I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17013

Can Reminders Promote Regular Pro-Environmental Behavior? Experimental Evidence from Peru

Hanna Fuhrmann-Riebel
Ben D'Exelle
Kristian López Vargas
Sebastian Tonke
Arjan Verschoor

MAY 2024

DISCUSSION PAPER SERIES

IZA DP No. 17013

Can Reminders Promote Regular Pro-Environmental Behavior? Experimental Evidence from Peru

Hanna Fuhrmann-Riebel
University of East Anglia and German Institute of Development and Sustainability

Ben D'Exelle
University of East Anglia and IZA

Kristian López Vargas
University of California, Santa Cruz

Sebastian Tonke
Max Planck Institute for Research on Collective Goods

Arjan Verschoor
University of East Anglia and Norwich Institute for Sustainable Development

MAY 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Can Reminders Promote Regular Pro-Environmental Behavior? Experimental Evidence from Peru*

Tackling environmental pollution requires a permanent change in regular, repeated behavior of households. Bringing about change in such behavior may require interventions that are not limited to a single point in time, yet little evidence exists on how frequently we need to target households to initiate behavioral change and to form new habits in regular proenvironmental behavior. To fill this gap, we investigate the impact of mobile text reminders on households' recycling behavior in urban Peru, by randomly varying the frequency of reminders over a nine-week treatment period. We find that reminders increase both the likelihood that households start to recycle, and the frequency of recycling among households that already recycled before the intervention. The effects are stronger if reminders are repeated over a longer period. Our findings suggest that low-cost mobile text reminders can support repeated pro-environmental behavior, and that some repetition may be needed to maximize their effectiveness.

JEL Classification: C93, D83, D90, D91, Q53, Q58
Keywords: reminders, limited attention, habit formation, recycling, Peru

Corresponding author:
Hanna Fuhrmann-Riebel
University of East Anglia
Research Park
Norwich NR4 7TJ
Great Britain
E-mail: hanna.fuhrmann-riebel@idos-research.de

* We thank Erwin Bulte, Ximeng Fang, Bereket Kebede, as well as participants of the IDOS-Uni Bonn Behavioral Economics Exchange Group, 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 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 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 pre-registered 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 Research on Collective Goods

1. Introduction

Environmental pollution, as a result of increasing global waste accumulation, poses a serious threat to ecosystems, biodiversity and human health, and contributes significantly to global climate change (UNEP, 2015).[1] Addressing environmental pollution requires a permanent shift in households' regular and repeated behavior. A good example is recycling behavior, which requires repeated action while significantly contributing to environmental protection. Weekly kerbside collection of recyclables, which is 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, however, remains limited in many low- and middle-income countries, where recycling habits have not established yet. To understand why that may be the case and what can be done to increase recycling, we distinguish two important behavioral factors. First, people may fail to follow through with their recycling intentions because of "limited attention" (Datta and Mullainathan, 2014; Karlan et al., 2016; Tiefenbeck et al., 2018). Limited cognitive capacity makes that one can only pay limited attention to recycling, which requires substantial organization (such as separating recyclables, putting them in a bag, putting the bag on the street on the correct day, etc.). Second, once people have recycled a couple of times, they may form a routine in recycling practices, and the barrier of limited attention may disappear.

The issue of 'limited attention' can be addressed by sending timely reminders that bring the specific behavior "to the top of mind" (Karlan et al., 2016).[2] However, it is not clear how effective reminders are when aiming to change *repeated* behavior. We may need to repeat reminders to overcome limited attention, yet the effectiveness of reminders may also decrease with repetition due to 'habituation', i.e., we may experience a weakened behavioral response due to repeated treatment. If that is the case, one may interrupt the series of reminders, which may recover the behavioral response to the treatment (Rankin et al., 2009). In addition, reminders may need to be repeated for habits to be formed, which – once established – will reduce the need for reminders. However, here again it

[1]As a result, sustainable waste and resource management have become a core interest of both researchers and policy makers over the last years (see, e.g., Miafodzyeva and Brandt, 2013; Thomas and Sharp, 2013; Kirakozian, 2016; Varotto and Spagnolli, 2017; Knickmeyer, 2020, for literature reviews and meta-analyses)

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

is *a priori* not clear how frequently people need to be treated for habits to be formed (Charness and Gneezy, 2009; Allcott and Rogers, 2014; Royer, Stehr and Sydnor, 2015). This leads to two important questions: 1) whether reminders can increase recycling, and 2) which frequency of reminders increases 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 of reminders.

To answer these questions, we conducted an 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 to 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 frequency to test their effect on recycling.

We collected data on households' recycling behavior over a period of 12 weeks, by accompanying the recycling trucks that collect the recycling bags on their daily routes. Having the addresses of registered households allows us to identify which households put a recycling bag on the street in which week, without needing to contact them. This 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 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 continuous reminders 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, reminders are generally effective at increasing recycling behavior: weekly reminders in the first three weeks of the intervention period increase the likelihood of recycling in all treatment groups compared to the control group. This initial effect is mainly driven by households that already recycled at baseline. Second, the reminder effect becomes stronger if the reminders are continued for another three weeks. Third, households that did not yet recycle at baseline 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 effects of reminders.

We contribute to the existing literature in the following ways. First, there is a growing yet still limited literature on the effectiveness of reminders in the context of regular, repeated behaviors (e.g., regular gym attendance (Calzolari and Nardotto, 2017; Muller and Habla, 2018) or energy conservation (Allcott and Rogers, 2014)) as opposed to infrequent or one-time decision-making. This literature has focused on 1) habit formation, which might mediate the effect of reminders on regular behavior, and 2) whether the frequency of reminders matters.[3] We contribute to this literature by generating new evidence on the role of the frequency with which reminders are used, by combining treatments that continue, interrupt or restart reminders. In contrast to previous studies, our design allows us to compare the 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.[4]

Second, evidence on the effectiveness of reminders on household recycling is scarce. Essl, Steffen and Staehle (2021) found a positive effect of weekly reminders in the form of stickers and flyers on the reuse of plastic bags in the context of weekly food delivery boxes in Switzerland. Chong et al. (2015) also focused on the effect of reminders on recycling in a recycling program in Peru. The authors did not find any significant effect of mobile text reminders on households' recycling activity, which was probably due to the lack of recycling bins, which they found was the main obstacle for households to recycle. Our main contribution is that we do find that reminders can significantly increase households' recycling activity, where the necessary infrastructure is in place.[5]

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

2. Design

2.1. 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 8,468 tons of garbage are generated daily in Lima, of which only 4% is recycled (WWF, 2018). Municipalities

[3]Some studies also generated evidence on the effectiveness of repeated reminders on one-time or infrequent decisions (Altmann and Traxler, 2014; Sonntag and Zizzo, 2015; Damgaard and Gravert, 2018).

[4]Allcott and Rogers (2014) compare persistence effects of sending reminders to a group that continues to receive reminders during that same time period as well, but without having a comparison group that was first interrupted and then restarted.

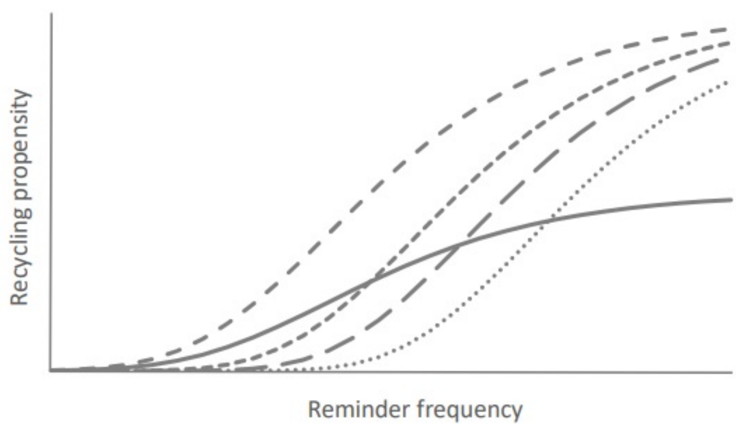
[5]By doing so, we also add to the growing literature on "green nudges" (Carlsson et al., 2021), which – given their high economic relevance – offer a promising tool for policy makers to address societal challenges (Grelle et al., 2024; He et al., 2023).

are responsible for coordinating recycling activities at the household level in Peru.[6] Some municipalities in Lima have established their own recycling program, 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, which we collaborated with for this study. Households that participate 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 recycling bags are collected once or twice a week (depending on the zones), always on the same day(s) at approximately the same time. Thus, paying attention to this specific day and time to place the recycling bags on the street is crucial for people to participate successfully.

2.2. Conceptual framework

Before presenting our experiment, it is important to conceptualize how we understand the role of reminders for recycling behavior. For this, we make the following assumptions. First, we assume that each respondent wants to recycle, since they voluntarily signed up to the recycling program. Second, they have limited cognitive capacity and can therefore pay only limited attention to recycling. As recycling requires substantial organisation to undertake the necessary tasks (such as separating recyclables, putting them in a bag, putting the bag on the street on the correct day, etc.), many fail to do so. Third, we expect that if we remind them on the recycling day, they will dedicate more attention to it, and this may be enough for them to perform the necessary recycling tasks. Fourth, once they have recycled a couple of times, they will form a routine, and the need for reminders decreases.

Figure 1: Recycling propensity by reminder frequency



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

Using these assumptions, we can plot the effect of the frequency of reminders on recycling propensity, as shown by Figure 1. The graph represents a sigmoid function, which has a convex and a concave part. The first part is convex, which implies that the effect of an additional reminder increases with the number of reminders. Here we need to overcome respondents' 'limited attention', by sending a sufficiently high number of reminders for an observable effect to occur. The second part is concave, which implies decreasing marginal returns due to 'habituation', i.e., a weakened behavioral response due to repetition (Rankin et al., 2009).

The different paths present different possibilities of how quickly one responds to reminders, how quickly habituation kicks in, and what recycling propensity one can reach. The recycling propensity relates to habit formation. It is possible that with sufficient reminders we reach a sufficiently high recycling propensity (or frequency) for recycling habits to be formed, so that no further reminders are needed. Equally plausible is that reminders are unable to increase recycling enough for habits to be formed (see the full line in Figure 1).

Given the different trajectories, several questions can be asked. First, it is important to ask *whether reminders are (at all) effective to encourage recycling behaviour*. It might be that reminders do not overcome the barrier of limited attention, and would therefore be unable to increase recycling. This is our first question. In addition, we do not know the optimal frequency of the reminders, which leads to our second question: *"What frequency of reminders supports repeated recycling most effectively?"*. This question consists of two sub-questions.

First, it might be that only a few reminders are needed for respondents to start recycling (this is the case for the first dashed line), or that a longer series of reminders is required (e.g., the dotted line). In addition, we do not know how quickly 'habituation' makes reminders less effective. Therefore, question 2a would be: *"Is a longer series of reminders more effective at maximizing recycling propensity than a shorter series of reminders?"* Second, reminders could be more effective after interrupting and restarting them. While there is a general tendency of decreasing effectiveness of reminders due to 'habituation', withholding the reminders could make the response to reminders recover over time. Such spontaneous recovery of the 'response to stimulus' has been described by Rankin et al. (2009). Interrupting and restarting reminders might better capture respondents' attention, which will increase the reminders' effectiveness. Graphically, this would make the relevant section of the graph less concave, which could help reach a higher recycling propensity. This leads to question 2b: *"Is a long series of reminders more effective at maximizing recycling propensity than an interrupted and restarted series of reminders?"*.

2.3. Treatments

To address our research questions, we use four treatments randomly assigned to individual respondents in a between-subject design. The four groups consist of one control group and three treatment groups (see Table 1). All groups do not receive any reminders in the baseline period. From period 2 onwards, we vary the frequency of reminders creating four groups as follows:

- Control group (T0): no reminders
- Continuous treatment group (T1): reminders over the whole nine weeks
- Interrupted treatment group (T2): reminders only for the first three weeks
- Restarted treatment group (T3): reminders for the first three weeks and for the last three weeks, with a three weeks' pause in between

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.

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 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 any 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 did in 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.[7]

The treatment message of the mobile text reminder contained a friendly greeting from the municipality and a simple reminder for people to put their recycling bags outside on that day. We also gave people a number they could call in case of doubts and referred 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.[8] 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].

2.4. 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.[9] We start with a comparison of recycling behavior in period 2. In this period, all treated households receive a weekly reminder. We expect that this will address people's limited attention about 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 habit formation may induce some households in both groups to 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

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

[8] Research has shown that varying the content of the reminder messages tends to have little or no effect (Apesteguia, Funk and Iriberri, 2013; Altmann and Traxler, 2014; Wallander, Ferraro and Higgins, 2017; Larochelle et al., 2019). 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.

[9] 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).

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 T2, T3 and T4 (thanks to recycling habits in T2, supported by the reinforcement of another series of reminders in T3 and T4) relative to the control group. As recycling habits weaken over time, recycling will be higher in the continuous treatment (T1) than in the interrupted treatment (T2). In addition, 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 in the long run. 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)*

2.5. Implementation

Sample

There are 7,183 households officially registered in the recycling program (end of March 2021). 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 7,183 households registered in the recycling program, 3,480 had a unique address.

When households register for the recycling program, they are asked to give a phone number. Households can decide whether they 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 of which mobile phone numbers are available. This is the case for 1,392 households of the 3,480 households that are registered with a unique address. This is the sample that we will use in our study.

Data collection

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.

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.[10] The recording of households' recycling behavior was done through audio recordings via headsets. 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. 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 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. In zones with two collection days per week, the reminders were sent in the morning of the first collection day of the week.

[10] 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).

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 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 already recycled during the baseline period. Balance tests confirm that these household characteristics are balanced across treatment groups.

Since sign-up years go back to 2015, we decided to focus in 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, this could be the case if people moved away from the district and are thus unable to participate in the recycling program. Second, people might have changed their phone numbers, and therefore be unable to receive our mobile text reminders. Ideally, we would have been able to remove those respondents from our data base. However, as the municipality does not update their data base 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 proportion of inactive respondents tends to be 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 1,392 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.[11]

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, where no recycling service was in place and thus no reminders were sent and no data could be collected. Therefore, the whole 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.

[11]To test the robustness of our results, we will also report all results using the full sample (see Appendix E).

3.2. Treatment effects by period

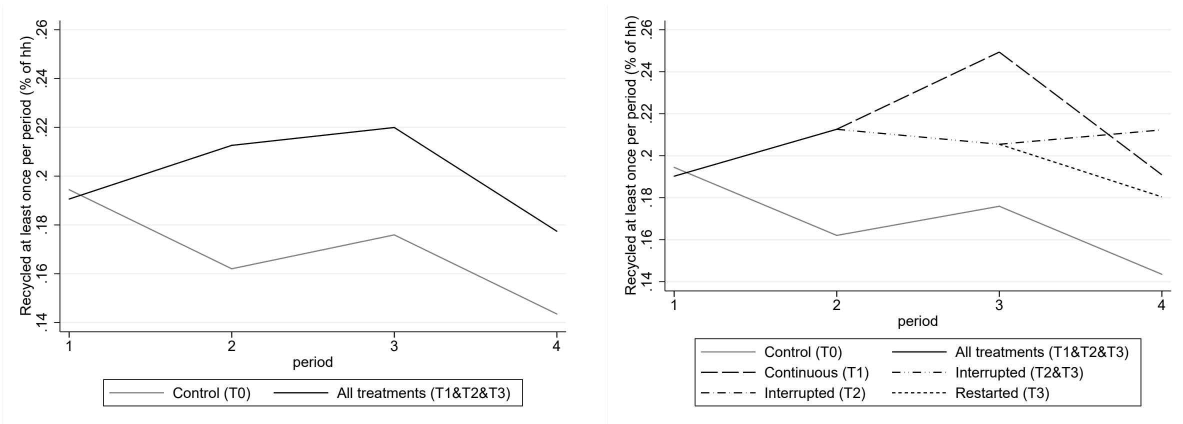
The main outcome variable of our analysis is a binary variable equal to one if a household recycled at least once per period of three weeks, zero otherwise. The periods of three weeks are the blocks along which we designed the different treatments and pre-registered our hypotheses. We focus on recycling behavior per period rather than week because of the irregular nature of households' recycling activity, as explained in Section 2.3. Histograms on households' recycling frequency in Appendix B confirm that hardly any household recycled every week in the baseline period.

To start our analysis, we analyse how treatment differences vary over the individual periods 2, 3 and 4. Figure 2 plots the proportion of households that recycled at least once in the respective period, by period and treatment. In Figure 2a, we observe that recycling levels increase after reminders are sent in period 2 in the treatment groups compared to the control group. Figure 2b shows that repeated reminders in the continuous treatment push recycling even higher in period 3. We also observe that recycling levels in all treatment groups are higher than in the control group throughout the whole intervention period. In addition, there is a general drop in recycling levels in period 4, which is caused by the public holiday and related travelling of people in week 11, as supported by the weekly recycling data presented in Figure C.1 in Appendix C.

Figure 2: Proportion of households that recycled at least once per period, by treatment

(a) All treatments vs. control

(b) Treatment split



Notes: The figure shows the proportions 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 ($\geq$ 2018). Sub-figure 2a compares all treatment groups pooled together with the control group; sub-figure 2b 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. For visualization purposes of the pooled treatments, the data for each treatment is first demeaned and then centered around the pooled groups mean.

To test whether these differences are statistically significant, we estimate a regression of the following form:

$$Y_i = ß_0 + ß_1 T_i + ß_2 X_i + u_i \tag{1}$$

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 multiple dwelling unit or unknown house type), and indicator variables for the different enumerators. We estimate regressions with OLS.[12]

We first look at the effect of all treatment groups pooled together over the whole intervention period. Figure 2a shows that recycling activity throughout the whole intervention period is higher among households that received weekly reminders than among households in the control group. The regression results reported in columns (1) of Table 2 confirm that this difference is statistically significant.

We next look at each of the intervention periods separately. Figure 2a shows that the reminders start working from the first period reminders are sent, i.e., period 2. Column (2) of Table 2 shows that the percentage of households that recycled at least once in period 2 is about 5 percentage points higher than in the control group, and that this difference is statistically significant at the 5% significance level. This finding provides support for hypothesis H1.

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). 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.

Figure 2b shows that the percentage of households that recycled at least once in period 3 is highest in the continuous treatment group. It also shows that the recycling activity in period 3 is higher in both the continuous and the interrupted treatment groups than in the control group, and that it is higher in the continuous treatment group than in the interrupted ones.

Column (3) in Table 2 confirms that in period 3, the percentage of households that recycled is 8.3 percentage points higher in the continuous treatment group than in the control group, which is significant at the 1% level. While the coefficient of T2&T3 is positive in column (3) as well, suggesting that households in the interrupted treatments recycled more than in the control group too, we find that this difference is not statistically significant. Moreover, when comparing treatment coefficients, we find that the coefficient of T1 is significantly larger than the coefficient of T2&T3 (p-value of a one-sided wald

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

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**	0.054**		
	(0.019)	(0.026)		
Continuous (T1)			0.083***	
			(0.030)	
Interrupted (T2&T3)			0.034	
			(0.025)	
Continuous (T1)				0.042
				(0.028)
Interrupted (T2)				0.049*
				(0.028)
Restarted (T3)				0.019
				(0.027)
Predicted Probability Control (T0)	0.158***	0.160***	0.172***	0.141***
	(0.017)	(0.023)	(0.020)	(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 in parentheses; standard errors clustered at the household level in column (1). ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Logit regressions are reported in Table D.1.

test = 0.033). This indicates that when only looking at period 3, the recycling activity of households in the continuous treatment group is significantly higher than among households in the interrupted groups.

In sum, we provide partial support for hypothesis H2. We observe a continuous reminder effect, as recycling is higher in the continuous treatment relative to the control group. While we also find some suggestive evidence for persistence effects, the difference in recycling levels between the interrupted treatments and the control group is not statistically significant. Furthermore, our analysis confirms that recycling in the continuous treatment group is significantly higher than in the interrupted treatment groups.

As a final step, we turn to period 4 where the treatment groups are separated into the continuous, interrupted and restarted treatment group. 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 we know that many people used this public holiday to travel, which might have led to even lower recycling rates during the whole week. Figure 2 shows that the recycling activity drops in period 4 compared to previous periods, and even falls below levels in the baseline period in case of the control group. Figure C.1 in Appendix C confirms that this is mainly driven by low recycling levels in week 11.

Despite the general drop in recycling activity in period 4, we can see in Figure 2b that the percentages of households that recycled remain higher in all treatment groups than in the control group. The regressions reported in column (4) of Table 2 confirm that the treatment coefficients are positive, but not statistically significant. The only exception is the interrupted treatment group, for which the difference is marginally significant at the 10% level. Looking at differences in recycling levels between the different treatment groups, we cannot identify any significant differences. Using a wald-test that compares the treatment coefficients, none of the differences are statistically significant (p-values for one-sided alternative hypotheses are: 0.414 for T2 > T1, 0.219 for T1 > T3, and 0.156 for T2 > T3). In sum, we do not find evidence in support of hypothesis H3.

3.3. Heterogeneity by recycling at baseline

The histogram presented in Figure 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 recycle yet at baseline might need more reminders to start recycling than households that already recycled recently.

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	Period 2	Period 3	Period 4	Period 2-4	Period 2	Period 3	Period 4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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^{**}				0.058^{**}	
			(0.095)				(0.029)	
Interrupted (T2&T3)			0.051				0.024	
			(0.089)				(0.023)	
Continuous (T1)				0.000				0.050^{**}
				(0.109)				(0.024)
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 in parentheses; standard errors clustered at individual household level in columns (1) and (5). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.036 for T1 > T2&T3; column (4) p=0.370 for T2 > T1, p=0.420 for T1 > T3, p=0.290 for T2 > T3; column (7) p=0.105 for T1 > T2&T3; column (8) p=0.469 for T1 > T2, p=0.155 for T1 > T3, p=0.169 for T2 > T3.. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

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 observe 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, we observe that their difference is substantial and statistically significant in the left panel, but not in the right panel.

In period 4, we observe significant positive effects of the continuous and the interrupted treatments among households that did not recycle at baseline, but not among households that already recycled before. The lack of significant effects among the respondents who already recycled in the baseline period might be partially due to the lower sample size.

3.4. Mechanisms

Tracking individual recycling behavior over two consecutive periods allows us to obtain insights into the importance of 'limited attention' and 'habit formation', which are the key behavioral mechanisms included in our conceptual framework. First, evidence in support of 'limited attention' would be provided if we observed that respondents who did not recycle in period 2 are more likely to start recycling in period 3 if they continue receiving reminders relative to the interrupted treatments. This would suggest that they need a sufficiently long series of reminders to overcome limited attention. Second, for evidence in support of 'habit formation', we could look for the following patterns. If we observed that respondents who recycled in period 2 are more likely to continue recycling in period 3 in the interrupted treatments relative to the control group, this would suggest some habit formation. If, in contrast, respondents who recycled in period 2 are less likely to continue recycling in period 3 if they stop receiving reminders, this would suggest that some respondents have not formed a firm recycling habit yet. Note that both options are not mutually exclusive, as we might have both types of respondents in the sample.

Table 4 shows the marginal probabilities of the continuous treatment and the interrupted treatments relative to the control group, on the following four discrete outcomes: the respondent (1) did not recycle at all in any of the two periods, (2) recycled only in period 3, (3) recycled only in period 2, (4) recycled at least once in both periods 2 and 3.

In column 1, we observe that both treatments reduce the likelihood that a respondent does not recycle in either period. The continuous treatment T1 has a stronger effect than the interrupted treatments, as confirmed by the wald test results reported in the notes under the table. A comparison between Panels B and C shows that both effects (as well as the difference between both treatments) is observed irrespective of whether respondents recycled in the baseline period or not. We also see that the effects are sizeable and stronger in the subsample of respondents who recycled at baseline. For example, among respondents who recycled in the baseline period, the continuous treatment T1 reduces the likelihood of no recycling by 23 percentage points, starting with an already low probability of 30.8%.

A decrease in the outcome 'neither p2 nor p3' should lead to an increase in one or more of the other outcomes, where respondents recycle in period 2, period 3 or both. An analysis of which outcomes become more likely across treatments can provide us insights in the behavioral mechanisms that influence recycling. Column 2 reports the effects on the likelihood that respondents only recycle in period 3 and not in period 2. We find that the continuous treatment has a stronger (positive) effect on this likelihood than the interrupted treatment (p-value of a one-sided wald test is 0.027). This suggests that some respondents need a sufficiently long series of reminders to start recycling, which provides support for the 'limited attention' mechanism.

Table 4: Treatment effects on recycling in period 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 (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 (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 (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 in parentheses. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Comparing the marginal probabilities of T1 and T2&T3 using a one-sided wald test gives the following p-values:
Panel A: column (1) p=0.023; column (2) p=0.027; column (3) p=0.202; column (4) p=0.281;
Panel B: column (1) p=0.084; column (2) p=0.068; column (3) p=0.140; column (4) p=0.375;
Panel C: column (1) p=0.030; column (2) p=0.126; column (3) p=0.068; column (4) p=0.335.

Column 3 reports the effects on the likelihood that respondents only recycle in period

2 and not in period 3. Panel B shows that the interrupted treatments increase this likelihood among respondents who recycled at baseline by 11 percentage points. This effect is sizeable, given a low 2.4 predicted probability in the control group. This effect is substantially stronger than in the continuous treatment, where the increase is only 4.9 percentage points. The lack of statistical significance of the difference between both effects is most likely due to the small size of this subsample. This provides suggestive evidence that some respondents have not established a firm recycling habit yet, when we stop sending reminders in period 3.

In column 4, we observe that respondents who recycled at baseline have a roughly 50% probability to recycle in periods 2 and 3, irrespective of whether we send reminders (Panel B). This suggests that these respondents have established firm recycling habits. In addition, the coefficients of both treatments are positive in all panels, and sizeable in Panel B. The coefficients, however, are not statistically significant, probably due to the small sample size. Note that we do not observe this pattern among respondents who did not recycle at baseline (Panel C), where the coefficients are extremely small.

3.5. Treatment effects over the entire intervention period

Next, we look at treatment effects on households' recycling behavior aggregated over the whole intervention period (periods 2-4).[13] 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.

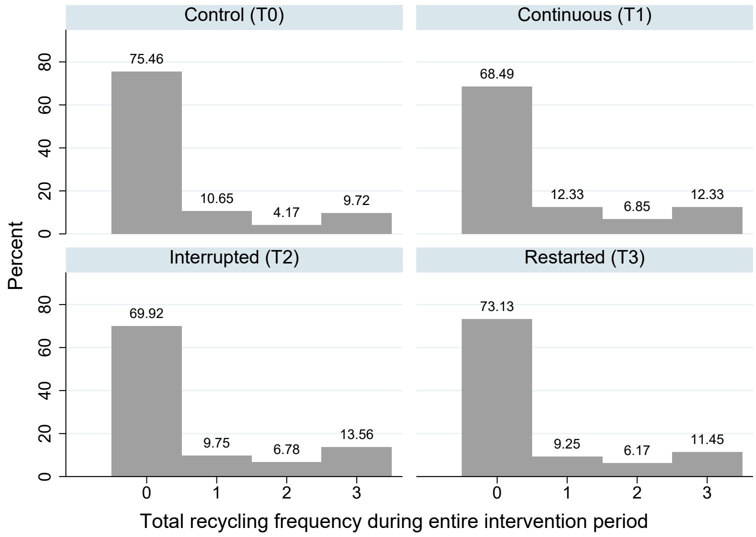
Figure 3 shows the distribution of the total recycling frequency over the whole intervention period, by treatment. Specifically, for each period we count whether a household recycled at least once, and we take the sum of this count across the three intervention periods (periods 2, 3, and 4). This means 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 Figure 3a, we observe a shift to the right (i.e., towards 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 Figures 3b and 3b, we observe that this holds irrespective of whether respondents already recycled before the intervention. Interestingly, while among respondents who did not recycle yet at baseline, the reminders are mostly successful at making respondents recycle in at least one of the subsequent three periods, among respondents who already recycled at baseline, the reminders increase the proportion of respondents who recycle two or three periods.

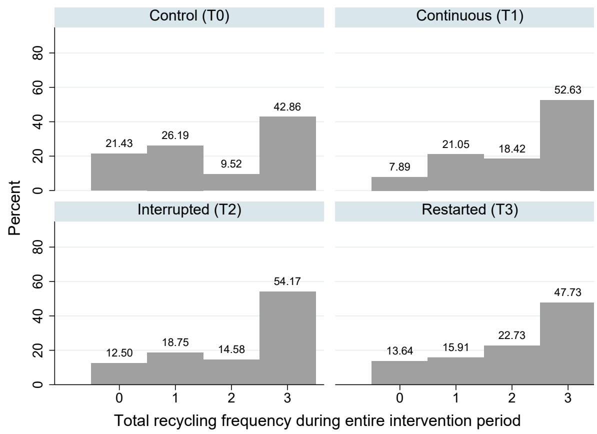
[13]Note that this analysis is exploratory and was not pre-registered.

Figure 3: Total recycling frequency during entire intervention period, by treatment

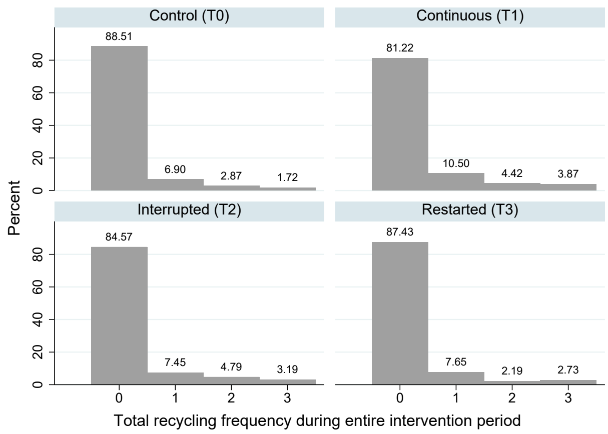
(a) Pooled sample



(b) Baseline recycling yes



(c) Baseline recycling no



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 ($\geq$ 2018). Sub-figure 3a for pooled sample: N(T0) = 216, N(T1) = 219, N(T2) = 236, N(T3) = 227; sub-figure 3b for households that recycled during the baseline period: N(T0) = 42, N(T1) = 38, N(T2) = 48, N(T3) = 44; sub-figure 3c for households that did not recycle during the baseline period: N(T0) = 174, N(T1) = 181, N(T2) = 188, N(T3) = 183.

To test whether treatment differences are statistically significant, we use an ordered probit regression, combining the four frequency categories as dependent variable. As right-hand side, we use the same variables used in previous regressions. Table 5 presents 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 sizeable. Comparing the treatment effects, we observe that the coefficients of

T1 are larger in absolute terms relative to treatments T2 and T3. Using a one-sided wald test that compares these coefficients, we can only reject the null of no difference between treatments T1 and T3 at the 10% significance level.

Panels B and C disaggregate the results by baseline recycling. All coefficients go in the expected direction, and suggest sizeable 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 5% significance level.

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 in parentheses. ***, **, * indicate significance levels at 1, 5, and 10%, respectively. Comparing the marginal probabilities of T1, T2 and T3, respectively, using a one-sided wald test gives the following p-values:
Panel A: column (1) T1 > T2 p=0.254, T1 > T3 p=0.051, T2 > T3 p=0.161; column (2) T1 > T2 p=0.254, T1 > T3 p=0.053, T2 > T3 p=0.162; column (3) T1 > T2 p=0.255, T1 > T3 p=0.054, T2 > T3 p=0.162; column (4) T1 > T2 p=0.255, T1 > T3 p=0.055, T2 > T3 p=0.162;
Panel B: column (1) T1 > T2 p=0.390, T1 > T3 p=0.255, T2 > T3 p=0.343; column (2) T1 > T2 p=0.391, T1 > T3 p=0.256, T2 > T3 p=0.343; column (3) T1 > T2 p=0.392, T1 > T3 p=0.269, T2 > T3 p=0.345; column (4) T1 > T2 p=0.391, T1 > T3 p=0.255, T2 > T3 p=0.343;
Panel C: column (1) T1 > T2 p=0.233, T1 > T3 p=0.038, T2 > T3 p=0.147; column (2) T1 > T2 p=0.233, T1 > T3 p=0.040, T2 > T3 p=0.148; column (3) T1 > T2 p=0.234, T1 > T3 p=0.045, T2 > T3 p=0.151; column (4) T1 > T2 p=0.235, T1 > T3 p=0.047, T2 > T3 p=0.152.

3.6. 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.2 in Appendix D.2 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.3). 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.4), 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.5 in Appendix D.3 presents regressions results using the number of bags recycled as dependent variable. Our analysis shows that none of the treatment effects are statistically significant.

Dis-aggregating 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.6). 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 more efficient in terms of fewer bags used. Among households that did not recycle at baseline (see Table D.7), 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 to 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, 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 to follow through with good intentions. Moreover, by repeating the reminders with varying frequency, we aim to encourage repeated recycling behavior over time and support 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 frequency of the reminders. We find that reminders are generally effective at increasing recycling. In the first period where all treatment groups receive weekly reminders (period 2) recycling increases relative to the control group. This effect is 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 performs 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 travelling 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 finding that the continuation of reminders increases the likelihood that respondents start recycling 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) suggests that some respondents have not established a firm recycling habit yet when 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. Second, people who did not yet recycle at baseline might need a longer series of reminders to start recycling. Third, as to 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.

We end with two final reflections. First, the estimated effects most likely present an underbound. We were unable to check whether phone numbers were active, so that a proportion of participants might not have received our reminder messages. Second, we do not know if our intervention period was sufficiently long for the development of permanent recycling habits. A longer period might have led to 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. 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.

References

Allcott, Hunt, and Todd Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review*, 104(10): 3003–37.

Altmann, Steffen, and Christian Traxler. 2014. "Nudges at the Dentist." *European Economic Review*, 72: 19–38.

Apesteguia, Jose, Patricia Funk, and Nagore Iriberri. 2013. "Promoting Rule Compliance in Daily-Life: Evidence from a Randomized Field Experiment in the Public Libraries of Barcelona." *European Economic Review*, 64: 266–284.

Borasino, Elena, and Hanna Fuhrmann-Riebel. 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.

Cadena, Ximena, and Antoinette Schoar. 2011. "Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments." *NBER Working Paper Series No. 17020.*

Calzolari, Giacomo, and Mattia Nardotto. 2017. "Effective Reminders." *Management Science*, 63(9): 2915–2932.

Carlsson, Fredrik, Christina Gravert, Olof Johansson-Stenman, and Verena Kurz. 2021. "The Use of Green Nudges as an Environmental Policy Instrument." *Review of Environmental Economics and Policy*, 15(2): 216–237.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to Exercise." *Econometrica*, 77(3): 909–931.

Chong, Alberto, Dean Karlan, Jeremy Shapiro, and Jonathan Zinman. 2015. "(Ineffective) Messages to Encourage Recycling: Evidence from a Randomized Evaluation in Peru." *The World Bank Economic Review*, 29(1): 180–206.

Dai, Hengchen, Silvia Saccardo, Maria A Han, Lily Roh, Naveen Raja, Sitaram Vangala, Hardikkumar Modi, Shital Pandya, and Daniel M Croymans. 2021. "Behavioral Nudges Increase COVID-19 Vaccinations." *Nature*, 579: 404–409.

Damgaard, Mette Trier, and Christina Gravert. 2018. "The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising." *Journal of Public Economics*, 157: 15–26.

Datta, Saugato, and Sendhil Mullainathan. 2014. "Behavioral Design: A New Approach to Development Policy." *Review of Income and Wealth*, 60(1): 7–35.

Dustan, Andrew, Juan Manuel Hernandez-Agramonte, and Stanislao Maldonado. 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, Andrea, Angela Steffen, and Martin Staehle. 2021. "Choose to Reuse! The Effect of Action-Close Reminders on Pro-Environmental Behavior." *Journal of Environmental Economics and Management*, 110: 102539.

Gilbert, Ben, and Joshua Graff Zivin. 2014. "Dynamic Salience with Intermittent Billing: Evidence from Smart Electricity Meters." *Journal of Economic Behavior and Organization*, 107: 176–190.

Gravert, Christina. 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.*

Grelle, Sonja, Sascha Kuhn, Hanna Fuhrmann-Riebel, and Wilhelm Hofmann. 2024. "The Role of Framing and Effort in Green Nudging Acceptance." *Behavioural Public Policy*, 1–16.

Grácio, Matilde, and Pedro C. Vicente. 2021. "Information, Get-Out-The-Vote Messages, and Peer Influence: Causal Effects on Political Behavior in Mozambique." *Journal of Development Economics*, 151: 102665.

He, Guojun, Yuhang Pan, Albert Park, Yasuyuki Sawada, and Elaine S. Tan. 2023. "Reducing Single-Use Cutlery with Green Nudges: Evidence from China's Food-Delivery Industry." *Science*, 381(6662): eadd9884.

Huck, Steffen, and Imran Rasul. 2010. "Transactions Costs in Charitable Giving: Evidence from Two Field Experiments." *The B.E. Journal of Economic Analysis & Policy*, 10(1).

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science*, 62(12): 3393–3411.

Kirakozian, Ankinée. 2016. "One Without the Other? Behavioural and Incentive Policies for Household Waste Management." *Journal of Economic Surveys*, 30(3): 526–551.

Knickmeyer, Doris. 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.

Larochelle, Catherine, Jeffrey Alwang, Elli Travis, Victor Hugo Barrera, and Juan Manuel Dominguez Andrade. 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.

Miafodzyeva, S., and N. Brandt. 2013. "Recycling Behaviour Among Householders: Synthesizing Determinants Via a Meta-Analysis." *Waste and Biomass Valorization*, 4(2): 221–235.

Milkman, Katherine L., Dena Gromet, Hung Ho, Joseph S. Kay, Timothy W. Lee, Pepi Pandiloski, Yeji Park, Aneesh Rai, Max Bazerman, John Beshears, Lauri Bonacorsi, Colin Camerer, Edward Chang, Gretchen Chapman, Robert Cialdini, Hengchen Dai, Lauren Eskreis-Winkler, Ayelet Fishbach, James J. Gross, Samantha Horn, Alexa Hubbard, Steven J. Jones, Dean Karlan, Tim Kautz, Erika Kirgios, Joowon Klusowski, Ariella Kristal, Rahul Ladhania, George Loewenstein, Jens Ludwig, Barbara Mellers, Sendhil Mullainathan, Silvia Saccardo, Jann Spiess, Gaurav Suri, Joachim H. Talloen, Jamie Taxer, Yaacov Trope, Lyle Ungar, Kevin G. Volpp, Ashley Whillans, Jonathan Zinman, and Angela L. Duckworth. 2021*a*. "Megastudies Improve the Impact of Applied Behavioural Science." *Nature*, 600(7889): 478–483.

Milkman, Katherine L., Mitesh S. Patel, Linnea Gandhi, Heather N. Graci, Dena M. Gromet, Hung Ho, Joseph S. Kay, Timothy W. Lee, Modupe Akinola, John Beshears, Jonathan E. Bogard, Alison Buttenheim, Christopher F. Chabris, Gretchen B. Chapman, James J. Choi, Hengchen Dai, Craig R. Fox, Amir Goren, Matthew D. Hilchey, Jillian Hmurovic, Leslie K. John, Dean Karlan, Melanie Kim, David Laibson, Cait Lamberton, Brigitte C. Madrian, Michelle N. Meyer, Maria Modanu, Jimin Nam, Todd Rogers, Renante Rondina, Silvia Saccardo, Maheen Shermohammed, Dilip Soman, Jehan Sparks, Caleb Warren, Megan Weber, Ron Berman, Chalanda N. Evans, Christopher K. Snider, Eli Tsukayama, Christophe Van den Bulte, Kevin G. Volpp, and Angela L. Duckworth. 2021*b*. "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, Paul, and Wolfgang Habla. 2018. "Experimental and Non-Experimental Evidence on Limited Attention and Present Bias at the Gym." *ZEW Discussion Paper No. 18-041.*

Rankin, Catharine H., Thomas Abrams, Robert J. Barry, Seema Bhatnagar, David F. Clayton, John Colombo, Gianluca Coppola, Mark A. Geyer, David L. Glanzman, Stephen Marsland, Frances K. McSweeney, Donald A.

Wilson, Chun-Fang Wu, and Richard F. Thompson. 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.

Rodríguez, Catherine, and Juan E. Saavedra. 2019. "The Persistent Effects of Youth Savings Reminders: Experimental Evidence from Text-Message Campaigns in Colombia." *Journal of Development Economics*, 139: 135–156.

Royer, Heather, Mark Stehr, and Justin Sydnor. 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, Axel, and Daniel John Zizzo. 2015. "On Reminder Effects, Drop-Outs and Dominance: Evidence from an Online Experiment on Charitable Giving." *PLOS ONE*, 10(8): 1–17.

Thomas, Christine, and Veronica Sharp. 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, Verena, Lorenz Goette, Kathrin Degen, Vojkan Tasic, Elgar Fleisch, Rafael Lalive, and Thorsten Staake. 2018. "Overcoming Salience Bias: How Real-Time Feedback Fosters Resource Conservation." *Management Science*, 64(3): 1458–1476.

UNEP. 2015. "United Nations Environment Programme: Global Waste Management Outlook."

Varotto, Alessandra, and Anna Spagnolli. 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, Steven, Paul Ferraro, and Nathaniel Higgins. 2017. "Addressing Participant Inattention in Federal Programs: A Field Experiment with the Conservation Reserve Program." *American Journal of Agricultural Economics*, 99(4): 914–931.

WWF. 2018. "Almost 90% of the Garbage Generated Daily Is Not Recycled."

Appendix

A. Balance tests and descriptive overview

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	1,392

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, 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	1,392

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=1,392).

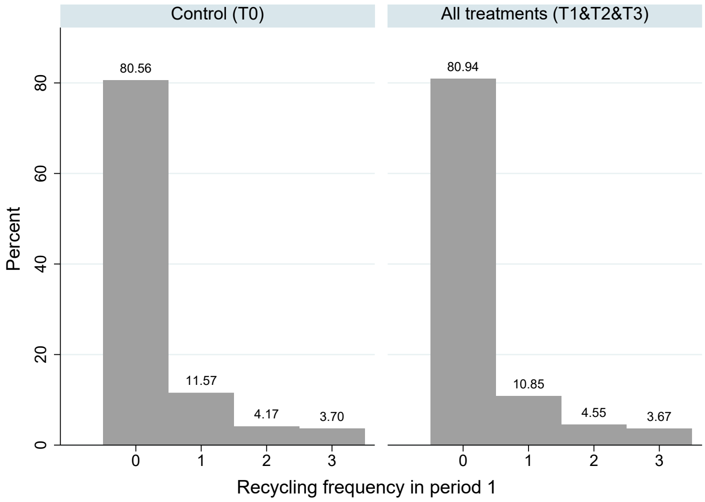
Table A.3: Household characteristics by group – recent sign-up years

		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.

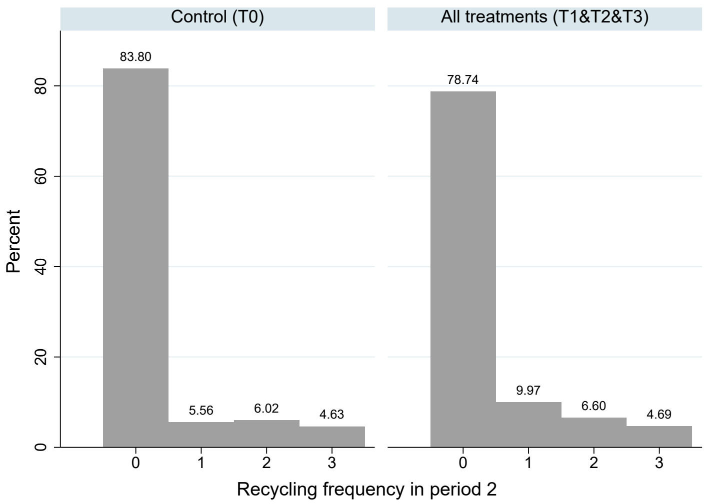
B. Histograms of recycling frequency per period

Figure B.1: Histograms of recycling frequency during the baseline period (week 1-3), comparing all treatment groups together with the control group



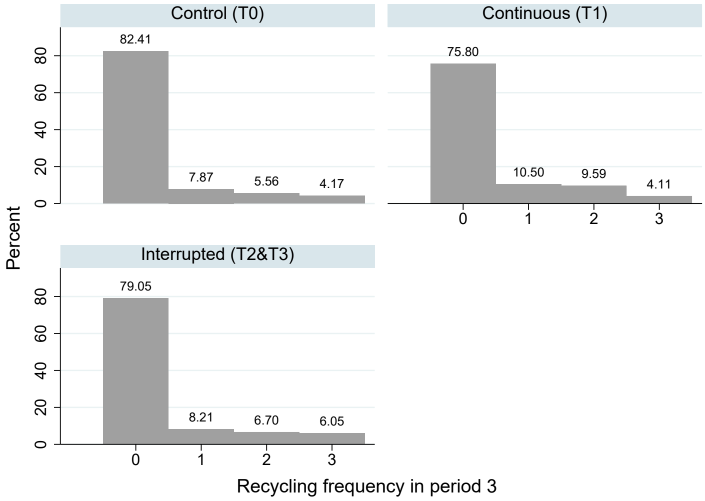
Notes: The figure shows histograms for the recycling frequency during the baseline period (from zero to three times) in the control group and in the aggregated treatment groups, focusing on households with a recent sign-up year ($\geq$ 2018). N in the control group (T0) = 216, the treatment groups together add up to N = 682.

Figure B.2: Histograms of recycling frequency during period 2 (week 4-6), comparing all treatment groups together with the control group



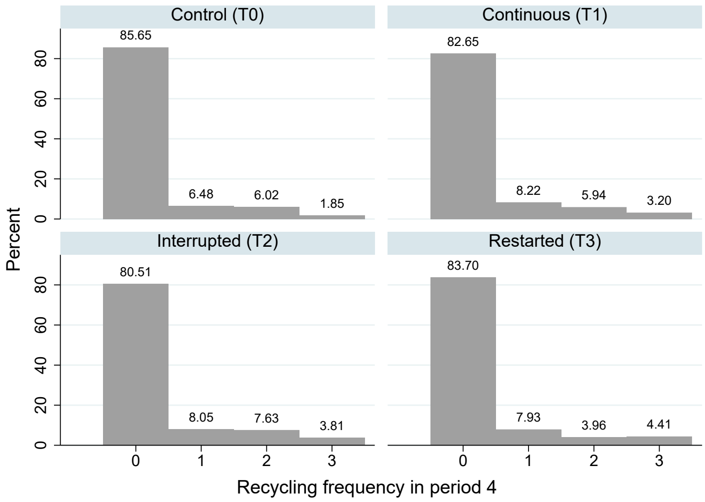
Notes: The figure shows histograms for the recycling frequency during period 2 (from zero to three times) in the control group and in the aggregated treatment groups, focusing on households with a recent sign-up year ($\geq$ 2018). N in the control group (T0) = 216, the treatment groups together add up to N = 682.

Figure 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 ($\geq$ 2018). N(T0) = 216, N(T1) = 219, N(T2&T3) = 463.

Figure B.4: Histograms of recycling frequency during period 4 (week 10-12), comparing the continuous, interrupted and restarted treatment groups with the control group

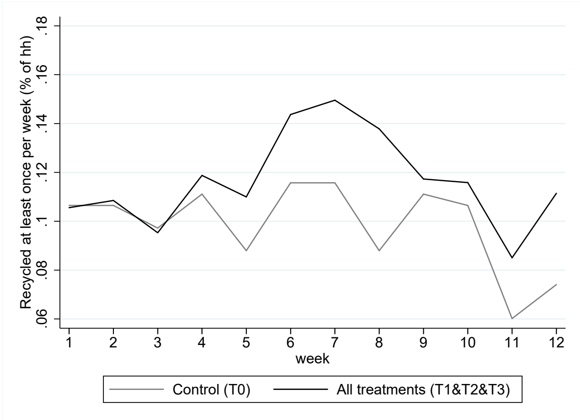


Notes: The figure shows histograms for the recycling frequency during period 4 (from zero to three times) in the control group and in the different treatment groups, focusing on households with a recent sign-up year ($\geq$ 2018). N(T0) = 216, N(T1) = 219, N(T2) = 236, N(T3) = 227.

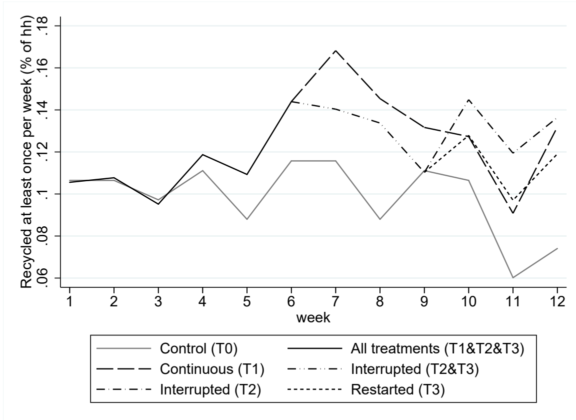
C. Graphs on weekly recycling behavior

Figure C.1: Percentages of households that recycled per week

(a) All treatments vs. control



(b) Treatment split



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 ($\geq$ 2018). Sub-figure C.1a compares all treatment groups pooled together with the control group; sub-figure C.1b 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.

D. Additional analyses

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

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***	
			(0.313)	
Interrupted (T2&T3)			0.397	
			(0.284)	
Continuous (T1)				0.495
				(0.320)
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 in parentheses; standard errors clustered at individual household level in column (1). ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

D.2. Recycling frequency per period

Table D.2: 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**	
			(0.057)	
Interrupted (T2&T3)			0.072	
			(0.049)	
Continuous (T1)				0.086*
				(0.048)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.073 for T1 > T2&T3; column (4) p=0.461 for T2 > T1, p=0.211 for T1 > T3, p=0.188 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table D.3: 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	
			(0.212)	
Interrupted (T2&T3)			0.169	
			(0.199)	
Continuous (T1)				0.097
				(0.208)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.169 for T1 > T2&T3; column (4) p=0.486 for T2 > T1, p=0.354 for T1 > T3, p=0.364 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table D.4: 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**	
			(0.050)	
Interrupted (T2&T3)			0.041	
			(0.037)	
Continuous (T1)				0.084**
				(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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.099 for T1 > T2&T3; column (4) p=0.483 for T2 > T1, p=0.178 for T1 > T3, p=0.164 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

D.3. Number of bags recycled per period

Table D.5: 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	
			(0.188)	
Interrupted (T2&T3)			-0.052	
			(0.165)	
Continuous (T1)				-0.008
				(0.153)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.070 for T1 > T2&T3; column (4) p=0.349 for T2 > T1, p=0.422 for T3 > T1, p=0.448 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table D.6: 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	
			(0.807)	
Interrupted (T2&T3)			-0.649	
			(0.762)	
Continuous (T1)				-0.918
				(0.737)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.224 for T1 > T2&T3; column (4) p=0.372 for T2 > T1, p=0.349 for T3 > T1, p=0.458 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

Table D.7: 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**	
			(0.134)	
Interrupted (T2&T3)			0.061	
			(0.089)	
Continuous (T1)				0.204**
				(0.079)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.045 for T1 > T2&T3; column (4) p=0.457 for T2 > T1, p=0.289 for T1 > T3, p=0.263 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

E. Full sample results

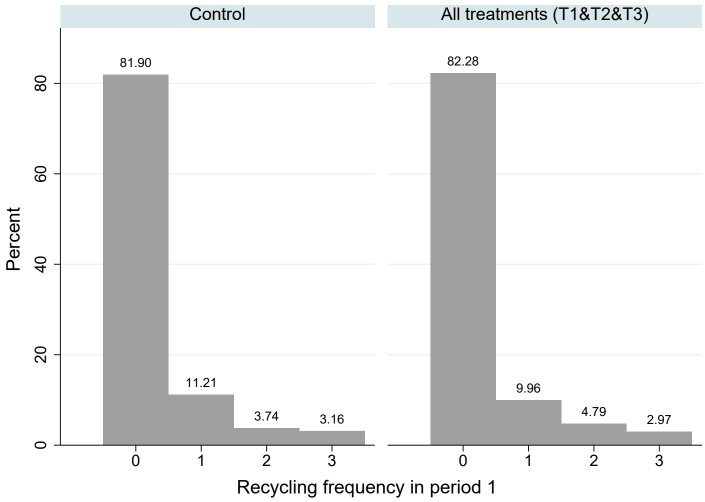
We conducted all analyses presented in this paper also with the full sample to verify whether and how any findings differ. As in our main analysis, we start by looking at the treatment effects by period (see Table E.1). Again, we find that reminders significantly increase households' recycling activity in period 2 among all treatment groups compared to the control group, and that this effect persists throughout the whole intervention period. In period 3, we can 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 can 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 Appendix 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. Again, the restarted treatment shows no significant effect on recycling levels. 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.

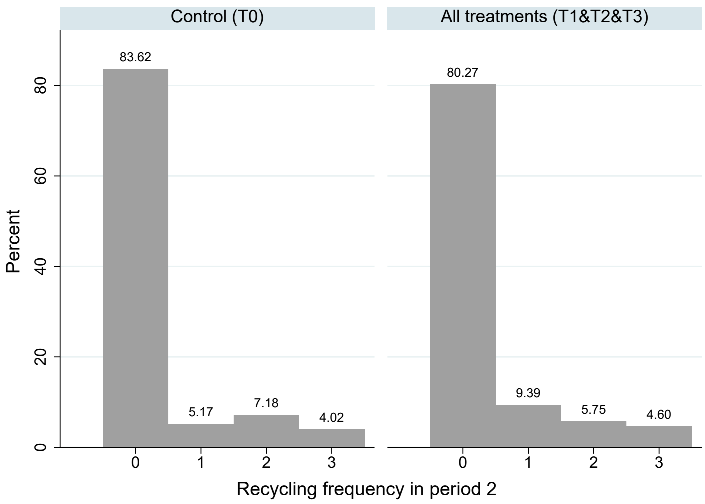
E.1. Full sample histograms of recycling frequency per period

Figure E.1: Histograms of recycling frequency during the baseline period (week 1-3), comparing all treatment groups together with the control group – full sample



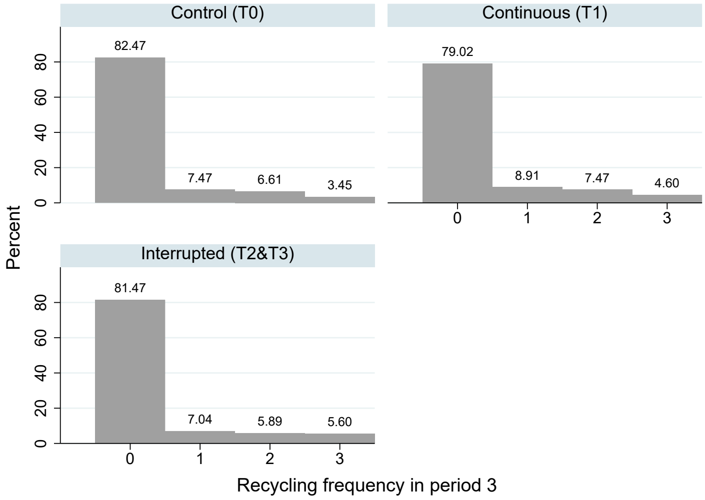
Notes: The figure shows histograms for the recycling frequency during the baseline period (from zero to three times) in the control group and in the aggregated treatment groups, including the full sample. N per group = 348, the treatment groups together add up to N = 1,044.

Figure E.2: Histograms of recycling frequency during period 2 (week 4-6), comparing all treatment groups together with the control group – full sample



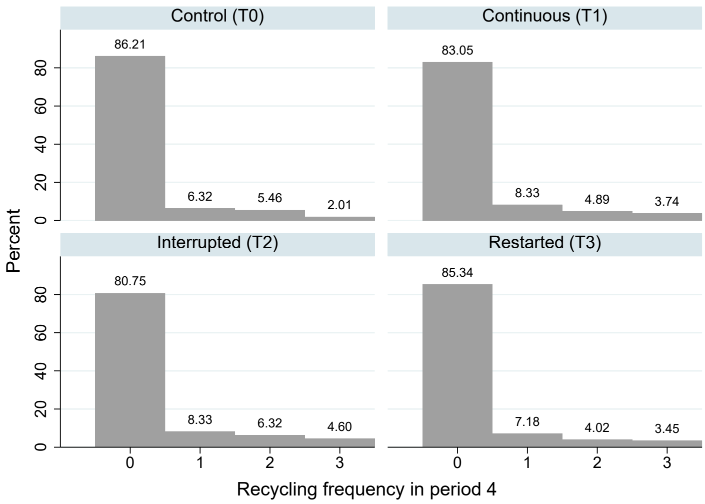
Notes: The figure shows histograms for the recycling frequency during period 2 (from zero to three times) in the control group and in the aggregated treatment groups, including the full sample. N per group = 348, the treatment groups together add up to N = 1,044.

Figure 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).

Figure 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

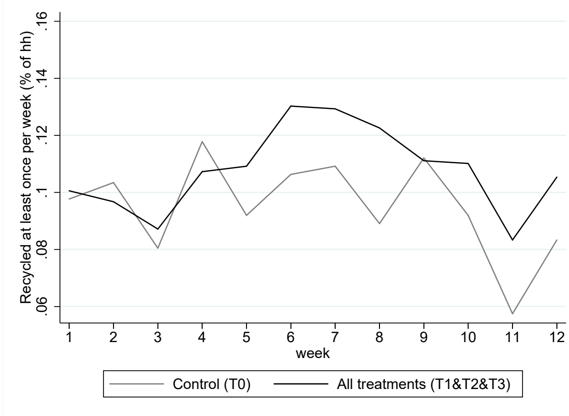


Notes: The figure shows histograms for the recycling frequency during period 4 (from zero to three times) in the control group and in the different treatment groups, including the full sample. N per group = 348.

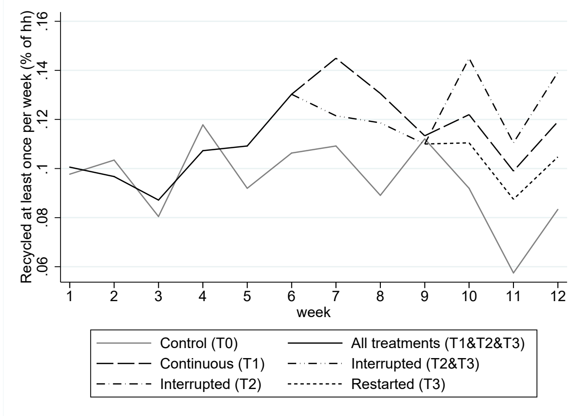
E.2. Full sample graphs on weekly recycling behavior

Figure E.5: Percentages of households that recycled per week – full sample

(a) All treatments vs. control



(b) Treatment split



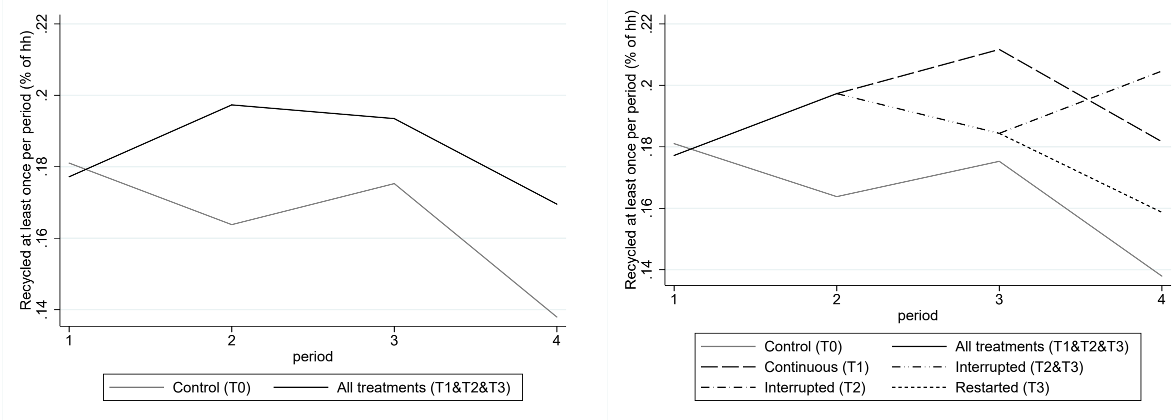
Notes: The figure shows the percentages of households that recycled at least once per week in the control group and in the different treatment groups over the whole study period, focusing on households with a recent sign-up year (>= 2019). Sub-figure E.5a compares all treatment groups pooled together with the control group; sub-figure E.5b splits up the different treatment groups based on the treatment variation by periods. 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

Figure E.6: Proportions of households that recycled at least once per period – full sample

(a) All treatments vs. control

(b) Treatment split



Notes: The figure shows the proportions 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. Sub-figure E.6a compares all treatment groups pooled together with the control group; sub-figure E.6b splits up the different treatment groups based on the treatment variation by periods. N(T0) = 348, N(T1&T2&T3) = 1,044, N(T1) = 348, N(T2&T3) = 698, N(T2) = 348, N(T3) = 348. The data is demeaned and re-centered around the pooled groups mean.

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^{**} (0.016)	0.037^{*} (0.020)		
Continuous (T1)			0.044^{*} (0.025)	
Interrupted (T2&T3)			0.014 (0.021)	
Continuous (T1)				0.040^{*} (0.023)
Interrupted (T2)				0.053^{**} (0.023)
Restarted (T3)				0.013 (0.022)
Constant	0.052^{***} (0.016)	0.058^{***} (0.020)	0.065^{***} (0.020)	0.031^{*} (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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.081 for T1 > T2&T3; column (4) p=0.291 for T2 > T1, p=0.125 for T1 > T3, p=0.044 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

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

	Period 2-4	Period 2	Period 3	Period 4
	(1)	(2)	(3)	(4)
All treatments (T1&T2&T3)	0.337**	0.372*		
	(0.170)	(0.210)		
Continuous (T1)			0.433*	
			(0.244)	
Interrupted (T2&T3)			0.157	
			(0.216)	
Continuous (T1)				0.448*
				(0.256)
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 in parentheses; standard errors clustered at individual household level in column (1). ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

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

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

	Period 2-4	Period 2	Period 3	Period 4
	(1)	(2)	(3)	(4)
All treatments (T1&T2&T3)	0.106*	0.154**		
	(0.058)	(0.072)		
Continuous (T1)			0.167**	
			(0.083)	
Interrupted (T2&T3)			0.075	
			(0.074)	
Continuous (T1)				0.061
				(0.089)
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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.099 for T1 > T2&T3; column (4) p=0.398 for T2 > T1, p=0.373 for T1 > T3, p=0.277 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

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

	Period 2-4	Period 2	Period 3	Period 4
	(1)	(2)	(3)	(4)
All treatments (T1&T2&T3)	0.015	0.011		
	(0.014)	(0.018)		
Continuous (T1)			0.019	
			(0.024)	
Interrupted (T2&T3)			-0.002	
			(0.019)	
Continuous (T1)				0.036*
				(0.020)
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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.153 for T1 > T2&T3; column (4) p=0.355 for T2 > T1, p=0.081 for T1 > T3, p=0.039 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

E.5. Full sample results: Mechanisms

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 (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 use fixed effects for enumerators. Full sample included. Standard errors in parentheses. Comparing treatment coefficients using one-sided wald tests gives following p-values: column (1) p=0.246 for T1 > T2&T3; column (2) p=0.109 for T1 > T2&T3; column (3) p=0.441 for T1 > T2&T3; column (4) p=0.123 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 (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 use fixed effects for enumerators. Full sample included. Only households included that recycled during the baseline period. Standard errors in parentheses. Comparing treatment coefficients using one-sided wald tests gives following p-values: column (1) p=0.272 for T1 > T2&T3; column (2) p=0.445 for T1 > T2&T3; column (3) p=0.030 for T1 > T2&T3; column (4) p=0.198 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

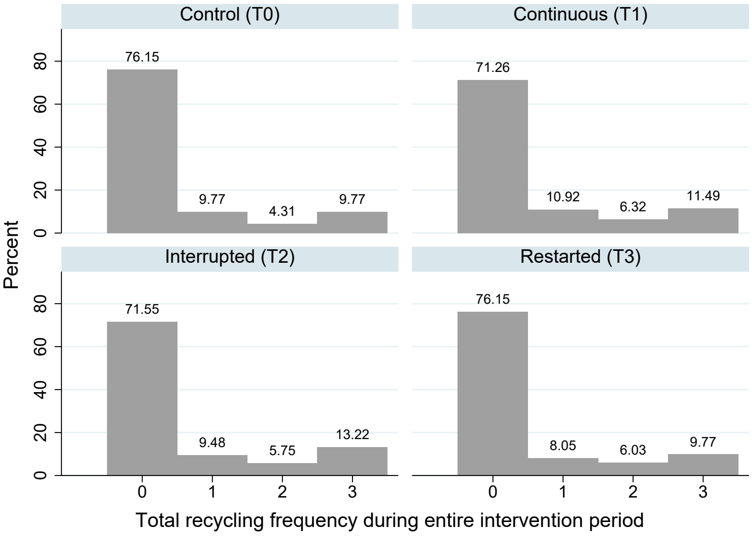
	Recycled p2 & p3 (1)	Recycled neither p2 nor p3 (2)	Recycled only p2 (3)	Recycled only p3 (4)
Continuous (T1)	0.007	-0.044	0.023	0.014
	(0.017)	(0.027)	(0.017)	(0.015)
Interrupted (T2&T3)	-0.001	-0.007	0.006	-0.003
	(0.014)	(0.021)	(0.013)	(0.012)
Predicted Probability Control (T0)	0.041***	0.904***	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 use fixed effects for enumerators. Full sample included. Only households included that did not recycle during the baseline period. Standard errors in parentheses. Comparing treatment coefficients using one-sided wald tests gives following p-values: column (1) p=0.288 for T1 > T2&T3; column (2) p=0.061 for T1 > T2&T3; column (3) p=0.101 for T1 > T2&T3; column (4) p=0.207 for T1 > T2&T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.

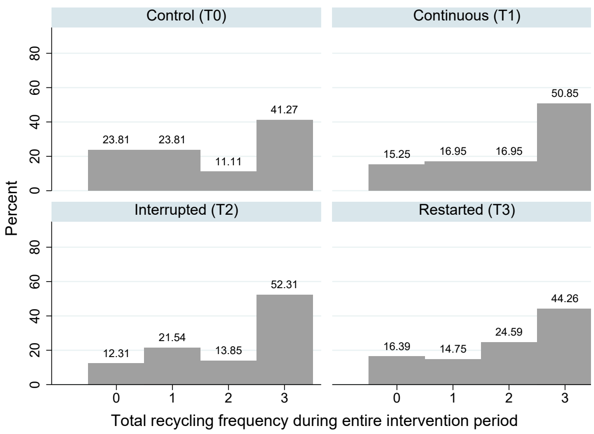
E.6. Full sample results: Treatment effects over the entire intervention period

Figure E.7: Total recycling frequency during entire intervention period, by treatment – full sample

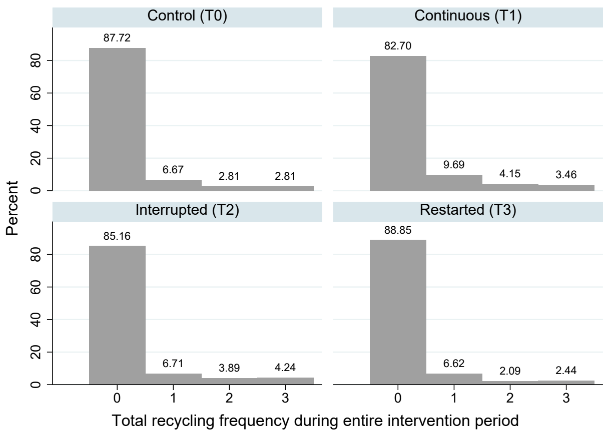
(a) Pooled sample



(b) Baseline recycling yes



(c) Baseline recycling no



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 ($\geq$ 2018). Sub-figure E.7a for pooled sample: N per group = 348; sub-figure E.7b for households that recycled during the baseline period: N(T0) = 63, N(T1) = 59, N(T2) = 65, N(T3) = 61; sub-figure E.7c 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 one-sided wald test gives the following p-values:
Panel A: column (1) T1 > T2 p=0.394, T1 > T3 p=0.026, T2 > T3 p=0.047; column (2) T1 > T2 p=0.394, T1 > T3 p=0.027, T2 > T3 p=0.049; column (3) T1 > T2 p=0.394, T1 > T3 p=0.028, T2 > T3 p=0.050; column (4) T1 > T2 p=0.394, T1 > T3 p=0.008, T2 > T3 p=0.050;
Panel B: column (1) T1 > T2 p=0.489, T1 > T3 p=0.281, T2 > T3 p=0.268; column (2) T1 > T2 p=0.489, T1 > T3 p=0.281, T2 > T3 p=0.267; column (3) T1 > T2 p=0.489, T1 > T3 p=0.291, T2 > T3 p=0.276; column (4) T1 > T2 p=0.489, T1 > T3 p=0.280, T2 > T3 p=0.267;
Panel C: column (1) T1 > T2 p=0.311, T1 > T3 p=0.017, T2 > T3 p=0.027; column (2) T1 > T2 p=0.312, T1 > T3 p=0.018, T2 > T3 p=0.054; column (3) T1 > T2 p=0.312, T1 > T3 p=0.021, T2 > T3 p=0.058; column (4) T1 > T2 p=0.312, T1 > T3 p=0.021, T2 > T3 p=0.058.

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

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

	Period 2-4	Period 2	Period 3	Period 4
	(1)	(2)	(3)	(4)
All treatments (T1&T2&T3)	0.052*	0.030		
	(0.030)	(0.036)		
Continuous (T1)			0.088*	
			(0.047)	
Interrupted (T2&T3)			0.042	
			(0.040)	
Continuous (T1)				0.078*
				(0.041)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.127 for T1 > T2&T3; column (4) p=0.311 for T2 > T1, p=0.106 for T1 > T3, p=0.041 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	Period 2	Period 3	Period 4
	(1)	(2)	(3)	(4)
All treatments (T1&T2&T3)	0.178	0.138		
	(0.122)	(0.156)		
Continuous (T1)			0.315*	
			(0.173)	
Interrupted (T2&T3)			0.210	
			(0.156)	
Continuous (T1)				0.156
				(0.178)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.241 for T1 > T2&T3; column (4) p=0.395 for T2 > T1, p=0.353 for T1 > T3, p=0.265 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	
			(0.041)	
Interrupted (T2&T3)			0.003	
			(0.033)	
Continuous (T1)				0.064**
				(0.032)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.134 for T1 > T2&T3; column (4) p=0.370 for T2 > T1, p=0.067 for T1 > T3, p=0.035 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

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	
			(0.151)	
Interrupted (T2&T3)			-0.041	
			(0.137)	
Continuous (T1)				0.079
				(0.132)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.108 for T1 > T2&T3; column (4) p=0.277 for T2 > T1, p=0.383 for T1 > T3, p=0.199 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	
			(0.676)	
Interrupted (T2&T3)			-0.300	
			(0.653)	
Continuous (T1)				-0.198
				(0.662)
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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.189 for T1 > T2&T3; column (4) p=0.274 for T2 > T1, p=0.476 for T3 > T1, p=0.311 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 (0.053)	0.070 (0.056)		
Continuous (T1)			0.123 (0.107)	
Interrupted (T2&T3)			-0.004 (0.081)	
Continuous (T1)				0.149* (0.078)
Interrupted (T2)				0.143* (0.078)
Restarted (T3)				0.066 (0.074)
Constant	0.196*** (0.057)	0.195*** (0.057)	0.271*** (0.079)	0.119* (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 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 in parentheses; standard errors clustered at individual household level in column (1). Comparing treatment coefficients using one-sided wald tests gives following p-values: column (3) p=0.098 for T1 > T2&T3; column (4) p=0.474 for T1 > T2, p=0.170 for T1 > T3, p=0.188 for T2 > T3. ***, **, * indicate significance levels at 1, 5, and 10%, respectively.