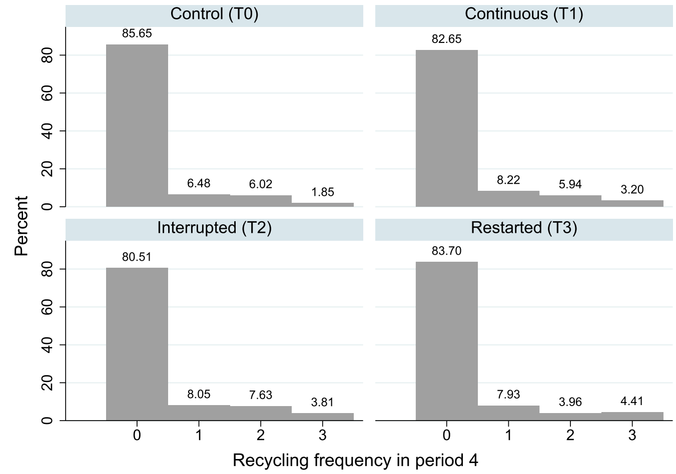

Fig. B.4. Histograms of recycling frequency during period 4 (week 10–12), comparing the continuous, interrupted and restarted treatment groups with the control group. *Notes:* The figure shows histograms for the recycling frequency during period 4 (from zero to three times) in the control group and in the different treatment groups, focusing on households with a recent sign-up year (≥ 2018). N(T0) = 216, N(T1) = 219, N(T2) = 236, N(T3) = 227.

Appendix C. Graphs on weekly recycling behavior

See Fig. C.1

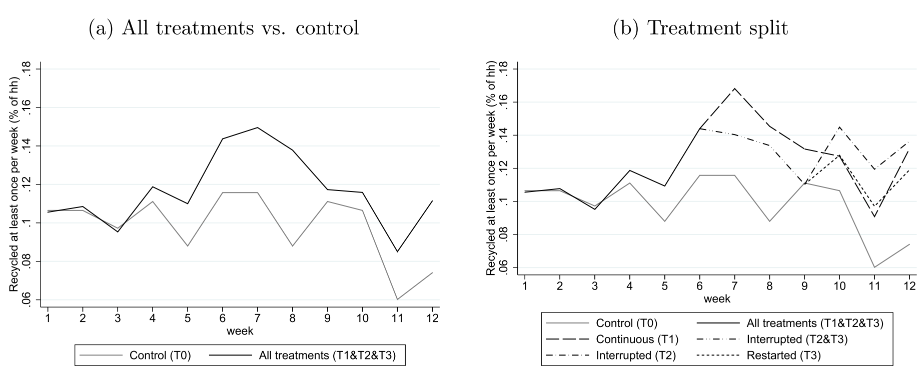

Fig. C.1. Percentages of households that recycled per week. *Notes:* The figure shows the percentages of households that recycled at least once per week in the control group and in the different treatment groups over the whole study period, focusing on households with a recent sign-up year (≥ 2018). (a) compares all treatment groups pooled together with the control group; (b) splits up the different treatment groups based on the treatment variation by periods. N(T0) = 216, N(T1&T2&T3) = 682, N(T1) = 219, N(T2&T3) = 463, N(T2) = 236, N(T3) = 227.

Appendix D. Additional analyses

D.1. Logit robustness check for treatment effects by period

See Tables D.1

Table D.1
Treatment effects on whether households recycled at least once, by period (logit).

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.509**	0.551**		
	(0.214)	(0.280)		
Continuous (T1)			0.858***	0.495
			(0.313)	(0.320)
Interrupted & Restarted (T2&T3)			0.397	
			(0.284)	
Interrupted (T2)				0.560*
				(0.307)
Restarted (T3)				0.265
				(0.313)
Constant	−3.063***	−2.945***	−3.070***	−3.264***
	(0.231)	(0.303)	(0.292)	(0.277)
Pseudo R^2	0.308	0.293	0.342	0.309
Observations	2694	898	898	898

Notes: The table reports logistic regressions with dependent variable equal to one if a household recycled at least once during the respective period. Only households included that signed up in 2018 or later. All regressions include a control for households' binary baseline recycling activity and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.4.

D.2. OLS with interactions with baseline recycling dummy

See Table D.2

Table D.2
Treatment effects on whether households recycled at least once, by period–interactions with baseline recycling yes vs. no.

	Period 2-4		Period 2		Period 3		Period 4	
	(1)		(2)		(3)		(4)	
	BR = yes	BR = no	BR = yes	BR = no	BR = yes	BR = no	BR = yes	BR = no
All treatments (T1&T2&T3)	0.100	0.027*	0.234**	0.009				
	(0.072)	(0.017)	(0.091)	(0.024)				
Continuous (T1)					0.145	0.056*	−0.036	0.049**
					(0.100)	(0.029)	(0.112)	(0.025)
Interrupted & Restarted (T2&T3)					0.055	0.023		
					(0.092)	(0.023)		
Interrupted (T2)							0.010	0.047*
							(0.106)	(0.024)
Restarted (T3)							−0.018	0.023
							(0.109)	(0.023)
Predicted Probability Control (T0)	0.564***	0.061***	0.490***	0.082***	0.617***	0.066***	0.586***	0.036**
	(0.063)	(0.014)	(0.078)	(0.021)	(0.075)	(0.018)	(0.076)	(0.014)
Adjusted R^2	0.358		0.356		0.400		0.338	
Observations	2694		898		898		898	

Notes: BR = baseline recycling. The table reports OLS regressions with dependent variable equal to one if a household recycled at least once during the respective period, columns (1)–(4). For each treatment combination, the table shows coefficients resulting from interactions with the baseline recycling dummy (which can be either yes or no). Only households included that signed up in 2018 or later. All regressions include a control for households' binary baseline recycling activity and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level. Using two-sided tests to compare treatment coefficients gives the following p-values: for column (3), T1 = T2&T3 when BR = yes gives p = 0.267, and p = 0.229 when BR = no. For column (4), T1 = T2 when BR = yes gives p = 0.671, and p = 0.943 when BR = no; for T1 = T3 when BR = yes gives p = 0.872, and p = 0.337 when BR = no; for T3 = T2 when BR = yes gives p = 0.791, and p = 0.365 when BR = no. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

D.3. Intensive margin - recycling frequency per period

See Table D.3–D.5.

Table D.3
Treatment effects on how often households recycled per period.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.075**	0.057		
	(0.037)	(0.047)		
Continuous (T1)			0.145**	0.086*
			(0.057)	(0.048)
Interrupted & Restarted (T2&T3)			0.072	
			(0.049)	
Interrupted (T2)				0.091*
				(0.049)
Restarted (T3)				0.042
				(0.051)
Constant	0.039	0.058	0.032	0.026
	(0.034)	(0.043)	(0.043)	(0.034)
Adjusted R^2	0.449	0.476	0.456	0.419
Observations	2694	898	898	898

Notes: The table reports OLS regressions with dependent variable being how often a household recycled during the respective period (never up to three times). Only households included that signed up in 2018 or later. All regressions include a control for households' baseline recycling activity (never up to three times) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.147 for T1 = T2&T3; column (4) p = 0.922 for T2 = T1, p = 0.422 for T1 = T3, p = 0.376 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.5.

Table D.4
Treatment effects on how often households recycled per period–baseline recycling yes.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.162	0.200		
	(0.153)	(0.200)		
Continuous (T1)			0.335	0.097
			(0.212)	(0.208)
Interrupted & Restarted (T2&T3)			0.169	
			(0.199)	
Interrupted (T2)				0.090
				(0.202)
Restarted (T3)				0.011
				(0.226)
Constant	0.295	0.252	0.384	0.254
	(0.250)	(0.322)	(0.306)	(0.275)
Adjusted R^2	0.226	0.217	0.226	0.191
Observations	516	172	172	172

Notes: The table reports OLS regressions with dependent variable being how often a household recycled during the respective period (never up to three times). Only households included that signed up in 2018 or later; only households that recycled during the baseline period. All regressions include a control for households' baseline recycling activity (never up to three times) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.338 for T1 = T2&T3; column (4) p = 0.971 for T2 = T1, p = 0.708 for T1 = T3, p = 0.728 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.6.

Table D.5
Treatment effects on how often households recycled per period–baseline recycling no.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.050*	0.017		
	(0.026)	(0.033)		
Continuous (T1)			0.104**	0.084**
			(0.050)	(0.037)
Interrupted & Restarted (T2&T3)			0.041	
			(0.037)	
Interrupted (T2)				0.086**
				(0.037)
Restarted (T3)				0.043
				(0.033)
Constant	0.034	0.070**	0.026	0.005
	(0.023)	(0.030)	(0.032)	(0.019)
Adjusted R^2	0.021	0.003	0.040	0.018
Observations	2178	726	726	726

Notes: The table reports OLS regressions with dependent variable being how often a household recycled during the respective period (never up to three times). Only households included that signed up in 2018 or later; only households that did not recycle during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.198 for T1 = T2&T3; column (4) p = 0.997 for T2 = T1, p = 0.355 for T1 = T3, p = 0.327 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.7.

D.4. Intensive margin-number of bags recycled per period

See Table D.6–D.8.

Table D.6
Treatment effects on the number of bags recycled by households per period.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.035	0.057		
	(0.121)	(0.129)		
Continuous (T1)			0.177	−0.008
			(0.188)	(0.153)
Interrupted & Restarted (T2&T3)			−0.052	
			(0.165)	
Interrupted (T2)				0.054
				(0.169)
Restarted (T3)				0.027
				(0.189)
Constant	0.254**	0.205	0.300*	0.257*
	(0.128)	(0.130)	(0.162)	(0.140)
Adjusted R^2	0.459	0.490	0.462	0.422
Observations	2694	898	898	898

Notes: The table reports OLS regressions with dependent variable being how many bags a household recycled during the respective period. Only households included that signed up in 2018 or later. All regressions include a control for households' baseline recycling activity (mean of recycled bags) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.139 for T1 = T2&T3; column (4) p = 0.698 for T2 = T1, p = 0.844 for T3 = T1, p = 0.895 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.8.

Table D.7
Treatment effects on the number of bags recycled by households per period–baseline recycling yes.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	−0.451	−0.124		
	(0.575)	(0.619)		
Continuous (T1)			−0.168	−0.918
			(0.807)	(0.737)
Interrupted & Restarted (T2&T3)			−0.649	
			(0.762)	
Interrupted (T2)				−0.687
				(0.799)
Restarted (T3)				−0.590
				(0.937)
Constant	2.039*	1.171	2.409*	2.547*
	(1.115)	(1.133)	(1.311)	(1.318)
Adjusted R^2	0.322	0.349	0.327	0.246
Observations	516	172	172	172

Notes: The table reports OLS regressions with dependent variable being how many bags a household recycled during the respective period. Only households included that signed up in 2018 or later; only households that recycled during the baseline period. All regressions include a control for households' baseline recycling activity (mean of recycled bags) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.448 for T1 = T2&T3; column (4) p = 0.744 for T2 = T1, p = 0.697 for T3 = T1, p = 0.915 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.9.

Table D.8
Treatment effects on the number of bags recycled by households per period–baseline recycling no.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.139**	0.096		
	(0.058)	(0.070)		
Continuous (T1)			0.280**	0.204**
			(0.134)	(0.079)
Interrupted & Restarted (T2&T3)			0.061	
			(0.089)	
Interrupted (T2)				0.216**
				(0.089)
Restarted (T3)				0.143
				(0.087)
Constant	0.095*	0.139**	0.119	0.026
	(0.050)	(0.062)	(0.078)	(0.041)
Adjusted R^2	0.018	−0.001	0.037	0.020
Observations	2178	726	726	726

Notes: The table reports OLS regressions with dependent variable being how many bags a household recycled during the respective period. Only households included that signed up in 2018 or later; only households that did not recycle during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.091 for T1 = T2&T3; column (4) p = 0.915 for T2 = T1, p = 0.578 for T1 = T3, p = 0.526 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.10.

Appendix E. Full sample results

We conducted all analyses presented in this paper also with the full sample to verify whether and how any findings differ. As in our main analysis, we start by looking at the treatment effects by period (see Tables E.1 and E.2). Again, we find that reminders significantly increase households' recycling activity in period 2 among all treatment groups compared to the control group, and that this effect persists throughout the entire intervention period. In period 3, we observe significant positive effects in the continuous treatment (but not in the interrupted ones). In period 4, we find significant positive effects in the continuous and the interrupted treatment (but not in the restarted group). Compared to results for households with a recent sign-up year, we note that coefficients tend to be smaller and significance levels weaker (many results are now significant at the 10% level). The pattern persists when splitting up the analysis by baseline recycling activity (see Appendix E.4) or when looking at treatment effects on the intensive margin (see Appendices E.7 and E.8).

The same pattern holds when looking at the overall treatment effects over the whole intervention period. We find again that results are fairly similar to the ones presented for households with a more recent sign-up year (see Table E.8). Both the continuous and the interrupted treatments have a significant positive effect on households' recycling activity over the whole intervention period, both regarding the number of periods a household recycled during the intervention period as well as whether a household recycled at all. Compared to the results presented for households that signed up more recently, we can note again that the coefficients for the full sample become slightly smaller.

The weaker effects that we observe among the full sample are not surprising as we now look at a larger group of households among which many may have changed their phone numbers by now or moved away from the district given their early sign-up years. Table A.2 in Appendix A confirms that the percentage of households that recycled at all during the intervention period was higher among households with a recent sign-up year than among those with an older sign up year (< 2018), indicating that those recently signed-up households were more responsive to our intervention. We interpret these findings as support for our decision to focus our main analysis on those households that signed up more recently.

E.1. Full sample histograms of recycling frequency per period

See Figs. E.1–E.4.

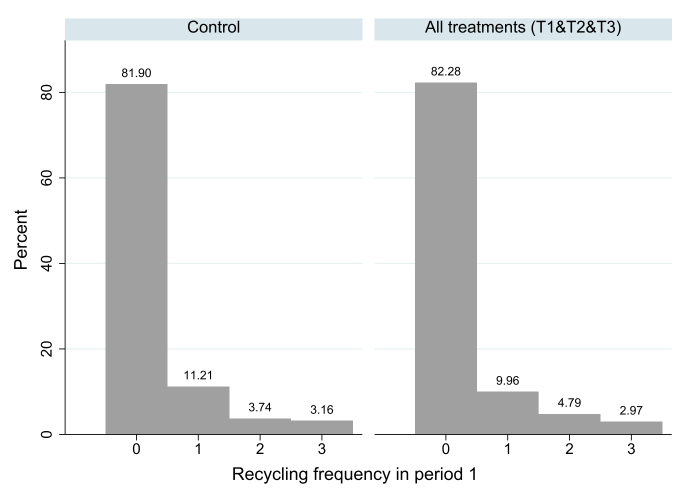

Fig. E.1. Histograms of recycling frequency during the baseline period (week 1–3), comparing all treatment groups together with the control group–full sample. *Notes:* The figure shows histograms for the recycling frequency during the baseline period (from zero to three times) in the control group and in the aggregated treatment groups, including the full sample. N per group = 348, the treatment groups together add up to N = 1,044.

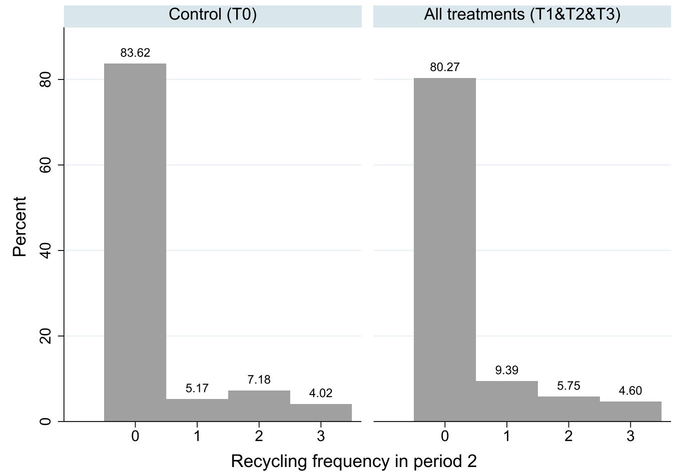

Fig. E.2. Histograms of recycling frequency during period 2 (week 4–6), comparing all treatment groups together with the control group – full sample. *Notes:* The figure shows histograms for the recycling frequency during period 2 (from zero to three times) in the control group and in the aggregated treatment groups, including the full sample. N per group = 348, the treatment groups together add up to N = 1044.

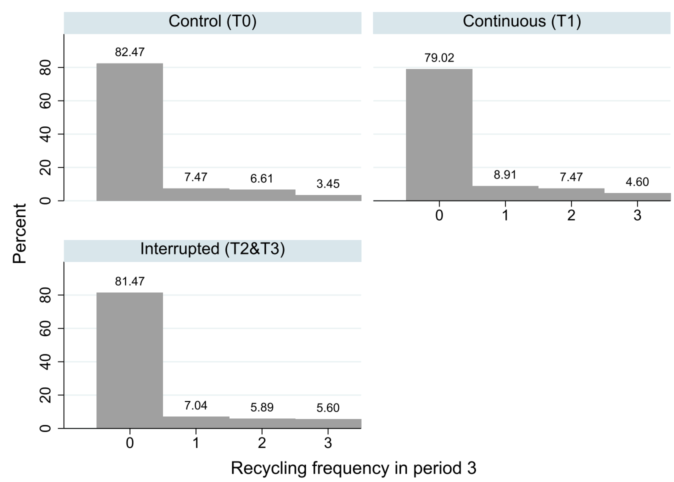

Fig. E.3. Histograms of recycling frequency during period 3 (week 7–9), comparing the continuous and the interrupted treatment groups with the control group – full sample. *Notes:* The figure shows histograms for the recycling frequency during period 3 (from zero to three times) in the control group and in the continuous and interrupted treatment groups, including the full sample. N per group = 348 (hence N = 696 for both interrupted groups together).

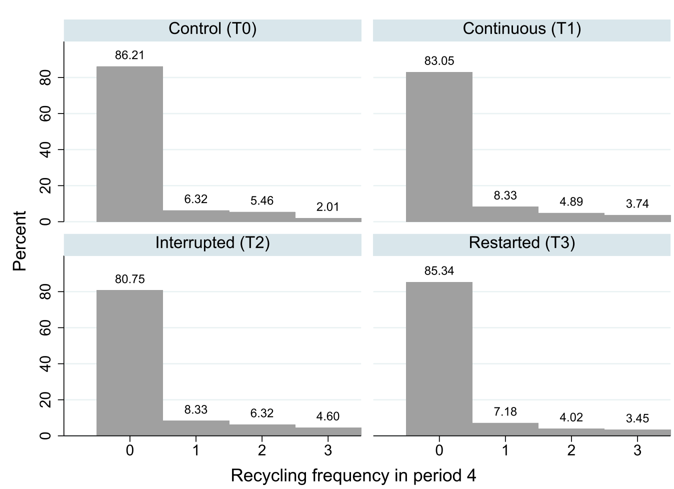

Fig. E.4. Histograms of recycling frequency during period 4 (week 10–12), comparing the continuous, interrupted and restarted treatment groups with the control group – full sample. *Notes:* The figure shows histograms for the recycling frequency during period 4 (from zero to three times) in the control group and in the different treatment groups, including the full sample. N per group = 348.

E.2. Full sample graphs on weekly recycling behavior

See Fig. E.5.

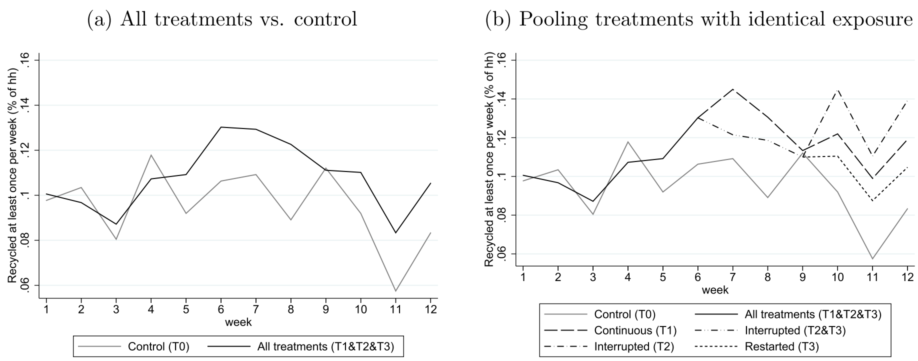

Fig. E.5. Percentages of households that recycled per week – full sample. *Notes:* The figure shows the percentages of households that recycled at least once per week in the control group and in the different treatment groups over the whole study period, focusing on the whole sample. (a) compares all treatment groups pooled together with the control group; (b) pools treatment groups that have identical exposure as of each period in the horizontal axis. N(T0) = 348, N(T1&T2&T3) = 1,044, N(T1) = 348, N(T2&T3) = 698, N(T2) = 348, N(T3) = 348.

E.3. Full sample results: treatment effects by period

See Fig. E.6, Tables E.1 and E.2.

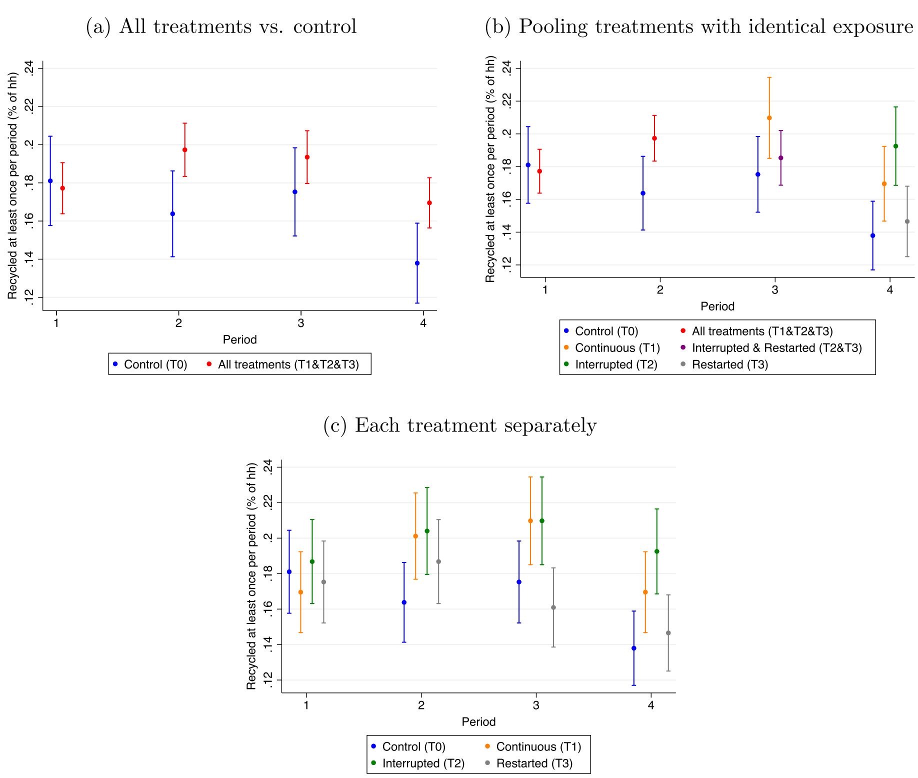

Fig. E.6. Proportion of households that recycled at least once per period–full sample. *Notes:* The figure shows the proportion of households that recycled at least once per period in the control group and in the different treatment groups over the whole study period, focusing on the whole sample. (a) compares all treatment groups pooled together with the control group; (b) pools treatment groups that have identical exposure as of each period in the horizontal axis; and (c) compares the control group with each treatment group separately. N(T0) = 348, N(T1&T2&T3) = 1,044, N(T1) = 348, N(T2&T3) = 698, N(T2) = 348, N(T3) = 348.

Table E.1
Treatment effects on whether households recycled at least once, by period–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.032**	0.037*		
	(0.016)	(0.020)		
Continuous (T1)			0.044*	0.040*
			(0.025)	(0.023)
Interrupted & Restarted (T2&T3)			0.014	
			(0.021)	
Interrupted (T2)				0.053**
				(0.023)
Restarted (T3)				0.013
				(0.022)
Constant	0.052***	0.058***	0.065***	0.031*
	(0.016)	(0.020)	(0.020)	(0.017)
Adjusted R^2	0.325	0.332	0.331	0.312
Observations	4176	1392	1392	1392

Notes: The table reports OLS regressions with dependent variable equal to one if a household recycled at least once during the respective period. Full sample included. All regressions include a control for households' binary baseline recycling activity and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.162 for T1 = T2&T3; column (4) p = 0.583 for T2 = T1, p = 0.250 for T1 = T3, p = 0.088 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.11.

Table E.2
Treatment effects on whether households recycle at least once, by period–full sample (logit).

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.337**	0.372*		
	(0.170)	(0.210)		
Continuous (T1)			0.433*	0.448*
			(0.244)	(0.256)
Interrupted & Restarted (T2&T3)			0.157	
			(0.216)	
Interrupted (T2)				0.581**
				(0.251)
Restarted (T3)				0.171
				(0.254)
Constant	−2.787***	−2.673***	−2.654***	−3.109***
	(0.193)	(0.237)	(0.222)	(0.241)
Pseudo R^2	0.282	0.284	0.290	0.287
Observations	4176	1392	1392	1392

Notes: The table reports logistic regressions with dependent variable equal to one if a household recycled at least once during the respective period. Full sample included. All regressions include a control for households' binary baseline recycling activity and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.12.

E.4. Full sample results: heterogeneity by recycling at baseline

See Tables E.3 and E.4.

Table E.3
Treatment effects on whether households recycled at least once, by period–baseline recycling yes–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.106*	0.154**		
	(0.058)	(0.072)		
Continuous (T1)			0.167**	0.061
			(0.083)	(0.089)
Interrupted & Restarted (T2&T3)			0.075	
			(0.074)	
Interrupted (T2)				0.084
				(0.086)
Restarted (T3)				0.031
				(0.089)
Constant	0.635***	0.622***	0.688***	0.601***
	(0.070)	(0.085)	(0.082)	(0.087)
Adjusted R^2	0.018	0.009	0.027	−0.007
Observations	744	248	248	248

Notes: The table reports OLS regressions with dependent variable equal to one if a household recycled at least once during the respective period. Full sample included; only households that recycled during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.198 for T1 = T2&T3; column (4) p = 0.795 for T2 = T1, p = 0.756 for T1 = T3, p = 0.553 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.13.

Table E.4
Treatment effects on whether households recycled at least once, by period–baseline recycling no–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.015	0.011		
	(0.014)	(0.018)		
Continuous (T1)			0.019	0.036*
			(0.024)	(0.020)
Interrupted & Restarted (T2&T3)			−0.002	
			(0.019)	
Interrupted (T2)				0.045**
				(0.021)
Restarted (T3)				0.007
				(0.018)
Constant	0.056***	0.070***	0.067***	0.030**
	(0.015)	(0.019)	(0.018)	(0.015)
Adjusted R^2	0.012	0.003	0.026	0.011
Observations	3432	1144	1144	1144

Notes: The table reports OLS regressions with dependent variable equal to one if a household recycled at least once during the respective period. Full sample included; only households that did not recycle during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.306 for T1 = T2&T3; column (4) p = 0.710 for T2 = T1, p = 0.161 for T1 = T3, p = 0.078 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Panel regression analog is reported in Table F.14.

E.5. Full sample results: mechanisms

See Tables E.5–E.7.

Table E.5
Treatment effects on recycling in period 2 and 3 combined–full sample.

	Recycled p2 & p3 (1)	Recycled neither p2 nor p3 (2)	Recycled only p2 (3)	Recycled only p3 (4)
Continuous (T1)	0.033	−0.077**	0.022	0.022
	(0.023)	(0.033)	(0.018)	(0.018)
Interrupted & Restarted (T2&T3)	0.019	−0.039	0.018	0.002
	(0.018)	(0.027)	(0.014)	(0.014)
Predicted Probability Control (T0)	0.072***	0.843***	0.040***	0.040***
	(0.015)	(0.022)	(0.011)	(0.012)
Observations	1392	1392	1392	1392

Notes: Marginal probabilities of a multinomial probit regressions on (1) whether households recycled at least once in both period 2 and period 3, (2) whether households recycled neither in period 2 nor in period 3, (3) whether households recycled only in period 2, (4) whether households recycled only in period 3. The regression includes a control for households' binary baseline recycling activity and the house type, and uses fixed effects for enumerators. Full sample included. Standard errors are in parentheses. Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (1) p = 0.492 for T1 = T2&T3; column (2) p = 0.217 for T1 = T2&T3; column (3) p = 0.822 for T1 = T2&T3; column (4) p = 0.246 for T1 = T2&T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table E.6
Treatment effects on recycling in period 2 and 3 combined–baseline recycling yes–full sample.

	Recycled p2 & p3 (1)	Recycled neither p2 nor p3 (2)	Recycled only p2 (3)	Recycled only p3 (4)
Continuous (T1)	0.151*	−0.148*	−0.019	0.016
	(0.091)	(0.077)	(0.049)	(0.065)
Interrupted & Restarted (T2&T3)	0.103	−0.140**	0.067	−0.031
	(0.079)	(0.069)	(0.049)	(0.052)
Predicted Probability Control (T0)	0.466***	0.314***	0.085**	0.135***
	(0.064)	(0.060)	(0.037)	(0.044)
Observations	248	248	248	248

Notes: Marginal probabilities of a multinomial probit regressions on (1) whether households recycled at least once in both period 2 and period 3, (2) whether households recycled neither in period 2 nor in period 3, (3) whether households recycled only in period 2, (4) whether households recycled only in period 3. The regression includes a control for the house type, and uses fixed effects for enumerators. Full sample included. Only households included that recycled during the baseline period. Standard errors are in parentheses. Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (1) p = 0.544 for T1 = T2&T3; column (2) p = 0.889 for T1 = T2&T3; column (3) p = 0.060 for T1 = T2&T3; column (4) p = 0.396 for T1 = T2&T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table E.7
Treatment effects on recycling in period 2 and 3 combined–baseline recycling no–full sample.

	Recycled p2 & p3 (1)	Recycled neither p2 nor p3 (2)	Recycled only p2 (3)	Recycled only p3 (4)
Continuous (T1)	0.007	−0.044	0.023	0.014
	(0.017)	(0.027)	(0.017)	(0.015)
Interrupted & Restarted (T2&T3)	−0.001	−0.007	0.006	−0.003
	(0.014)	(0.021)	(0.013)	(0.012)
Predicted Probability Control (T0)	0.041***	0.904***	0.028***	0.027***
	(0.012)	(0.018)	(0.010)	(0.010)
Observations	1144	1144	1144	1144

Notes: Marginal probabilities of a multinomial probit regressions on (1) whether households recycled at least once in both period 2 and period 3, (2) whether households recycled neither in period 2 nor in period 3, (3) whether households recycled only in period 2, (4) whether households recycled only in period 3. The regression includes a control for the house type, and uses fixed effects for enumerators. Full sample included. Only households included that did not recycle during the baseline period. Standard errors are in parentheses. Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (1) p = 0.576 for T1 = T2&T3; column (2) p = 0.123 for T1 = T2&T3; column (3) p = 0.242 for T1 = T2&T3; column (4) p = 0.413 for T1 = T2&T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

E.6. Full sample results: treatment effects over the entire intervention period

See Fig. E.7 and Table E.8.

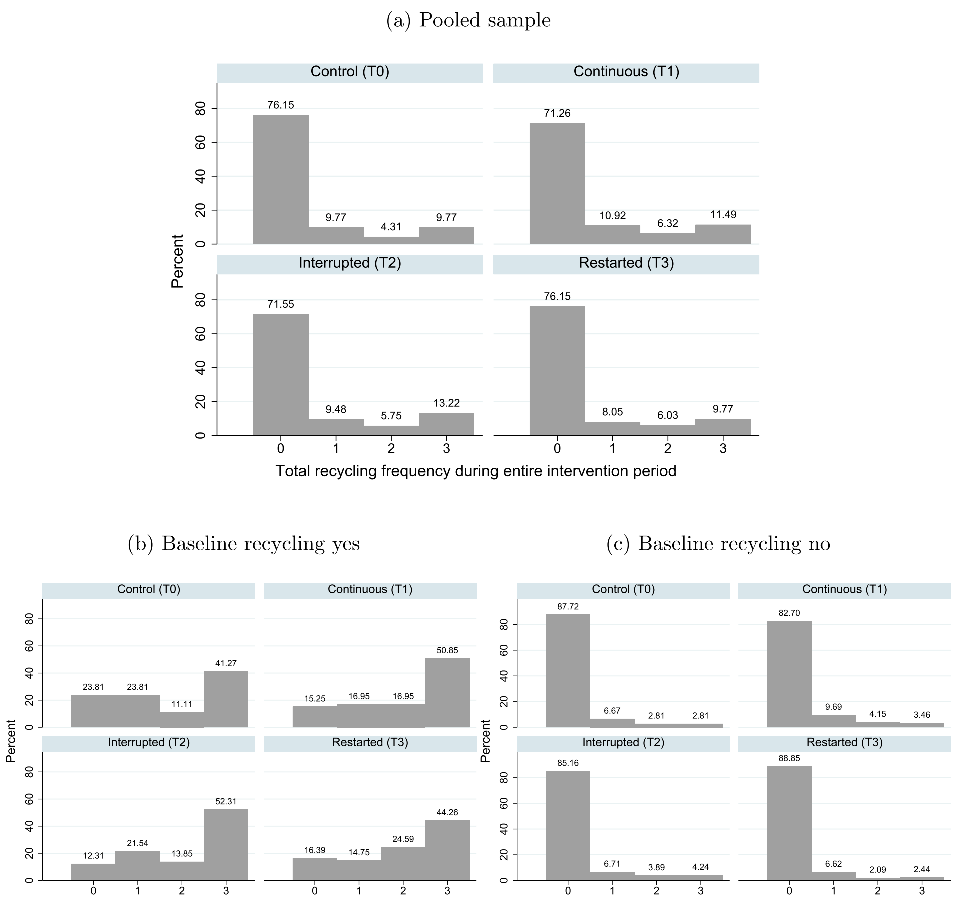

Fig. E.7. Total recycling frequency during entire intervention period, by treatment – full sample. *Notes:* The figure shows histograms for the total recycling frequency during the entire intervention period (period 2-4). Concretely, it shows in how many of the different periods of the intervention period households recycled at least once (from zero to three periods). Only households included with a recent sign-up year (≥ 2018). (a) for pooled sample: N per group = 348; (b) for households that recycled during the baseline period: N(T0) = 63, N(T1) = 59, N(T2) = 65, N(T3) = 61; (c) for households that did not recycle during the baseline period: N(T0) = 285, N(T1) = 289, N(T2) = 283, N(T3) = 287.

Table E.8
Treatment effects on recycling over entire intervention period-full sample.

Recycled in …	no period (1)	one period (2)	two periods (3)	three periods (4)
Panel A: Pooled sample (N = 1,392)				
Continuous (T1)	−0.075**	0.029**	0.019**	0.027**
	(0.033)	(0.013)	(0.009)	(0.012)
Interrupted (T2)	−0.654**	0.025**	0.017**	0.023**
	(0.033)	(0.013)	(0.009)	(0.012)
Restarted (T3)	−0.010	0.004	0.003	0.003
	(0.031)	(0.013)	(0.008)	(0.010)
Predicted Probability Control (T0)	0.807***	0.106***	0.046***	0.041***
	(0.022)	(0.012)	(0.007)	(0.008)
Panel B: Recycled at baseline (N = 248)				
Continuous (T1)	−0.085	−0.039	−0.005	0.130
	(0.053)	(0.025)	(0.007)	(0.080)
Interrupted (T2)	−0.086*	−0.040	−0.005	0.132*
	(0.052)	(0.025)	(0.007)	(0.078)
Restarted (T3)	−0.057	−0.023	−0.001	0.081
	(0.056)	(0.023)	(0.004)	(0.080)
Predicted Probability Control (T0)	0.222***	0.220***	0.171***	0.386***
	(0.045)	(0.031)	(0.024)	(0.057)
Panel C: No recycling at baseline (N = 1,144)				
Continuous (T1)	−0.050*	0.022*	0.012*	0.016*
	(0.028)	(0.013)	(0.007)	(0.009)
Interrupted (T2)	−0.036	0.016	0.009	0.011
	(0.028)	(0.013)	(0.007)	(0.009)
Restarted (T3)	0.009	−0.004	−0.002	−0.002
	(0.026)	(0.012)	(0.006)	(0.007)
Predicted Probability Control (T0)	0.886***	0.065***	0.026***	0.023***
	(0.019)	(0.010)	(0.006)	(0.006)

Notes: Marginal probabilities of ordered probit regressions on (1) whether households did not recycle at all during the entire intervention period, (2) whether households recycled in one period of the intervention period, (3) whether households recycled in two periods of the intervention period, (4) whether households recycled in three periods of the intervention period. All regressions include a control for the house type, and use fixed effects for enumerators. The regressions for the pooled sample also include a control for households' binary baseline recycling activity. Full sample included. Standard errors in parentheses. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Comparing the marginal probabilities of T1, T2 and T3, respectively, using a two-sided Wald test gives the following p-values:
Panel A: column (1) T1 = T2 p = 0.788, T1 = T3 p = 0.051, T2 = T3 p = 0.095; column (2) T1 = T2 p = 0.788, T1 = T3 p = 0.054, T2 = T3 p = 0.097; column (3) T1 = T2 p = 0.788, T1 = T3 p = 0.055, T2 = T3 p = 0.099; column (4) T1 = T2 p = 0.788, T1 = T3 p = 0.056, T2 = T3 p = 0.099;
Panel B: column (1) T1 = T2 p = 0.979, T1 = T3 p = 0.561, T2 = T3 p = 0.536; column (2) T1 = T2 p = 0.979, T1 = T3 p = 0.562, T2 = T3 p = 0.534; column (3) T1 = T2 p = 0.979, T1 = T3 p = 0.581, T2 = T3 p = 0.551; column (4) T1 = T2 p = 0.979, T1 = T3 p = 0.560, T2 = T3 p = 0.534;
Panel C: column (1) T1 = T2 p = 0.623, T1 = T3 p = 0.033, T2 = T3 p = 0.106; column (2) T1 = T2 p = 0.623, T1 = T3 p = 0.036, T2 = T3 p = 0.109; column (3) T1 = T2 p = 0.623, T1 = T3 p = 0.041, T2 = T3 p = 0.116; column (4) T1 = T2 p = 0.623, T1 = T3 p = 0.041, T2 = T3 p = 0.116.

E.7. Full sample results: intensive margin - recycling frequency per period

See Tables E.9–E.11.

Table E.9
Treatment effects on how often households recycled per period–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.052*	0.030		
	(0.030)	(0.036)		
Continuous (T1)			0.088*	0.078*
			(0.047)	(0.041)
Interrupted & Restarted (T2&T3)			0.042	
			(0.040)	
Interrupted (T2)				0.100**
				(0.041)
Restarted (T3)				0.024
				(0.039)
Constant	0.073**	0.088**	0.100***	0.030
	(0.030)	(0.037)	(0.037)	(0.031)
Adjusted R^2	0.430	0.481	0.417	0.396
Observations	4176	1392	1392	1392

Notes: The table reports OLS regressions with dependent variable being how often a household recycled during the respective period (never up to three times). Full sample included. All regressions include a control for households' baseline recycling activity (never up to three times) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.255 for T1 = T2&T3; column (4) p = 0.621 for T2 = T1, p = 0.212 for T1 = T3, p = 0.082 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table E.10
Treatment effects on how often households recycled per period–baseline recycling yes–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.178	0.138		
	(0.122)	(0.156)		
Continuous (T1)			0.315*	0.156
			(0.173)	(0.178)
Interrupted & Restarted (T2&T3)			0.210	
			(0.156)	
Interrupted (T2)				0.204
				(0.173)
Restarted (T3)				0.085
				(0.186)
Constant	0.125	0.084	0.273	0.026
	(0.201)	(0.258)	(0.238)	(0.216)
Adjusted R^2	0.240	0.238	0.269	0.193
Observations	744	248	248	248

Notes: The table reports OLS regressions with dependent variable being how often a household recycled during the respective period (never up to three times). Full sample included; only households that recycled during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.481 for T1 = T2&T3; column (4) p = 0.791 for T2 = T1, p = 0.707 for T1 = T3, p = 0.529 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table E.11
Treatment effects on how often households recycled per period–baseline recycling no–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.025	0.007		
	(0.023)	(0.027)		
Continuous (T1)			0.045	0.064**
			(0.041)	(0.032)
Interrupted & Restarted (T2&T3)			0.003	
			(0.033)	
Interrupted (T2)				0.077**
				(0.033)
Restarted (T3)				0.013
				(0.027)
Constant	0.077***	0.101***	0.094***	0.036
	(0.024)	(0.028)	(0.031)	(0.023)
Adjusted R^2	0.009	0.000	0.016	0.010
Observations	3432	1144	1144	1144

Notes: The table reports OLS regressions with dependent variable being how often a household recycled during the respective period (never up to three times). Full sample included; only households that did not recycle during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.268 for T1 = T2&T3; column (4) p = 0.741 for T2 = T1, p = 0.133 for T1 = T3, p = 0.070 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

E.8. Full sample results: intensive margin-number of bags recycled per period

See Tables E.12–E.14.

Table E.12
Treatment effects on the number of bags recycled by households per period–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.050	0.049		
	(0.103)	(0.129)		
Continuous (T1)			0.108	0.079
			(0.151)	(0.132)
Interrupted & Restarted (T2&T3)			−0.041	
			(0.137)	
Interrupted (T2)				0.161
				(0.140)
Restarted (T3)				0.037
				(0.141)
Constant	0.309**	0.289*	0.435***	0.200*
	(0.125)	(0.171)	(0.139)	(0.121)
Adjusted R^2	0.442	0.470	0.446	0.407
Observations	4176	1392	1392	1392

Notes: The table reports OLS regressions with dependent variable being how many bags a household recycled during the respective period. Full sample included. All regressions include a control for households' baseline recycling activity (mean of recycled bags) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.215 for T1 = T2&T3; column (4) p = 0.554 for T2 = T1, p = 0.765 for T1 = T3, p = 0.399 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table E.13
Treatment effects on the number of bags recycled by households per period–baseline recycling yes–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	−0.070	−0.013		
	(0.514)	(0.688)		
Continuous (T1)			0.147	−0.198
			(0.676)	(0.662)
Interrupted & Restarted (T2&T3)			−0.300	
			(0.653)	
Interrupted (T2)				0.204
				(0.681)
Restarted (T3)				−0.156
				(0.723)
Constant	1.600	1.426	2.158*	1.243
	(1.023)	(1.512)	(1.100)	(0.974)
Adjusted R^2	0.316	0.320	0.357	0.242
Observations	744	248	248	248

Notes: The table reports OLS regressions with dependent variable being how many bags a household recycled during the respective period. Full sample included; only households that recycled during the baseline period. All regressions include a control for households' baseline recycling activity (mean of recycled bags) and the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.379 for T1 = T2&T3; column (4) p = 0.547 for T2 = T1, p = 0.951 for T3 = T1, p = 0.622 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table E.14
Treatment effects on the number of bags recycled by households per period–baseline recycling no–full sample.

	Period 2-4 (1)	Period 2 (2)	Period 3 (3)	Period 4 (4)
All treatments (T1&T2&T3)	0.076	0.070		
	(0.053)	(0.056)		
Continuous (T1)			0.123	0.149*
			(0.107)	(0.078)
Interrupted & Restarted (T2&T3)			−0.004	
			(0.081)	
Interrupted (T2)				0.143*
				(0.078)
Restarted (T3)				0.066
				(0.074)
Constant	0.196***	0.195***	0.271***	0.119*
	(0.057)	(0.057)	(0.079)	(0.068)
Adjusted R^2	0.008	0.000	0.014	0.007
Observations	3432	1144	1144	1144

Notes: The table reports OLS regressions with dependent variable being how many bags a household recycled during the respective period. Full sample included; only households that did not recycle during the baseline period. All regressions include a control for the house type, and use fixed effects for enumerators. Robust standard errors are in parentheses; standard errors are clustered at individual household level in column (1). Comparing treatment coefficients using two-sided Wald tests gives following p-values: column (3) p = 0.196 for T1 = T2&T3; column (4) p = 0.947 for T1 = T2, p = 0.339 for T1 = T3, p = 0.375 for T2 = T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Appendix F. Panel regression analyses

We estimate a set of random-effects (RE) panel regressions[25] using the following specification:

$$Y_{it} = \alpha + \sum_{g=1}^{3} \beta_g \cdot T_{ig} + \sum_{t=2}^{4} \delta_t \cdot P_t + \sum_{g=1}^{3}\sum_{t=2}^{4} \gamma_{gt} \cdot (T_{ig} \cdot P_t) + \mathbf{X}_i'\theta + u_i + \varepsilon_{it} \quad (2)$$

with Y_{it} being a binary variable equal to one if household i recycled at least once in period t, zero otherwise. T_{ig} is a set of indicator variables for the treatment groups, using the control group as reference. P_t includes period fixed effects, with period 1 as the reference. $T_{ig} \cdot P_t$ are interaction terms capturing the effect of each treatment group in each post-treatment period. The vector $\mathbf{X}_i$ includes controls for households' baseline recycling activity, the house type, and indicator variables for the different enumerators. u_i is an unobserved, time-invariant household-specific random effect (random intercept) and ϵ_{it} is an idiosyncratic error that varies across households and periods.

We begin by reporting Table F.1, which presents the direct results from our random–effects (RE) panel specification. As in our primary analysis, the sample is restricted to households that enrolled in 2018 or later. Coefficients on the treatment–by–period interactions capture the effect of each experimental arm (T1, T2, T3) relative to the control within each post-treatment period. To facilitate comparison with the sequence of results in the main analysis, Table F.2 then summarizes these RE estimates in a similar layout as Tables 2 and 3 and the complementary OLS and logit specifications in Appendices D and E. In Table F.2 and subsequent tables, Panel (1) above the horizontal line reports the

Table F.1
Random-effects panel regression with treatment-by-period interactions.

	Coef.	SE
Period effects (baseline: Period 1)		
Period 2	−0.0324	(0.0274)
Period 3	−0.0185	(0.0237)
Period 4	−0.0509**	(0.0220)
Treatments (baseline: Control)		
Continuous (T1)	−0.0051	(0.0117)
Interrupted (T2)	0.0056	(0.0118)
Restarted (T3)	−0.0008	(0.0120)
Interactions: Treatment × Period		
Continuous × Period 2	0.0735**	(0.0373)
Continuous × Period 3	0.0870***	(0.0332)
Continuous × Period 4	0.0509	(0.0334)
Interrupted × Period 2	0.0451	(0.0357)
Interrupted × Period 3	0.0439	(0.0337)
Interrupted × Period 4	0.0425	(0.0326)
Restarted × Period 2	0.0456	(0.0346)
Restarted × Period 3	0.0141	(0.0330)
Restarted × Period 4	0.0201	(0.0324)
Other covariates		
Baseline recycler (=1)	0.6847***	(0.0224)
House	0.0515***	(0.0169)
Constant	0.0476***	(0.0121)
Observations	3592	
Households (groups)	898	

Notes: Random-effects GLS; robust standard errors clustered at the household level. Dependent variable equals 1 if a household recycled at least once in the respective period; period 1 is the baseline. Treatment indicators are relative to the control arm; enumerator indicators are included but not reported. Only households that signed up in 2018 or later are included. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

[25] We conducted a Hausman test (not reported here) comparing fixed- and random-effects panel specifications. The test indicated no systematic difference between the two estimators ($\chi^2(3) = 0.69,\ p = 0.8755$), supporting the use of the random-effects model. The specification used for the Hausman test is not exactly the same as the more complete RE model presented in this section as it only included time-variant regressors.

Table F.2
Treatment effects on whether households recycled at least once, by period (panel regression version of Table 2).

	(1)				(2)
	Period 1	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	−0.005	0.074**	0.087***	0.051	0.070**
	(0.012)	(0.037)	(0.033)	(0.033)	(0.029)
Interrupted (T2)	0.006	0.045	0.044	0.042	0.044
	(0.012)	(0.036)	(0.034)	(0.033)	(0.029)
Restarted (T3)	−0.001	0.046	0.014	0.020	0.027
	(0.012)	(0.035)	(0.033)	(0.032)	(0.028)
All treatments (T1&T2&T3)	0.000	0.054*	0.048*	0.038	0.047*
	(0.010)	(0.030)	(0.027)	(0.026)	(0.024)
Interrupted & Restarted (T2&T3)	0.002	0.045	0.029	0.031	0.035
	(0.010)	(0.032)	(0.029)	(0.028)	(0.025)
Predicted Probability Control (T0)	0.194***	0.162***	0.176***	0.144***	0.160***
	(0.008)	(0.024)	(0.021)	(0.019)	(0.012)
Overall R^2			0.487		
Intraclass correlation			0.265		
Observations			3592		

Notes: The table reports a random effects panel regression where the dependent variable equals one if a household recycled at least once during the respective period. The regression includes controls for households' binary baseline recycling activity, house type, and enumerator fixed effects. Only households that signed up in 2018 or later are included. In Panel (1) and above the horizontal line, we report the main coefficients from a random-effects panel regression with treatment-by-period interactions and the prespecified controls; the T1, T2, and T3 rows are the interaction coefficients read directly from the regression. The remaining entries in that panel (below the horizontal line: pooled arms and the control row) are obtained with Wald tests. Panel (2) reports average effects over periods 2–4; for each arm this is the simple mean of the three period-specific interaction coefficients. When pooling arms (T2&T3 or All treatments), coefficients are combined using sample-size weights from the subsample. Standard errors in parentheses are cluster-robust at the household level; for post-estimation entries they reflect the delta method from the corresponding linear combinations. Two-sided tests. Only households that signed up in 2018 or later are included. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.3
Treatment effects on whether households recycled at least once, by period–baseline recycling yes vs. no (panel regression version of Table 3).

	(1)						(2)	
	Period 2		Period 3		Period 4		Period 2–4	
	BR = yes	BR = no	BR = yes	BR = no	BR = yes	BR = no	BR = yes	BR = no
Continuous (T1)	0.211**	0.030	0.199**	0.053*	0.010	0.048*	0.140*	0.044*
	(0.107)	(0.031)	(0.095)	(0.030)	(0.110)	(0.025)	(0.084)	(0.023)
Interrupted (T2)	0.229**	0.005	0.086	0.038	0.051	0.045*	0.122	0.029
	(0.101)	(0.029)	(0.098)	(0.029)	(0.103)	(0.024)	(0.083)	(0.022)
Restarted (T3)	0.273***	−0.009	0.039	0.008	−0.004	0.026	0.102	0.008
	(0.100)	(0.028)	(0.103)	(0.027)	(0.107)	(0.022)	(0.083)	(0.021)
All treatments (T1&T2&T3)	0.237***	0.008	0.107	0.033	0.019	0.040*	0.121*	0.027
	(0.100)	(0.024)	(0.083)	(0.022)	(0.087)	(0.018)	(0.070)	(0.017)
Interrupted & Restarted (T2&T3)	0.251***	−0.002	0.063	0.023	0.024	0.036*	0.112	0.019
	(0.090)	(0.025)	(0.088)	(0.024)	(0.092)	(0.019)	(0.074)	(0.018)
Predicted Probability Control (T0)	0.490***	0.083***	0.633***	0.066***	0.585***	0.037***	0.569***	0.062***
	(0.078)	(0.021)	(0.074)	(0.018)	(0.076)	(0.014)	(0.063)	(0.014)
Overall R^2					0.523			
Observations					3592			

Notes: BR = baseline recycling. The table reports a random effects panel regression with dependent variable equal to one if a household recycled at least once during the respective period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. The regression includes a control for households' binary baseline recycling activity, house type, and enumerator fixed effects. Only households that signed up in 2018 or later are included. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

treatment–by–period interaction coefficients read directly from the RE regression, while below the horizontal line pooled entries (combined arms and the control row) are obtained via Wald tests; Panel (2) reports average effects across periods 2–4, computed as simple means of the corresponding interaction coefficients, with pooled averages weighted by arm-specific sample sizes in the sample.

Table F.4
Treatment effects on whether households recycled at least once, by period (panel logit regression version of Table D.1).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	1.313*	1.460***	0.984	1.252**
	(0.671)	(0.563)	(0.636)	(0.530)
Interrupted (T2)	0.822	0.762	0.834	0.806
	(0.646)	(0.586)	(0.600)	(0.520)
Restarted (T3)	0.840	0.256	0.389	0.495
	(0.638)	(0.604)	(0.625)	(0.524)
All treatments (T1&T2&T3)	0.986*	0.818*	0.734	0.846*
	(0.566)	(0.490)	(0.498)	(0.440)
Interrupted & Restarted (T2&T3)	0.831	0.514	0.616	0.654
	(0.583)	(0.520)	(0.526)	(0.460)
Predicted Probability Control (T0)	0.162***	0.175***	0.145***	0.161***
	(0.022)	(0.019)	(0.017)	(0.011)
Log pseudolikelihood	−915.481			
Intraclass correlation	0.560			
Observations	3592			

The table reports a random-effects panel logit regression with dependent variable equal to one if a household recycled at least once during the respective period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later. The regression includes a control for households' baseline recycling activity, house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.5
Treatment effects on how often households recycled per period (panel regression version of Table D.3).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.096	0.151**	0.101*	0.116**
	(0.059)	(0.061)	(0.054)	(0.049)
Interrupted (T2)	0.059	0.076	0.078	0.071
	(0.059)	(0.060)	(0.054)	(0.049)
Restarted (T3)	0.022	0.048	0.039	0.036
	(0.060)	(0.060)	(0.058)	(0.051)
All treatments (T1&T2&T3)	0.058	0.091*	0.072*	0.074*
	(0.050)	(0.050)	(0.043)	(0.041)
Interrupted & Restarted (T2&T3)	0.041	0.062	0.059	0.054
	(0.053)	(0.052)	(0.047)	(0.044)
Predicted Probability Control (T0)	0.315***	0.315***	0.241***	0.290***
	(0.042)	(0.040)	(0.031)	(0.022)
Overall R^2	0.562			
Intraclass correlation	0.335			
Observations	3592			

Notes: The table reports a random effects panel regression with dependent variable being how often a household recycled per period (never up to three times). Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later. The regression includes a control for households' baseline recycling activity (never up to three times), house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.6
Treatment effects on how often households recycled per period–baseline recycling yes (panel regression version of Table D.4).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.273	0.357	0.129	0.253
	(0.239)	(0.230)	(0.219)	(0.189)
Interrupted (T2)	0.280	0.149	0.086	0.172
	(0.238)	(0.226)	(0.203)	(0.189)
Restarted (T3)	0.132	0.175	0.001	0.103
	(0.250)	(0.242)	(0.242)	(0.210)
All treatments (T1&T2&T3)	0.228	0.219	0.070	0.172
	(0.208)	(0.198)	(0.177)	(0.166)
Interrupted & Restarted (T2&T3)	0.209	0.161	0.046	0.139
	(0.217)	(0.208)	(0.190)	(0.176)
Predicted Probability Control (T0)	1.190***	1.238***	1.071***	1.167***
	(0.178)	(0.163)	(0.138)	(0.093)
Overall R^2	0.346			
Intraclass correlation	0.322			
Observations	688			

Notes: The table reports a random effects panel regression with dependent variable being how often a household recycled per period (never up to three times). Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later; only households that recycled during the baseline period. The regression includes a control for households' baseline recycling activity (never up to three times), house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.7
Treatment effects on how often households recycled per period–baseline recycling no (panel regression version of Table D.5).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.046	0.096*	0.081**	0.074*
	(0.045)	(0.052)	(0.037)	(0.038)
Interrupted (T2)	0.008	0.062	0.082**	0.051
	(0.041)	(0.048)	(0.037)	(0.035)
Restarted (T3)	−0.005	0.017	0.047	0.020
	(0.040)	(0.044)	(0.033)	(0.033)
All treatments (T1&T2&T3)	0.016	0.058	0.070***	0.048*
	(0.033)	(0.037)	(0.025)	(0.027)
Interrupted & Restarted (T2&T3)	0.002	0.040	0.065**	0.036
	(0.035)	(0.039)	(0.028)	(0.028)
Predicted Probability Control (T0)	0.103***	0.092***	0.040**	0.079***
	(0.028)	(0.029)	(0.017)	(0.015)
Overall R^2	0.041			
Intraclass correlation	0.385			
Observations	2904			

Notes: The table reports a random effects panel regression with dependent variable being how often a household recycled per period (never up to three times). Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later; only households that did not recycle during the baseline period. The regression includes a control for the house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.8
Treatment effects on the number of bags recycled by households per period (panel regression version of Table D.6).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.172	0.170	−0.010	0.111
	(0.165)	(0.187)	(0.153)	(0.143)
Interrupted (T2)	0.024	0.008	−0.006	0.009
	(0.163)	(0.194)	(0.169)	(0.153)
Restarted (T3)	−0.067	−0.175	−0.004	−0.082
	(0.173)	(0.190)	(0.194)	(0.159)
All treatments (T1&T2&T3)	0.041	−0.001	−0.007	0.011
	(0.130)	(0.156)	(0.132)	(0.120)
Interrupted & Restarted (T2&T3)	−0.021	−0.082	−0.005	−0.036
	(0.141)	(0.166)	(0.148)	(0.131)
Predicted Probability Control (T0)	0.782***	0.917***	0.741***	0.813***
	(0.108)	(0.137)	(0.107)	(0.068)
Overall R^2	0.557			
Intraclass correlation	0.383			
Observations	3592			

Notes: The table reports a random effects panel regression with dependent variable how many bags a household recycled per period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later. The regression includes a control for households' baseline recycling activity (mean of recycled bags), house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.9
Treatment effects on the number of bags recycled by households per period–baseline recycling yes (panel regression version of Table D.7).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	−0.091	−0.249	−0.990	−0.444
	(0.688)	(0.802)	(0.742)	(0.627)
Interrupted (T2)	0.161	−0.360	−0.824	−0.341
	(0.780)	(0.879)	(0.767)	(0.712)
Restarted (T3)	−0.676	−0.972	−0.663	−0.771
	(0.785)	(0.871)	(0.935)	(0.745)
All treatments (T1&T2&T3)	−0.196	−0.535	−0.818	−0.516
	(0.609)	(0.717)	(0.646)	(0.571)
Interrupted & Restarted (T2&T3)	−0.240	−0.653	−0.747	−0.547
	(0.663)	(0.765)	(0.709)	(0.621)
Predicted Probability Control (T0)	3.310***	3.810***	3.524***	3.548***
	(0.512)	(0.614)	(0.529)	(0.319)
Overall R^2	0.425			
Intraclass correlation	0.383			
Observations	688			

Notes: The table reports a random effects panel regression with dependent variable being how many bags a household recycled per period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later; only households that recycled during the baseline period. The regression includes a control for households' baseline recycling activity (mean of recycled bags), house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.10
Treatment effects on the number of bags recycled by households per period-baseline recycling no (panel regression version of Table D.8).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.220*	0.262*	0.196**	0.226**
	(0.128)	(0.139)	(0.080)	(0.099)
Interrupted (T2)	−0.008	0.101	0.202**	0.099
	(0.069)	(0.110)	(0.089)	(0.075)
Restarted (T3)	0.079	0.017	0.155*	0.084
	(0.100)	(0.109)	(0.088)	(0.082)
All treatments (T1&T2&T3)	0.096	0.126	0.185***	0.135**
	(0.071)	(0.092)	(0.055)	(0.060)
Interrupted & Restarted (T2&T3)	0.035	0.059	0.179***	0.091
	(0.070)	(0.094)	(0.066)	(0.063)
Predicted Probability Control (T0)	0.172***	0.218***	0.069**	0.153***
	(0.048)	(0.071)	(0.030)	(0.030)
Overall R^2	0.036			
Intraclass correlation	0.370			
Observations	2904			

Notes: The table reports a random effects panel regression with dependent variable being how many bags a household recycled per period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Only households included that signed up in 2018 or later; only households that did not recycle during the baseline period. The regression includes a control for the house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.11
Treatment effects on whether households recycled at least once, by period-full sample (panel regression version of Table E.1).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.049*	0.046*	0.043	0.046*
	(0.029)	(0.028)	(0.027)	(0.024)
Interrupted (T2)	0.034	0.029	0.049*	0.037
	(0.028)	(0.028)	(0.027)	(0.024)
Restarted (T3)	0.029	−0.009	0.014	0.011
	(0.027)	(0.027)	(0.026)	(0.023)
All treatments (T1&T2&T3)	0.037	0.022	0.035*	0.032
	(0.023)	(0.023)	(0.021)	(0.020)
Interrupted & Restarted (T2&T3)	0.032	0.010	0.032	0.024
	(0.024)	(0.024)	(0.023)	(0.021)
Predicted Probability Control (T0)	0.164***	0.175***	0.138***	0.159***
	(0.018)	(0.018)	(0.016)	(0.010)
Overall R^2	0.460			
Intraclass correlation	0.287			
Observations	5568			

Notes: The table reports a random effects panel regression with dependent variable equal to one if a household recycled at least once during the respective period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Full sample included. The regression includes a control for households' binary baseline recycling activity, house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.12
Treatment effects on whether households recycle at least once, by period-full sample (panel logit regression version of Table E.2).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.867*	0.792	0.836	0.832*
	(0.510)	(0.484)	(0.510)	(0.432)
Interrupted (T2)	0.618	0.501	0.938*	0.686
	(0.495)	(0.500)	(0.503)	(0.433)
Restarted (T3)	0.536	−0.186	0.240	0.197
	(0.497)	(0.515)	(0.519)	(0.436)
All treatments (T1&T2&T3)	0.674	0.369	0.671	0.571
	(0.422)	(0.416)	(0.414)	(0.360)
Interrupted & Restarted (T2&T3)	0.577	0.158	0.589	0.441
	(0.440)	(0.442)	(0.440)	(0.380)
Predicted Probability Control (T0)	0.163***	0.173***	0.139***	0.158***
	(0.017)	(0.017)	(0.014)	(0.009)
Log pseudolikelihood	−1403.081			
Intraclass correlation	0.585			
Observations	5568			

Notes: The table reports a random-effects panel logit regression with dependent variable equal to one if a household recycled at least once during the respective period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Full sample included. The regression includes a control for households' binary baseline recycling activity, house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.13
Treatment effects on whether households recycled at least once, by period-baseline recycling yes-full sample (panel regression version of Table E.3).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.122	0.158*	0.055	0.112
	(0.088)	(0.085)	(0.090)	(0.072)
Interrupted (T2)	0.168**	0.105	0.091	0.121*
	(0.085)	(0.085)	(0.087)	(0.070)
Restarted (T3)	0.182**	0.052	0.035	0.090
	(0.085)	(0.088)	(0.090)	(0.071)
All treatments (T1&T2&T3)	0.158**	0.105	0.061	0.108*
	(0.072)	(0.071)	(0.073)	(0.059)
Interrupted & Restarted (T2&T3)	0.175**	0.079	0.063	0.106*
	(0.075)	(0.076)	(0.077)	(0.062)
Predicted Probability Control (T0)	0.556***	0.587***	0.556***	0.566***
	(0.063)	(0.061)	(0.061)	(0.036)
Overall R^2	0.147			
Intraclass correlation	0.317			
Observations	992			

Notes: The table reports a random effects panel regression with dependent variable equal to one if a household recycled at least once during the respective period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Full sample included; only households that recycled during the baseline period. The regression includes a control for households' binary baseline recycling activity, house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table F.14
Treatment effects on whether households recycled at least once, by period-baseline recycling no-full sample (panel regression version of Table E.4).

	(1)			(2)
	Period 2	Period 3	Period 4	Periods 2–4
Continuous (T1)	0.027	0.016	0.034*	0.026
	(0.024)	(0.024)	(0.020)	(0.019)
Interrupted (T2)	0.008	0.015	0.043**	0.022
	(0.023)	(0.024)	(0.021)	(0.019)
Restarted (T3)	−0.008	−0.025	0.007	−0.009
	(0.022)	(0.022)	(0.018)	(0.017)
All treatments (T1&T2&T3)	0.009	0.002	0.028*	0.013
	(0.019)	(0.019)	(0.015)	(0.015)
Interrupted & Restarted (T2&T3)	−0.000	−0.005	0.025	0.006
	(0.019)	(0.020)	(0.016)	(0.015)
Predicted Probability Control (T0)	0.077***	0.084***	0.046***	0.069***
	(0.016)	(0.016)	(0.012)	(0.009)
Overall R^2	0.035			
Intraclass correlation	0.334			
Observations	4576			

Notes: The table reports a random effects panel regression with dependent variable equal to one if a household recycled at least once during the respective period. Coefficients in column (1) are from the panel regression; column (2) is estimated via post-estimation Wald tests. Full sample included; only households that did not recycle during the baseline period. The regression includes a control for the house type, and enumerator fixed effects. Robust standard errors are in parentheses; standard errors are clustered at the household level. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Data availability

Data will be made available on request.

References

Aker, J. C., Sawyer, M., Goldstein, M., O'Sullivan, M., & McConnell, M. (2020). Just a bit of cushion: The role of a simple savings device in meeting planned and unplanned expenses in rural Niger. *World Development*, *128*, 104772.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9), 1082–1095. Special Issue: The Role of Firms in Tax Systems.

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, *104*(10), 3003–3037.

Altmann, S., & Traxler, C. (2014). Nudges at the dentist. *European Economic Review*, *72*, 19–38.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495.

Andor, M. A., Gerster, A., Peters, J., & Schmidt, C. M. (2020). Social norms and energy conservation beyond the us. *Journal of Environmental Economics and Management*, *103*, 102351.

Antinyan, A., Asatryan, Z., Dai, Z., & Wang, K. (2021). Does the frequency of reminders matter for their effectiveness? A randomized controlled trial. *Journal of Economic Behavior & Organization*, *191*, 752–764.

Apesteguia, J., Funk, P., & Iriberri, N. (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.

Ashraf, N., Karlan, D., & Yin, W. (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics*, *121*(2), 635–672.

Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural india: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ*, *340*, c2220.

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507.

Bonan, J., Cattaneo, C., d'Adda, G., Galliera, A., & Tavoni, M. (2025). Social norms and tariff salience: An experimental study on household waste management. *Journal of Environmental Economics and Management*, *130*, 103124.

Bonan, J., Cattaneo, C., d'Adda, G., & Tavoni, M. (2020). The interaction of descriptive and injunctive social norms in promoting energy conservation. *Nature Energy*, *5*, 900–909.

Borasino, E., & Fuhrmann-Riebel, H. (2022). New kids on the recycling block: The role of supermarkets and bodegas for sustainable consumer behaviour in Lima. *Circular Economy and Sustainability*, *2*, 857–881.

Byrne, D. P., Goette, L., Martin, L. A., Jones, A., Miles, A., Schob, S., Staake, T., & Tiefenbeck, V. (2024). *How nudges create habits: Theory and evidence from a field experiment*. Working paper, SSRN.

Cadena, X., & Schoar, A. (2011). Remembering to pay? Reminders VS. Financial incentives for loan payments. NBER Working Paper Series No. 17020.

Calzolari, G., & Nardotto, M. (2017). Effective reminders. *Management Science*, *63*(9), 2915–2932.

Carlsson, F., Gravert, C., Johansson-Stenman, O., & Kurz, V. (2021). The use of green nudges as an environmental policy instrument. *Review of Environmental Economics and Policy*, *15*(2), 216–237.

Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, *77*(3), 909–931.

Chong, A., Karlan, D., Shapiro, J., & Zinman, J. (2015). (ineffective) messages to encourage recycling: Evidence from a randomized evaluation in Peru. *The World Bank Economic Review*, *29*(1), 180–206.

Dai, H., Saccardo, S., Han, M. A., Roh, L., Raja, N., Vangala, S., Modi, H., Pandya, S., & Croymans, D. M. (2021). Behavioral nudges increase Covid-19 vaccinations. *Nature*, *579*, 404–409.

Damgaard, M. T., & Gravert, C. (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics*, *157*, 15–26.

Datta, S., & Mullainathan, S. (2014). Behavioral design: A new approach to development policy. *Review of Income and Wealth*, *60*(1), 7–35.

Dupas, P., & Robinson, J. (2013). Savings constraints and microenterprise development: Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics*, *5*(1), 163–192.

Dustan, A., Hernandez-Agramonte, J. M., & Maldonado, S. (2023). Motivating bureaucrats with behavioral insights when state capacity is weak: Evidence from large-scale field experiments in Peru. *Journal of Development Economics*, *160*, 102995.

Essl, A., Steffen, A., & Staehle, M. (2021). Choose to reuse! the effect of action-close reminders on pro-environmental behavior. *Journal of Environmental Economics and Management*, *110*, 102539.

Fang, X., Goette, L., Rockenbach, B., Sutter, M., Tiefenbeck, V., Schoeb, S., & Staake, T. (2023). Complementarities in behavioral interventions: Evidence from a field experiment on resource conservation. *Journal of Public Economics*, *228*, 105028.

Ferraro, P. J., & Price, M. K. (2013). Adaptive linear step-up procedures that control the false discovery rate. *The Review of Economics and Statistics*, *95*(1), 64–73.

Fuhrmann-Riebel, H., D'Exelle, B., Lopez Vargas, K., Tonke, S., & Verschoor, A. (2024). Correcting misperceptions about trends and norms to address weak collective action — experimental evidence from a recycling program. *Journal of Environmental Economics and Management*, *128*, 103046.

Gilbert, B., & Zivin, J. G. (2014). Dynamic salience with intermittent billing: Evidence from smart electricity meters. *Journal of Economic Behavior and Organization*, *107*, 176–190.

Gravert, C. (2022). Reminders as a tool for behavior change. A version of this article is in Mazar N. and Soman D. (in press), Behavioral Science in the Wild, Toronto, Canada: University of Toronto Press, Forthcoming Available at SSRN.

Gravert, C., & Olsson Collentine, L. (2021). When nudges aren't enough: Norms, incentives and habit formation in public transport usage. *Journal of Economic Behavior & Organization*, *190*, 1–14.

Grelle, S., Kuhn, S., Fuhrmann-Riebel, H., & Hofmann, W. (2024). The role of framing and effort in green nudging acceptance. *Behavioural Public Policy*, 1–16.

Grácio, M., & Vicente, P. C. (2021). Information, get-out-the-vote messages, and peer influence: Causal effects on political behavior in Mozambique. *Journal of Development Economics*, *151*, 102665.

He, G., Pan, Y., Park, A., Sawada, Y., & Tan, E. S. (2023). Reducing single-use cutlery with green nudges: Evidence from china's food-delivery industry. *Science*, *381*(6662), eadd9884.

Huck, S., & Rasul, I. (2010). Transactions costs in charitable giving: Evidence from two field experiments. *The B.E. Journal of Economic Analysis & Policy*, *10*(1).

Huffman, A. H., Van Der Werff, B. R., Henning, J. B., & Watrous-Rodriguez, K. (2014). When do recycling attitudes predict recycling? An investigation of self-reported versus observed behavior. *Journal of Environmental Psychology*, *38*, 262–270.

Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics*, *80*, 47–58.

Karlan, D., McConnell, M., Mullainathan, S., & Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, *62*(12), 3393–3411.

Kaza, S., Yao, L. C., Bhada-Tata, P., & Van Woerden, F. (2018). *What a waste 2.0: A global snapshot of solid waste management to 2050*. Washington, DC: World Bank. License: CC BY 3.0 IGO.

Kirakozian, A. (2016). One without the other? Behavioural and incentive policies for household waste management. *Journal of Economic Surveys*, *30*(3), 526–551.

Knickmeyer, D. (2020). Social factors influencing household waste separation: A literature review on good practices to improve the recycling performance of urban areas. *Journal of Cleaner Production*, *245*, 118605.

Kremer, M., Rao, G., & Schilbach, F. (2021). Behavioral development economics. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of behavioral economics: Applications and foundations 2* (pp. 345–458). Amsterdam: Elsevier.

Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed? Modelling habit formation in the real world. *European Journal of Social Psychology*, *40*(6), 998–1009.

Larochelle, C., Alwang, J., Travis, E., Barrera, V. H., & Andrade, J. M. D. (2019). Did you really get the message? Using text reminders to stimulate adoption of agricultural technologies. *The Journal of Development Studies*, *55*(4), 548–564.

Mertens, S., Herberz, M., Hahnel, U. J. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, *119*(1), e2107346118.

Miafodzyeva, S., & Brandt, N. (2013). Recycling behaviour among householders: Synthesizing determinants via a meta-analysis. *Waste and Biomass Valorization*, *4*(2), 221–235.

Milkman, K. L., Ellis, S. F., Gromet, D. M., DeMay, I. M., Graci, H. N., Jung, Y., Mobarak, R. S., Zumaran, R. A. S., Simmons, M. N., den Bulte, C. V., Benartzi, S., Hilchey, M., Goodyear, L., Karlan, D., Mazar, N., Mochon, D., Shah, A. M., Soman, D., Zinman, J., & Duckworth, A. L. (2025). Can reminder emails compel Americans to save? A two-million-person megastudy. *PNAS Nexus*, *4*(9), pgaf280.

Milkman, K. L., Gromet, D., Ho, H., Kay, J. S., Lee, T. W., Pandiloski, P., Park, Y., Rai, A., Bazerman, M., Beshears, J., Bonacorsi, L., Camerer, C., Chang, E., Chapman, G., Cialdini, R., Dai, H., Eskreis-Winkler, L., Fishbach, A.Gross, J. J. et al. Megastudies improve the impact of applied behavioural science. *Nature*, *600*(7889), 478–483.

Milkman, K. L., Patel, M. S., Gandhi, L., Graci, H. N., Gromet, D. M., Ho, H., Kay, J. S., Lee, T. W., Akinola, M., Beshears, J., Bogard, J. E., Buttenheim, A., Chabris, C. F., Chapman, G. B., Choi, J. J., Dai, H., Fox, C. R., Goren, A.Hilchey, M. D. et al. A megastudy of text-based nudges encouraging patients to get vaccinated at an upcoming doctor's appointment. *Proceedings of the National Academy of Sciences*, *118*(20), e2101165118.

Muller, P., & Habla, W. (2018). Experimental and non-experimental evidence on limited attention and present bias at the gym. ZEW Discussion Paper No. 18-041.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F. K., Wilson, D. A., Wu, C.-F., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, *92*(2), 135–138. Special Issue: Neurobiology of Habituation.

Rockenbach, B., Tonke, S., & Weiss, A. R. (2025). A large-scale field experiment to reduce nonpayments for water: From diagnosis to treatment. *The Review of Economics and Statistics*, *107*(5), 1233–1246.

Rodríguez, C., & Saavedra, J. E. (2019). The persistent effects of youth savings reminders: Experimental evidence from text-message campaigns in Colombia. *Journal of Development Economics*, *139*, 135–156.

Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics*, *7*(3), 51–84.

Sonntag, A., & Zizzo, D. J. (2015). On reminder effects, drop-outs and dominance: Evidence from an online experiment on charitable giving. *PLOS ONE*, *10*(8), 1–17.

Thomas, C., & Sharp, V. (2013). Understanding the normalisation of recycling behaviour and its implications for other pro-environmental behaviours: A review of social norms and recycling. *Resources, Conservation and Recycling*, *79*, 11–20. SI: Resourceful Behaviours.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., & Staake, T. (2018). Overcoming salience bias: How real-time feedback fosters resource conservation. *Management Science*, *64*(3), 1458–1476.

Tiefenbeck, V., Wörner, A., Schöp, S., Fleisch, E., & Staake, T. (2019). Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives. *Nature Energy*, *4*(1), 35–41.

Tonke, S. (2024). Providing procedural knowledge: A field experiment to encourage resource conservation in Namibia. *Journal of Development Economics*, *166*, 103202.

Tonke, S. (2025). Shaping identity: Evidence from a large-scale field experiment. *Journal of Political Economy Microeconomics*, *forthcoming*.

UNEP. (2024). *Global waste management outlook 2024: Beyond an age of waste – turning rubbish into a resource*. Technical report, United Nations Environment Programme.

Varotto, A., & Spagnolli, A. (2017). Psychological strategies to promote household recycling. A systematic review with meta-analysis of validated field interventions. *Journal of Environmental Psychology*, *51*, 168–188.

Wallander, S., Ferraro, P., & Higgins, N. (2017). Addressing participant inattention in federal programs: A field experiment with the conservation reserve program. *American Journal of Agricultural Economics*, *99*(4), 914–931.

Wood, W., & Neal, D. T. (2007). A new look at habits and the habit–goal interface. *Psychological Review*, *114*(4), 843–863.

WWF. 2018. Almost 90% of the garbage generated daily is not recycled. https://www.wwf.org.pe/en/?328834/Almost-90-of-the-garbage-generated-daily-is-not-recycled.