

TI 2012-061/1
Tinbergen Institute Discussion Paper



The Power of a Bad Example – A Field Experiment in Household Garbage Disposal

*Robert Dur*¹

*Ben Vollaard*²

¹Erasmus University Rotterdam, Tinbergen Institute, CESifo, and IZA;

²Tilburg University, CentER, and TILEC.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

The Power of a Bad Example – A Field Experiment in Household Garbage Disposal*

Robert Dur[†] and Ben Vollaard[‡]

April 8, 2014

Abstract

Field-experimental studies have shown that people litter more in more littered environments. Inspired by these findings, many cities around the world have adopted policies to quickly remove litter. While such policies may prevent people from following the bad example of litterers, they may also invite free-riding on public cleaning services. We are the first to show that both forces are at play. We conduct a natural field experiment where, in a randomly assigned part of a residential area, the frequency of cleaning was drastically reduced during a three-month period. We find evidence that some people start to clean up after themselves when public cleaning services are diminished. However, the tendency to litter more dominates. We also find that these responses continue to exist for some time after the treatment has ended.

JEL-codes: C93, H40, K42

Keywords: littering, public services, free-riding, field experiment.

*We gratefully acknowledge comments and suggestions from the Editor in Chief, two referees, Olivier Marie, Jan Stoop, Christian Traxler, participants of the Advances with Field Experiments 2011 Workshop at the University of Chicago, the New Directions in Welfare 2011 conference at the OECD in Paris, the 3rd Transatlantic Workshop on the Economics of Crime at Collegio Carlo Alberto in Moncalieri, the 10th Tiber Symposium on Psychology and Economics at Tilburg University, and the WZB Field Days in Berlin. We also thank seminar audiences at the University of Munich, Erasmus University Rotterdam, Stockholm University, and Tilburg University. Special thanks to the City of Rotterdam, the Charlois city district, and the community service providers Roteb and Stadstoezicht who made the field experiment possible. We are most grateful to Maya von Harras for providing daily project management, to Herman Coomans and Arthur Andjelic for collecting and processing the data, and to Peter van Dalen for his role in the early stages of the project.

[†]Erasmus University Rotterdam, Tinbergen Institute, CESifo, IZA. E-mail: dur@ese.eur.nl. Corresponding author. Postal address: Erasmus School of Economics H9-15, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

[‡]Tilburg University, CentER, TILEC. E-mail: b.a.vollaard@uvt.nl. Postal address: Kamer K 331, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

1. Introduction

When a person breaks the law, he often inflicts more damage to society than just the direct negative consequences of his act. The transgression may be an invitation for others to also break the law. This spillover effect constitutes an additional negative externality from illegal behavior (see e.g. Keizer et al. 2008).

As a case in point, consider littering. In most countries, littering is illegal and punishable by a fine. A litterer spoils a neighborhood's appearance and, in the longer run, harms the environment. In addition to these direct negative externalities, litterers set a bad example for others to follow. Mimicking may arise for several reasons. First, some people may simply imitate other people's behavior, thinking that "If everyone is doing it, it must be a sensible thing to do." (Cialdini et al. 1990: 1015). That is, the presence of litter affects the perceived descriptive norm for littering, resulting in a normative influence on behavior. Second, people may infer from observing litter that littering is tolerated by the police and by the community at large, thus reducing the fear for formal and social sanctions (Kahan 1997). In other words, litter can affect people's beliefs about the prevailing injunctive norm for littering. Such norms might change behavior via social pressure or providing information or both. Third, people may behave in a conditionally cooperative manner (Gächter 2007), and so may be discouraged to contribute to a clean environment when they observe that others did not do so. Last, preferences for cleanliness may be such that the decline in well-being from an additional piece of litter is relatively small when the environment is already littered as compared to when the environment is clean (Anderson and Francois 1997).

These behavioral spillover effects can give rise to negative feedback loops, transforming a clean and orderly neighborhood into a dirty and messy place. A series of experimental studies

have provided evidence supporting this idea. In these studies, researchers experimentally induced variation in the amount of litter present at places like grocery stores (Geller et al. 1977), picnic areas (Crump et al. 1977), a waiting room for participants in a lab (Kraus et al. 1978), a parking garage (Reiter and Samuel 1980, Cialdini et al. 1990, and Reno et al. 1993), an amusement park (Cialdini et al. 1990), the lobby of a dormitory on a university campus (Cialdini et al. 1990), an academic department's common room (Ramos and Torgler 2012), and an alley in a large shopping area (Keizer et al. 2011).¹ With a few exceptions (Crump et al. 1977 and Reno et al. 1993), these experimental studies find that people litter significantly more often in littered environments as compared to clean environments.² In most studies, the effect is substantial. For instance, in Cialdini et al. (1990: 1017)'s study into littering of a handbill that was tucked under the windshield wiper of parked cars, 11 percent of people litter in a clean environment whereas 41 percent of people do so in a littered environment.

Based on the findings from these one-shot, small-scale field experiments, it has often been argued that frequent clean-ups of public places yield a double dividend (Reiter and Samuel 1980, Cialdini et al. 1990, Huffman et al. 1995, and Anderson and Francois 1997). First, the frequent clean-ups ensure that litter is quickly removed. Second, the resulting litter-free environment may keep people from further littering. That is, government provision of cleaning efforts can encourage rather than discourage private contributions to a clean environment. Several local governments have followed this advice and have implemented intensive periodic cleaning programs in residential and nonresidential areas alike. Many practitioners believe that

¹ Finnie (1973) and Schultz et al. (2013) also study littering in clean and dirty areas, but the cleanliness of the area is not experimentally varied, making it impossible to infer causal relationships.

² Cialdini et al. (1990) provide some evidence that the effect is nonlinear: a single piece of litter seems to make an anti-litter norm more salient than a completely clean environment. They also find that the positive effect of a heavily littered environment on littering is significantly greater when people observe an experimental confederate littering (see also Reno et al. 1993). Keizer et al. (2008) show that also other signs of disorder, such as unreturned shopping carts standing around in disarray in a parking garage and illegally set off fireworks cause littering.

such programs have been highly effective in discouraging littering (Lewis et al. 2009, Schultz and Stein 2009, and Ouiller and Sauneron, 2011).

The results from one-shot experiments may not hold in settings where people experience cleaning services over a longer time period, however. In line with the standard theory of private provision of public goods (Bergstrom et al. 1986), the frequent clean-ups may invite the public to free-ride on government-provided services. Such behavior is particularly likely to occur over time, when citizens have the time to learn about the provision of government services and adjust their behavior accordingly. From what people state in surveys, free-riding on government services seems a relevant concern in this context. For example, surveys conducted in the US reviewed by R.W. Beck Inc. (2007) suggest that people litter more if they know that someone else will clean up after them. Similar findings are reported by Lewis et al. (2009) for the UK. They note that “many people think that if someone else is paid to clean up [...] then littering can be justified. More specifically, those questioned often named this “someone else” as the local authority.” (Lewis et al. 2009: 19). Clearly, if many people think along these lines, government provision of cleaning efforts will *increase* rather than decrease littering. In other words, a policy of frequent or sustained cleaning may backfire.

This paper provides the first test of the littering response to publicly provided cleaning services in the longer term. We report the results of a natural field experiment where, in a randomly assigned part of a residential area, public cleaning services were drastically reduced during a period of three months. In contrast to the existing one-shot experiments, which focus on the spontaneous behavioral response, our setting allows for learning effects among residents. The experiment took place in Rotterdam, the second largest city in The Netherlands, from December 2010 to February 2011. Using high-frequency data on litter at treated and control

locations before, during, and after the experiment, we examine whether people litter more when the environment is less often cleaned up, how this behavior develops over time, and whether it continues to exist after the treatment has ended.³

The focus of our study is on littering around underground containers for collection of household garbage in a densely populated area. The containers are placed on the sidewalk and feature a lid that is to be opened to dispose of garbage (see Figure 1). A large, clearly visible warning label on the container states that putting garbage on the street is illegal and subject to a 115 euro fine. Discarded household items that are too big to be put into the container should either be brought to a nearby garbage collection depot or be picked up by municipality workers after making an appointment by telephone, at no charge.

Despite the fine and the free legal alternatives, littering around underground containers occurs frequently – to the annoyance of many residents. In surveys, littered streets feature high on residents’ lists of greatest annoyances, together with speeding, dog feces, illegal parking, and nuisance from youth (IVM 2009). In response to this public concern and inspired by some of the above-mentioned experimental studies, the municipality decided a few years ago to provide cleaning services at all container locations in the area on a daily basis rather than at the regular frequency of two or three times a week. Our experimental treatment is to abandon these extra services and to return to the regular cleaning frequency in a randomly assigned part of the area for a period of three months, while keeping the frequency of emptying the container constant. The natural setting of the experiment does not allow us to isolate the specific behavioral mechanisms driving the response to a change in public cleaning services. Our

³ We do not consider effects on other offenses or other outcomes because of data limitations. Thus, unlike for instance Keizer et al. (2008), our paper is not a test of Wilson and Kelling’s (1982) broken windows theory.

primary aim is to put previous work on behavioral mechanisms to the test within people's own habitat and over a long period of time.

Our paper contributes to several strands in the literature. First, we make a contribution to the empirical literature on littering behavior, a literature that we reviewed above. Existing studies are limited to short-term interventions (lasting one day or less) and focused on people's spontaneous response (to litter or not to litter). Our field experiment is the first to study the effects of an intervention that lasted several months, while also measuring behavior in the month after the treatment. This relatively long duration allows us to study possible learning effects. Arguably, it takes a while before people realize that they can no longer free ride on frequent public cleaning services.

Second, our paper contributes to a recent literature in economics showing that people tend to follow examples set by others. This has for instance been shown in the context of donations to charities (Frey and Meier 2004, Martin and Randal 2008, Shang and Croson 2009), tax evasion (Fellner et al. 2013), welfare participation (Bertrand et al. 2000), and in public good games played in laboratory experiments (Burlando and Guala 2005, Gächter 2007, Engel et al. 2011 and 2014).

Lastly, our paper is related to the empirical literature on private contributions to public goods (see Nyborg and Rege 2003 and Payne 2009 for surveys). Our findings point to a delicate balance between publicly provided services and voluntary contributions from citizens. Too low a level of public services may result in a bad equilibrium in which most citizens choose not to cooperate. Too high a level of public services may discourage private contributions, and leave it up to the state to provide the public good.

2. Method

A. Littering and cleaning activities

The experiment was conducted in a residential area of some 4,000 households within the city of Rotterdam, from November 29, 2010 until March 3, 2011. The residential area is part of Charlois, a densely populated and relatively poor city district. The area counts 41 locations with underground garbage containers.⁴ The containers are placed on the sidewalk. The part of the container that is visible at street level features a lid that is to be opened to dispose of garbage (see Figure 1). Each container serves some 60 households.

It is illegal to place garbage next to the container. Every container includes a large warning label stating the 115 euro fine for illegal disposal of garbage next to the container. Residents can make an appointment by telephone for pickup of discarded household items that are too big to fit in the container, at no charge. Alternatively, the items can be disposed of freely at a nearby garbage collection depot (at approximately 7 miles distance from this area). According to data provided to us by community service provider Roteb, every week on average 18 households from this area make an appointment per week, which is equal to 0.5 percent of all households in the area.

Residents can break the law, first of all by putting large household items such as broken furniture next to the container. It is tempting to take the easy way out: it takes time and effort to bring garbage to a nearby collection depot or to set up an appointment for pick-up.

⁴ The total number of containers in the area equals 70: the 41 locations include 15 with one container, 23 with two containers, and three with three containers.

Moreover, when making an appointment, it often takes a few days before garbage can be picked up, implying that people need to store garbage in their house or backyard for some time (as a matter of fact, appointments could be made for Tuesdays only). When garbage is illegally placed on the sidewalk, it is almost always put next to the container rather than somewhere else. Supposedly, the container is a safe location: at this common collection point it is close to impossible to trace the offender, quick removal is secured since this is where service workers stop to clean up, and neighbors may be less annoyed by items placed next to the container rather than someplace else (e.g. in front of someone's door).

Second, it is also tempting for some people to place a garbage bag next to a container rather than in a container. It is a way of avoiding the effort of opening the lid and pulling the lever. That is a small effort, but it requires the use of some force and one may get his hands wet or dirty. In rare cases, doing the right thing is impossible, because the container is stuck or full or other garbage blocks access to the container. From data we collected in a way that we discuss later in this section, we know that 90 percent of the time containers are neither stuck nor full. We also know that residents are almost always able to reach the container and open the lid: in 90 percent of the cases that litter is present, the number of bags and types of discarded household items is equal to 3 or less. Further note that, even if a container is stuck or full or cannot be accessed, it is still illegal to put garbage next to the container. In those cases, citizens are supposed to use a nearby container. Alternative container locations are rarely more than one block away. To conclude, placing garbage next to the container is primarily the result of taking the easy way out rather than technical reasons. That makes our setting comparable to settings used in previous studies into littering.

The garbage containers are emptied two or three times a week, either in the early morning or the late afternoon, according to a fixed schedule. Most of the containers in the area are emptied on Monday and Thursday afternoon. Before emptying, the area around the containers is cleaned of all litter. That is necessary to prevent litter from falling into the exposed underground space when the container is being emptied. In addition to the clean-up before emptying, the area around the containers is cleaned every morning by a crew of two, both of them driving a small truck. Taken together, it was standard practice to clean the area around the containers at least once a day, and sometimes even twice a day. This has been the case since November 2007 when the daily cleaning activities by the crew of two were introduced (i.e., more than three years before the start of our experiment). Hence, we are confident that past policies no longer have a noticeable influence on residents' behavior.

B. Data collection

Littering around the garbage containers was recorded twice daily on weekdays from Monday, September 20, 2010 until Friday, April 1, 2011. During these six months, the same two non-uniformed observers made a walk through the area, following a route designed by us, firstly from 8.00 to 9.30 am, which is before the daily cleaning round in the morning, and secondly from 1.30 to 3.00 pm. During each round, the observers stopped by each container location, and noted the number of garbage bags next to the container. They also noted whether discarded household items were present, split between six different categories (audiovisual equipment; household appliances; furniture; home improvement material; paper and cardboard; other). Finally, they noted whether the container was stuck or full. Whether a container is stuck rather than full is not always easy to determine, this is why we created one indicator for the two conditions.

In the two months before the start of the intervention, on average 31 percent of locations were littered with one or more bags or household items at the time of the morning round and 11 percent of locations at the time of the afternoon round. The containers were stuck or full 13 percent of the time in the morning and 6 percent in the afternoon on average during the baseline period. The distribution of littering over locations was roughly stable, but heavily skewed. The reasons for why some locations tend to be more littered than others within the same neighborhood and sometimes even the same street are not clearly understood. Correlational evidence suggests that the presence of shops and particular types of containers are positively related to the degree of accumulation of garbage (see Appendix).

C. Treatment, hypotheses, and main outcome variable

To examine how littering behavior changes in response to public cleaning services, the frequency of cleaning around the garbage containers was reduced for a period of three months in a randomly assigned part of the area. The daily cleaning by the crew of two in the morning was cancelled in the treatment area. As a consequence, cleaning was reduced from one or two times a day to two or three times a week. The treatment was not communicated to residents so as to avoid confounding effects arising from communication. Hence, the main channel by which residents learn about cleaning activities is the cleanliness of the area around the garbage container. Since most residents dispose of their garbage multiple times per week, and also live in close proximity to the container, they have frequent exposure to the condition of the container location (maximum walking distance to the nearest container is 75 meters or 250 feet, see Figure 2 for a map of the area).⁵ Anecdotal evidence from service workers suggests that residents have one particular container that they always use. We test the response to the

⁵ Some residents may also have noticed that the two service workers no longer clean up litter around the garbage container in the morning.

conditions at this specific container location. Residents in both the treatment and control area are also exposed to conditions at container locations that they do not use, which may lead to treatment dilution. We do not expect different effects for container locations that are at the border of the treatment and control area: in all cases except one, container locations in one area do not provide a direct view of container locations in the other area.

We focus the empirical analysis on the accumulation of garbage next to the container from the early afternoon until the next morning. During these hours accumulation of litter is not distorted by cleaning activities, except when the container is emptied late afternoon. We drop all observations of accumulation from afternoon to next morning that are distorted by cleaning activities. This way the number of hours during which garbage is accumulated is similar across observations. As discussed above, we know the number of bags next to the container, but have an indicator variable for the presence of six different categories of household items. Thus we measure whether the number of categories of household items increased. Given the heterogeneity in discarded household items, ranging from a broken chair to a cardboard box, the chance of having another item of a similar category is small. This implies that a change in the number of categories of household items gives a good indication of accumulation of this type of garbage.

Garbage can be either bags or household items. We are mainly interested in the overall response in littering behavior to the treatment. Given the heterogeneity in accumulation of the number of garbage bags and the six categories of household items, we create an indicator variable which is one in case of accumulation of any type of garbage and zero otherwise. Advantages of using an indicator variable rather than the number of accumulated items are that the results become less sensitive to outliers in the number of accumulated items and that

the indicator variable is not distorted by the disposal of several items by one individual at the same time. However, in the sensitivity analysis, we also report results for the total number of accumulated items as well as the number of garbage bags and the number of types of household items.

Figure 3 illustrates how the indicator is defined. The first two diagrams show that reduced cleaning in and of itself does not affect the indicator. That is only the case when less cleaning results in *additional* garbage in the early morning, as in the third, fourth and fifth diagram. The latter three diagrams show that we treat accumulation of one or more bags and one or more disposed household items as equivalent. Even though the indicator variable is a summary measure of the build-up of garbage, it allows for both variation in accumulation of garbage between locations and variation within the same container location. The average incidence of accumulation of garbage by container location varied between 0 and 50 percent in the treatment area during the baseline period, leaving ample room for more frequent accumulation of garbage during the experiment.

If litter begets litter, then we expect more frequent accumulation of garbage when the cleaning frequency is reduced compared to the control area. If people learn to clean up after themselves, then we expect to see less frequent accumulation of garbage compared to the control area in the long run.

For every week, we have four observations of the accumulation of garbage per container location: Monday to Tuesday, Tuesday to Wednesday, Wednesday to Thursday, and Thursday

to Friday.⁶ As mentioned above, we exclude observations for locations that were cleaned in the afternoon due to scheduled emptying of the container. These regular cleaning activities distort the accumulation of garbage from the early afternoon until the next morning. The 2 or 3 days per week that the containers are emptied differ from location to location. Since a great majority of the containers are emptied on Monday and Thursday afternoon, we exclude the accumulation from Monday to Tuesday and from Thursday to Friday completely. Another reason to exclude these observations is enforcement activities in the afternoon that were scheduled for Mondays and Thursdays in both the control and treatment area. These enforcement activities, albeit at a low level and conducted haphazardly, may occasionally distort the accumulation of garbage, as they involve removing and opening garbage bags so as to search for the identity of the offenders. In total, we have observations for 50 days for the period September 20, 2010 until April 1, 2011. The total number of observations including the month after the treatment is equal to 1,730.⁷

D. Randomization procedure

Randomization happened at the group level rather than the individual level. Logistical constraints made co-location of treatment locations necessary. That meant that the residential area had to be cut in two parts. The cut of choice was the one for which the two parts were most similar in terms of presence of illegally disposed garbage in the early afternoon in the months before the start of the intervention (see Figure 2). A coin toss determined which of the two halves was chosen as treatment group. The treatment group includes 21 container

⁶ We exclude data on accumulation of garbage from Friday to Monday because cleaning activities take place throughout the neighborhood on Sundays.

⁷ Over the course of four months, we miss one day due to sickness of the observers and one day due to a late start of their morning round. We miss one day in the week before the start of the treatment, because of miscommunication to the service workers about the nature of the treatment. Due to adverse weather conditions in December, cleaning activities were altered during two days, resulting in another two missing days (December 1 and 8, 2010). Finally, we miss one day due to closure of garbage containers around New Year's Eve.

locations; the other 20 locations serve as control group. In the sensitivity analysis, we show spatial correlation between co-located container locations not to be an issue: our results are robust to collapsing the data to two groups – treatment and control.

E. Baseline characteristics

Table 1 summarizes the baseline characteristics of container locations in our sample. As discussed above, the randomization procedure guaranteed balance between the treatment and control area with respect to the presence of garbage in the early afternoon. The percentage of locations that were littered in the afternoon is indeed similar between the two areas before the start of the intervention (September 20-November 26, 2010). In the treatment group, litter is observed at 11 percent of locations in the early afternoon; in the control group at 12 percent of locations. A t-test shows the two averages not to be statistically significantly different. We find similar patterns when we split litter into garbage bags and disposed household items (row 2 and 3), when we focus on the number of bags or household items (row 4 and 5), and also for the incidence of a stuck or full container (row 6). Our key outcome variable, the accumulation of garbage from early afternoon to early morning the next day, is significantly larger in the control area than in the treatment area (see row 7). The observation that residents in the control area litter more than residents in the treatment area is confirmed by the relatively high number of fines for illegal garbage disposal in the control area (row 8). The number of telephone appointments for garbage pick-up is also statistically significantly higher in the control area than in the treatment area (row 9). Naturally, in the empirical analysis, we control for pre-existing differences in the accumulation of garbage by including container location fixed-effects (we use the statistical software package STATA 13 and xtreg as our primary estimation

command).⁸ We also provide evidence for similar trends in the accumulation of garbage prior to the start of the treatment.

Table 2 provides some further information on possible pre-existing differences between the treatment and control area based on another source of data. In 2009, residents in both areas were interviewed as part of a national crime survey (IVM 2009). 131 randomly selected households from the area were interviewed for the survey, 70 in the treatment area and 61 in the control area. Table 2 presents the items from the survey that are related to our study. Perceived disorder from littering, worry about crime, age of residents, and employment status are not statistically significantly different between the two areas.

3. Results

Before we start our main analysis, let us first examine how the lower frequency of cleaning in the morning affected the presence of litter in the early afternoon. During the experiment, with the morning cleaning round cancelled in the treatment area, presence of litter in the early afternoon strongly increased in the treatment group (from 11 to 27 percent of locations on average) and remained roughly stable in the control group (from 12 to 15 percent of locations on average). Thus, quite naturally, taking away cleaning activities in the morning resulted in a much more littered environment early afternoon. In the remainder of this section, we examine how this affected littering behavior from early afternoon until the next morning.

⁸ Occasionally, the amount of garbage declined from the early afternoon to the morning the next day. This happened 4 percent of the time during the baseline period. In the empirical analysis, these observations are classified as “no accumulation of garbage”. Our results do not change when we code these observations as missing.

A. Non-parametric evidence

To analyze how littering behavior changed as a result of the reduced frequency of cleaning, we non-parametrically estimate the difference in accumulation of garbage between the control and treatment area for each day before, during and after the experiment. This flexible estimation framework allows us to analyze trends before the start of the intervention, changes in the behavioral response during the months of the intervention, and continuation of the behavioral response after the cleaning frequency returned to its original level. We estimate the following equation:

$$I(\Delta L_{i,t}=1) = \sum_{t=1}^Z D_{i,t} \alpha_t + \gamma_i + \delta_t + \varepsilon_{i,t} \quad (1)$$

$I(\Delta L_{i,t}=1)$ is an indicator function which is one in case of accumulation of garbage from the early afternoon until the next morning of day t at container location i and zero otherwise. $D_{i,t}$ are daily dummy variables which are one for container locations in the treatment area and zero for container locations in the control area. We include container location-fixed effects γ_i to control for location-specific factors driving accumulation of litter that are constant over time. We include day-fixed effects δ_t to account for common shocks in littering, such as weather conditions and official holidays. $\varepsilon_{i,t}$ is the idiosyncratic error. To allow for serial correlation in observations by container location, we cluster the standard errors at the container level using the Huber/White/sandwich estimator.

In Figure 4 the daily coefficients, α_t , from estimating equation (1) are plotted in light gray – as open dots during the treatment period and closed dots outside the treatment period. On the horizontal axis, time 0 corresponds with the first day of the experiment (November 29, 2010). The vertical axis shows the percentage point difference in accumulation of garbage between

the treatment and control area. The coefficients are estimated relative to the first day for which we have data (September 22, 2010).^{9,10} The separately estimated monthly coefficients, plotted in black and superimposed on the daily coefficients in Figure 4, illustrate the major trend in littering behavior before, during, and after the intervention.

First of all, Figure 4 shows the absence of a difference in trend between treatment and control before the start of the treatment, which is in line with the common trend assumption underlying causal interpretation of the parameter of interest. Second, we find littering to go up sharply as soon as cleaning services are reduced. Accumulation of litter at container locations goes up by some 15 percentage-points on average. The figure suggests that the frequency of positive accumulation increased by about 75 percent (the baseline frequency is some 20 percent, see Table 1). Below, we conduct a test whether the average rate of littering during the experiment in the treatment area relative to the control area is statistically significantly different from zero. Finally, Figure 4 suggests persistence of the behavioral response: littering does not return to its original level once the cleaning round in the morning is re-established – at least not in the first month after the intervention for which we have data. Residents show an immediate response to less frequent cleaning of litter, but seem slow in adjusting to more frequent cleaning.

B. Average treatment effect

Next, we test whether the treatment had on average a statistically significant effect on littering behavior. We estimate the following linear probability model:

⁹ As discussed in Section 2, monitoring started two days earlier. We exclude Mondays and Tuesdays, which is why the first observation is for Wednesday, September 22.

¹⁰ We scale the coefficients to the mean difference in accumulation during the baseline period.

$$I(\Delta L_{i,t}=1) = \alpha T_{i,t} + \gamma_i + \delta_t + \varepsilon_{i,t} \quad (2)$$

where T is the treatment dummy, which is one in the experimental area during the intervention period and zero otherwise. Other than replacing the daily dummies by T , estimation equation (2) is identical to equation (1). Parameter of interest α is the estimated change in incidence of accumulation of garbage from the afternoon to the morning the next day.

Table 3 presents the estimation results. We find that the incidence of litter accumulation is on average 15 percentage points higher when the cleaning frequency is reduced. Given litter accumulation at some 20 percent of locations during the baseline period, this implies a 75 percent increase. The effect is statistically significant at the 99 percent confidence level. In the second column of Table 3, we allow the size of the treatment effect to vary during the intervention period. As can be expected from the pattern in littering behavior observed in Figure 4, the estimated coefficient for the second half of the period is similar to the coefficient for the first half of the period. A Wald-test shows that the difference between the two parameter estimates is not statistically significantly different from zero.¹¹

C. Persistence

We test for post-intervention continuation of the treatment effect in the third column of Table 3. Some of the theories discussed in the introduction naturally predict that treatment effects do not disappear immediately when the frequency of cleaning returns to its regular level. Particularly, it may take a while before citizens' beliefs about the risk of formal and social sanctions and about other people's contributions to a clean environment adjust to the new context. Furthermore, it may take some time before people realize that the local government

¹¹ Our conclusions do not change when we divide the treatment period into three rather than two periods.

is again cleaning up on a daily basis. Lastly, people may form habits in response to the treatment. We find evidence for post-intervention continuation of the treatment effect. During the first month after the treatment has ended, the treatment effect is of similar size as the effect during the intervention. The coefficient is statistically significantly at the 99 percent confidence level.

D. Evidence for discouragement

The results so far lend strong support to the hypothesis that public cleaning services encourage private contributions. The data on litter at treatment and control locations clearly show that when public cleaning services are diminished, private contributions to a clean environment go down on average. The average treatment effect may hide contrasting behavioral responses, however. To further analyze such effects, we focus on the use of a legal alternative for disposal of garbage: making a telephone appointment for pick-up of discarded oversized household items. If public provision encourages private contributions, then we expect to see a fall in the number of telephone appointments in response to the treatment. In the case of discouragement, we expect to see the reverse. Community service provider Roteb provided weekly data on these telephone appointments. The data relate to all of the treatment area and all of the control area rather than to specific container locations.

Similar to the analysis of accumulation of illegally disposed garbage, first we non-parametrically estimate the effect of the treatment on the number of appointments. We estimate the following equation:

$$A_{i,t} = \sum_{t=1}^Z D_{i,t} \alpha_t + \gamma_i + \delta_m + \varepsilon_{i,t} \quad (3)$$

where $A_{i,t}$ is the number of appointments for pick-up of discarded household items made in area i and week t . $D_{i,t}$ are weekly dummy variables that are one for the treatment area and zero for the control area. We include area-fixed effects γ_i and month-fixed effects δ_m .¹²

Figure 5 plots the weekly coefficients, α_t , from estimating equation (3) in light gray. On the horizontal axis, time 0 corresponds with the first week of the experiment (first week of December, 2010). For purposes of illustration, we extend the time window that we have used so far with an additional 6 weeks before and 6 weeks after the period of the daily monitoring of illegal disposal of garbage (August, 2010-May, 2011). The vertical axis shows the difference in the number of telephone appointments for pick-up of discarded household items between the treatment and control area. The coefficients are estimated relative to the first week for which we have data on appointments.¹³ To illustrate the major trend in the number of appointments, we separately estimate monthly coefficients, which are plotted in black in Figure 5 and superimposed on the weekly coefficients.

In line with the common trend assumption and similar to our analysis of accumulation of illegally disposed garbage in Figure 4, we do not observe a trend before the start of the treatment in Figure 5. As soon as cleaning services are reduced, we find the number of telephone appointments to *go up* in the treatment area relative to the control area. This increase in the number of appointments is in line with the discouragement hypothesis. The high frequency of cleaning before the treatment started apparently discouraged citizens to make telephone appointments for pick-up of their garbage.

¹² Given the number of parameters to be estimated in equation (3), we opt for month-fixed effects rather than week-fixed effects. Later, when we estimate the average treatment effect, we use week-fixed effects (equation 4).

¹³ We scale the coefficients to the mean difference in the number of appointments during the baseline period.

This finding stands in contrast to the average treatment effect for illegal disposal of garbage that we found above. The effect is also small relative to our previous result. In the treatment area, the number of appointments for pick-up of discarded household items went up by 2 to 3 *per week*, while the number of container locations with accumulation of garbage went up by about 3 *per day*.¹⁴ The behavioral effect is not insignificant within the context of this legal alternative, however. Given an average number of appointments in the treatment area before the start of the intervention of 8.6 per week, this amounts to about a one-third increase.

Again, in line with the findings for accumulation of illegally disposed garbage, the behavioral effect of the treatment seems to continue after the treatment has ended. The number of telephone appointments in the treatment area relative to the control area remains higher even when the morning cleaning rounds are re-established.

To test whether the effect of the treatment on this type of legal disposal of household garbage is statistically significantly different from zero, we estimate the following equation:

$$A_{i,t} = \alpha T_{i,t} + \gamma_i + \delta_t + \varepsilon_{i,t} \quad (4)$$

where T is the treatment dummy, which is one in the experimental area during the treatment period and zero otherwise. Parameter of interest α is the estimated change in the number of telephone appointments during the period of the treatment. We include area-fixed effects γ_i and week-fixed effects δ_t .

¹⁴ As discussed in Section 3B, the frequency of accumulation of garbage went up by 15 percentage-points, from about 20 percent in the baseline period. With 21 container locations in the treatment area, this translates into about 3 additional sites with accumulation of garbage.

Table 4 presents the estimation results. In the first column, we estimate the average treatment effect using data for the same period as in Table 3 (September, 2010-March, 2011). We find the number of appointments for pick-up of discarded household items to go up by about 3 in response to the reduced frequency of cleaning, an increase of about a third. The effect is statistically significant at the 95% confidence level. In the second column of Table 4, we allow the size of the effect to vary over the treatment period. The point estimates are not statistically significantly different from each other. In the third column, we also consider the month after the treatment. We find some evidence for continuation of the treatment effect on appointments after the treatment has ended. The point estimates for the treatment effect and the effect during the period after the intervention are similar in size; the latter is, however, not statistically significantly different from zero. In the last column, similar to Figure 5, we extend the time window of the analysis with 6 weeks before and 6 weeks after the period of the daily monitoring. The estimation results are similar to those based on the shorter period reported in column (3). The post-treatment effect on appointments becomes statistically significant when using this extended time window, providing further evidence for continuation of the behavioral effect.

D. Sensitivity analysis

So far, we assumed variation in littering to be independent across the 41 container locations. Given the randomization at the area level rather than the container level, for reasons discussed in Section 2, the variation in littering among co-located container locations within the treatment area and within the control area may not be fully independent, rendering our statistical tests of the treatment effect invalid. To see how robust our results are to the possibility of spatial correlation among container locations within the treatment and control area, we collapse the data by area. In the second row of Table 5 we present the results. Even

though we greatly reduce the number of observations that we use to identify the effect of the treatment, we find the results to be similar. This finding provides support for our assumption that spatial correlation due to co-location of treatment and control locations does not bias our results.

A particular feature of our experimental design is the great number of time periods that we use to estimate the treatment effect. The period before and during the treatment encompasses observations for 42 days over a period of four months. If we would ignore the potential serial correlation in observations by container location, then we would be likely to underestimate the standard errors (Bertrand, Duflo, and Mullainathan 2004). To address this issue, we have clustered the standard errors by container location in the default specification. Not clustering the standard errors would not lead to different conclusions; it hardly affects the standard errors. An alternative solution to the potential problem of serial correlation is to collapse the time series information into a pre- and post-period. The results are presented in the third row of Table 5. Even though the number of observations is greatly reduced, we still find the effect to be statistically significantly different from zero at the 95 percent confidence level. The standard errors are similar to those of the default specification, suggesting that we successfully correct for potential serial correlation within clusters.

In the fourth and fifth row of Table 5, we estimate the treatment effect separately for garbage bags and for discarded household items. We find both estimated coefficients to be positive, but only the coefficient for discarded household items is statistically significant. When studying the data at this disaggregated level, the incidence of illegal behavior becomes too small to identify the treatment effect in a very precise way. The coefficients are of similar size given the higher accumulation of household items relative to garbage bags during the baseline

period (Table 1). This suggests that residents respond similarly to the treatment irrespective of the type of garbage, justifying our outcome measure that combines the two.

As an additional test, we estimate the effect on the accumulation of the number of garbage bags and the number of types of household items rather than a binary indicator. The outcome variable is now based on count data truncated at zero, and the data are characterized by overdispersion. For these two reasons, we use a negative binomial regression model. In line with Allison and Waterman (2002), we estimate an unconditional negative binomial regression estimator with dummy variables to represent the fixed effects.¹⁵

The results are reported in rows (6) and (7). In both cases, we find a positive effect from the treatment, which is statistically significant at the 95 percent level. The first coefficient implies a 140 percent increase in the accumulation of garbage bags due to the treatment (from a base level of 0.12 in the treatment area). This is almost twice as high as the previously estimated increase in the indicator variable. Apparently, litter accumulation at treatment locations also becomes more intense. The second coefficient implies a 75 percent increase in the accumulation of types of discarded household items (from a base level of 0.19 in the treatment area). This estimate is not directly comparable to the default estimate based on the indicator variable, however, since we measure the number of types of household items next to the container rather than the actual number of household items (see Section 2).¹⁶

¹⁵ Our findings are robust to the specific regression model: a poisson model and even a linear model provide qualitatively similar results.

¹⁶ If we allow the coefficients for these two alternative outcome variables to vary between the first and the second half of the treatment period, then we find the parameters to be of similar size. The differences between the estimates for the first and the second half of the treatment period are not statistically significant. This is in line with our previous findings reported in Table 3.

E. Possible confounds

The two service workers whose morning round was cancelled during the treatment did not only remove garbage, but also enforced laws pertaining to garbage disposal and occasionally checked whether containers were stuck. This may confound our results in two ways.

First, the treatment lowered the level of law enforcement in the treatment area because the two service workers regularly open garbage bags to search for the identity of offenders. If they find sufficient evidence that links a bag to a household, they impose a 115 euro fine. The fine arrives in the mail two weeks later. People who illegally dispose household items such as a couch or an appliance can only be caught *in flagrante*, which almost never happens. Community service provider Stadstoezicht Rotterdam provided us with container-level data on fines for illegal disposal of garbage. The data show that the number of fines per container location per month was more than halved in the treatment area during the experiment, whereas it remained more or less stable in the control area. However, the effect on littering is likely to be small, because of the very low number of fines. Before the experiment, per month some 25 out of the 4,000 households received a fine for illegal disposal of garbage, which comes down to about one per location every two months on average. Given the number of bags observed, the implied chance of getting caught is smaller than 1 percent. Data for previous years show similar low number of fines. Still, to prevent a possible estimation bias, we include the number of fines per location as a covariate in the estimation equation. We lag the effect by 14 days because of the time it takes to process the fine. We assume the fine to have an effect on littering for 14 days. In other words, we count fines written at most four weeks ago and at least two

weeks ago.¹⁷ Row (8) of Table 5 reports the results. Controlling for fines does not affect the estimated treatment effect.

In theory, the brief presence in the morning of two service workers may have a deterrent effect in and of itself. If so, there may be an upward bias in the estimated effect of reduced cleaning on littering behavior. This potential bias is limited because of the low level of law enforcement, the short time of the visit of the two workers, the fact that they also had to drive through a couple of treatment streets on their daily route during the experiment, and the focus of the empirical analysis on what happens from the afternoon until the next morning, i.e. in the absence of the service workers.

Second, it is possible that containers are stuck more often in the treatment area during the treatment period. Occasionally, the service workers pull the lever of the container when it seems stuck, fix a container that has become stuck, or take out garbage that is sitting in the opening of a stuck container. With the morning round cancelled during the treatment, greater accumulation of garbage next to the container from the early afternoon to the next morning may simply be the result of a greater frequency of stuck containers in the early afternoon in the treatment area relative to the control area. If so, our results are mainly driven by the specific design of the container, limiting the external validity of our results. The monitoring data suggest that the share of containers that were stuck or full in the treatment area relative to the control area went up only slightly during the treatment. To exclude this channel, we include an indicator of a container being stuck or full in the early afternoon as an additional covariate in the estimation equation. Row (9) of Table 5 reports the results. We find the estimated effect

¹⁷ Our results are similar when we assume the fines to have an effect on littering behavior for more than two weeks.

not to be changed, indicating that this specific design feature of the garbage containers is not driving our results.¹⁸

Lastly, from November to January, two large cargo containers were placed during the weekend as an extra means to dispose of discarded household items. So as not to distort the experiment, one container was placed in the center of the treatment area and one container in the center of the control area. It could be that for some unknown reason residents in one area made greater use of the cargo container than residents in the other area. In that case, estimation of the treatment effect could be biased. Community service provider Roteb provided us with the exact weights of the garbage collected at the two cargo containers. When we include these weights as covariate in the empirical analysis, we find the same treatment effect (see the last row of Table 5), suggesting that the two temporarily placed cargo containers do not affect our results.¹⁹

4. Discussion

Litter is a serious problem in many communities around the world (Lewis et al. 2009; Schultz and Stein 2009). Quick removal of litter is seen by many as an effective policy response, not only because it ensures a clean environment, but also because the presence of litter may encourage further littering. In this view, government provision of cleaning services encourages private contributions to a clean environment, and so yields a ‘double dividend’. Others have argued that government provision may discourage private contributions. Why clean up after

¹⁸ Some residents may see the presence of garbage next to the container as a signal of the container being stuck or full. They may also use this as an easy excuse not to do the right thing. This could be an additional channel explaining our results.

¹⁹ Our results are also robust to including the three additional covariates (number of fines, stuck or full container, garbage disposed in cargo containers) at the same time.

yourself if government service workers stop by frequently to do the job? As described in the introduction, existing empirical work on this issue provides mixed evidence. On the one hand, a number of small-scale, one-shot field experiments show that litter begets litter. On the other hand, survey evidence suggests that many people think that littering can be justified if the local authority cleans up.

We have reported the results from a large-scale and long-lasting field experiment, in which we experimentally reduced the frequency of public cleaning services for a period of 3 months in people's own habitat. We find evidence that some residents start to clean up after themselves when public cleaning services are diminished, as suggested by the increase in the number of appointments for pick-up of discarded household items, a legal alternative to littering. The tendency to litter more dominates, however. Even in the longer term, government provision of cleaning services mainly encourages private contributions. In areas where the frequency of cleaning was reduced, the tendency to litter went up by some 75 percent. Our unique contribution is to show that both forces are at play.

We also find evidence for continuation of the treatment effects after the treatment has ended. People are quick to respond to less frequent cleaning, in the manner described above, but slow to respond when cleaning returns to its original frequency. During the first month after the treatment ended, littering as well as telephone appointments for pick-up of discarded household items remained at elevated levels. In other words, people do not immediately litter less if the environment is cleaner. This finding stands in stark contrast to results from existing one-off experiments in the psychology literature. A possible explanation for this difference might be the length of the intervention, with our experiment exposing people for a period of three months to a more littered environment, while the existing experiments are one-off.

Further, it might be that in a residential area learning effects are stronger than in non-residential areas.

A limitation of our study is that, for practical reasons, we could not randomly assign households or streets to treatment and control. Instead, we randomized over two adjacent areas, and controlled for baseline differences in the statistical analysis. Further, since the experiment took place in a relatively small neighborhood, we cannot completely rule out that the treatment had some spillover effects to the control area. Relatedly, we do not know whether additional litter is caused by a single litterer or multiple people. Lastly, a weakness of our study is that our data cannot meaningfully disentangle the four possible mechanisms behind the ‘litter begets litter’ effect that we find. Future studies could address these issues by running a larger scale experiment and setting up more extensive data collection. We hope that our study will stimulate policy makers and social scientists to collaborate more often so as to thoroughly learn about people’s short-run and long-run responses to anti-littering policies through experimental studies.

References

Allison, Paul D., and Richard Waterman (2002), Fixed effects negative binomial regression models, In: Ross M. Stolzenberg (ed.), *Sociological Methodology*, Oxford: Basil Blackwell.

Anderson, Siwan, and Patrick Francois (1997), Environmental cleanliness as a public good: Welfare and policy implications of nonconvex preferences, *Journal of Environmental Economics and Management*, 34(3): 256-274.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian (1986), On the private provision of public goods, *Journal of Public Economics*, 29(1): 25-49.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), How much should we trust differences-in-differences estimates?, *Quarterly Journal of Economics*, 119(1): 249-275.

Bertrand, Marianne, Erzo F.P. Luttmer, and Sendhil Mullainathan (2000), Network effects and welfare cultures, *Quarterly Journal of Economics*, 115(3): 1019-1055.

Burlando, Roberto M., and Francesco Guala (2005), Heterogeneous agents in public goods experiments, *Experimental Economics*, 8(1): 35-54.

Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren (1990), A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places, *Journal of Personality and Social Psychology*, 58(6): 1015-1026.

Crump, S. Larry, Dennis L. Nunes, and E.K. Crossman (1977), The effects of litter on littering behavior in a forest environment, *Environment and Behavior*, 9(1): 137-146.

Engel, Christoph, Sebastian Kube, and Michael Kurschilgen (2011), Can we manage first impressions in cooperation problems? An experimental study on “broken (and fixed) windows”, Preprints of the Max Planck Institute for Research on Collective Goods, Bonn 2011/5.

Engel, Christoph, Martin Beckenkamp, Martin, Andreas Glöckner, Bernd Irlenbusch, Heike Hennig-Schmidt, Sebastian Kube, Michael Kurschilgen, Alexander Morell, Andreas Nicklisch, Hans-Theo Normann, and Emanuel Towfigh (2014), First impressions are more important than early intervention: Qualifying broken windows theory in the lab, *International Review of Law and Economics*, 37(1): 126-136.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler (2013), Testing enforcement strategies in the field: Threat, moral appeal, and social information, *Journal of the European Economic Association*, 11(3): 634-660.

Finnie, William C. (1973), Field experiments in litter control, *Environment and Behavior*, 5(2): 123-144.

Frey, Bruno S., and Stephan Meier (2004), Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment, *American Economic Review*, 94(5): 1717-1722.

Gächter, Simon (2007), Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications, In: Bruno S. Frey and Alois Stutzer (Eds.), *Economics and psychology; A promising new cross-disciplinary field*, Cambridge, MA: MIT Press.

Geller, E. Scott, Jill F. Witmer, and Margaret A. Tuso (1977), Environmental interventions for litter control, *Journal of Applied Psychology*, 62(3): 344-351.

Huffman, Kim T., William F. Grossnickle, John G. Cope, and Kenneth P. Huffman (1995), Litter reduction: A review and integration of the literature, *Environment and Behavior*, 27(2): 153-183.

IVM [Integrale Veiligheidsmonitor] (2009), Buurtproblemen in Rotterdam-Rijnmond.

Kahan, Dan M. (1997), Social influence, social meaning, and deterrence, *Virginia Law Review*, 83: 349-395.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg (2008), The spreading of disorder, *Science*, 322: 1681-1685.

Keizer, Kees, Siegwart Lindenberg, and Linda Steg (2011), The reversal effect of prohibition signs, *Group Processes and Intergroup Relations*, 14(5): 681-688.

Krauss, Robert M., Jonathan L. Freedman, Morris Whitcup (1978), Field and laboratory studies of littering, *Journal of Experimental Social Psychology*, 14(1): 109-122.

Lewis, Alan, Polly Turton, and Thomas Sweetman (2009), Litterbugs. How to deal with the problem of littering, London: Policy Exchange.

Martin, Richard, and John Randal (2008), How is donation behaviour affected by the donations of others?, *Journal of Economic Behavior and Organization*, 67(1): 228-238.

Nyborg, Karine, and Mari Rege (2003), Does public policy crowd out private contributions to public goods?, *Public Choice*, 115(3-4): 397-418.

Ouiller, Olivier, and Sarah Sauneron (2011), 'Green nudges': New incentives for ecological behavior, La Note d'Analyse No. 216, Centre d'Analyse Stratégique/Office of the Prime Minister, Paris.

Payne, A. Abigail (2009), Does government funding change behavior? An empirical analysis of crowd out, In: Jeffrey R. Brown and James M. Poterba (Eds.), *Tax policy and the economy*, Volume 23, Chicago: University of Chicago Press.

R.W. Beck Inc. (2007), Literature review – Litter; A review of litter studies, attitude surveys and other litter-related literature, report for Keep America Beautiful, Seattle.

Ramos, João, and Benno Torgler (2012), Are academics messy? Testing the broken windows theory with a field experiment in the work environment, *Review of Law and Economics*, 8(3): 563-577.

Reiter, Susan M., and William Samuel (1980), Littering as a function of prior litter and the presence or absence of prohibitive signs, *Journal of Applied Social Psychology*, 10(1): 45-55.

Reno, Raymond R., Robert B. Cialdini, and Carl A. Kallgren (1993), The transsituational influence of social norms, *Journal of Personality and Social Psychology*, 64(1): 104-112.

Schultz, P. Wesley, and Steven R. Stein (2009), Litter in America: National findings and recommendations, Washington, DC: Keep America Beautiful.

Schultz, P. Wesley, Renée J. Bator, Lori Brown Large, Coral M. Bruni, and Jennifer J. Tabanico (2013), Littering in context: Personal and environmental predictors of littering behavior, *Environment and Behavior*, 45(1): 35-59.

Shang, Jen, and Rachel Croson (2009), A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods, *Economic Journal*, 119: 1422-1439.

Wilson, James Q., and George L. Kelling (1982), Broken windows. The police and neighborhood safety, *The Atlantic Monthly*, 249 (3): 29–38.

Table 1. *Baseline characteristics and randomization check*

	Treatment area	Control area	Difference
Illegally disposed garbage present (%) ¹	10.6 (30.8)	11.9 (32.4)	-1.3
Illegally disposed garbage bags present (%) ¹	2.7 (16.3)	3.7 (18.9)	-1.0
Illegally disposed household items present (%) ¹	8.5 (27.9)	9.4 (29.1)	-0.9
Number of illegally disposed garbage bags present ¹	0.046 (0.38)	0.066 (0.43)	-0.020
Number of types of illegally disposed household items present ¹	0.095 (0.33)	0.111 (0.37)	-0.015
Stuck or full container (%) ¹	6.9 (25.3)	5.4 (22.7)	1.5
Accumulation of garbage (%) ²	16.6 (37.2)	30.9 (46.3)	-14.3***
Number of fines for illegal garbage disposal ³	0.26 (0.50)	0.38 (0.68)	-0.12***
Number of telephone appointments for garbage pick-up	7.6 (2.9)	10.7 (2.9)	-3.1**

Note. All but the last row based on daily data for 41 container locations during baseline period (September-November, 2010), excluding days at which a container location was emptied. Last row based on weekly data for treatment and control area during baseline period. Number of observations is 987 for treatment area and 940 for control area, except for accumulation of garbage: 308 and 326, respectively, and the last row: 10 observations for treatment area and also for control area (weekly data). (1) Condition in the early afternoon; (2) Accumulation from the early afternoon to the early morning from Tuesday to Wednesday and from Wednesday to Thursday; (3) Fines written at most 4 and at least 2 weeks ago, per container location. Standard deviations between parentheses. ** $p < .05$; *** $p < .01$.

Table 2. *Perceived crime and disorder and background characteristics of residents in treatment and control area*

	Treatment area	Control area	Difference
'In this neighborhood, roads, paths and squares are well-kept' (% agree)	46.4 (50.2)	48.3 (50.4)	-2.0
'In this neighborhood, streets are frequently littered' (% agree)	58.6 (49.6)	55.7 (50.1)	2.8
'In this neighborhood, the area around garbage containers is frequently littered' (% agree)	67.1 (47.3)	60.1 (49.3)	6.5
'I am frequently worried about my safety in this neighborhood' (% agree)	14.3 (35.3)	12.1 (32.9)	2.2
Age of residents†	40.6 (14.5)	41.4 (14.2)	-0.8
Paid work for more than 12 hours per week	61.4 (49.0)	57.4 (49.9)	4.1
Number of observations	70	61	

Source: IVM (2009).

Note. † Only residents aged 12 or over were interviewed. Standard deviations between parentheses. ** $p < .05$; *** $p < .01$.

Table 3. *The effect of reduced cleaning frequency on accumulation of garbage*

Dependent variable: accumulation of garbage next to container from early afternoon until early morning next day (indicator)	Main specification	Allowing for varying treatment effect over time	Including post-intervention data and allowing for post-intervention effect
Treatment	0.153 (0.053)***		0.153 (0.053)***
Treatment – first half		0.156 (0.066)**	
Treatment – second half		0.150 (0.057)***	
Treatment – post intervention			0.169 (0.061)***
Number of observations	1,474	1,474	1,730

Note. Observations by container location and by day. Standard errors between parentheses corrected for clustering at the container level. Other covariates are container location-fixed effects and day-fixed effects. ** $p < .05$; *** $p < .01$.

Table 4. *The effect of reduced cleaning frequency on appointments for pick-up of discarded household items*

Dependent variable: number of telephone appointments for pick-up of garbage	(1)	(2)	(3)	(4)
Treatment	3.18 (1.23)**		3.18 (1.32)**	2.71 (1.22)**
Treatment – first half		3.90 (1.59)**		
Treatment – second half		2.67 (1.43)*		
Treatment – post intervention			2.70 (1.69)	2.53 (1.25)**
Number of observations	44	44	54	78

Note. Observations by area and by week. Standard errors between parentheses. Other covariates are area-fixed effects and week-fixed effects. * $p < .10$; ** $p < .05$; *** $p < .01$.

Table 5. *Sensitivity analysis*

Dependent variable: accumulation of garbage	
(1) Default	0.153 (0.053)***
(2) Collapse to two areas	0.153 (0.044)***
(3) Collapse to two periods	0.129 (0.053)**
(4) Garbage bags only	0.043 (0.039)
(5) Household items only	0.110 (0.047)**
(6) Number of garbage bags	0.876 (0.411)**
(7) Number of types of household items	0.562 (0.255)**
(8) Control for fines	0.151 (0.054)***
(9) Control for stuck or full container	0.151 (0.053)***
(10) Control for use of cargo containers	0.153 (0.053)***

Note. Observations by container location and by day. Number of observations equal to 1,474, except row (2) with 84 observations, and row (3) with 82 observations. Standard errors between parentheses corrected for clustering at the container level, except for (2) and (3). All rows report regression results of a linear probability model, except for rows (6) and (7) which are based on a negative binomial regression model. Row (8) includes as a covariate the number of fines for illegal disposal of garbage written at most four weeks ago and at least two weeks ago. Row (9) includes an indicator of a stuck or full container in the early afternoon as covariate, row (10) the weights of garbage disposed at temporarily placed cargo containers. Other covariates are area-fixed effects and day-fixed effects (row 2), container location-fixed effects and period-fixed effects (row 3), and container location-fixed effects and day-fixed effects (all other columns). ** $p < .05$; *** $p < .01$.

Figure 1. Underground garbage container with warning label (enlarged)



Note. The warning label on the lid says: ‘Call 14 010 for disposed items, inquiries, and complaints. Fine as of 115 euro for illegal disposal’.

Figure 2. Map of the area



Note. The black dots denote the location of the 41 garbage containers.

Figure 3. *Defining accumulation of garbage*

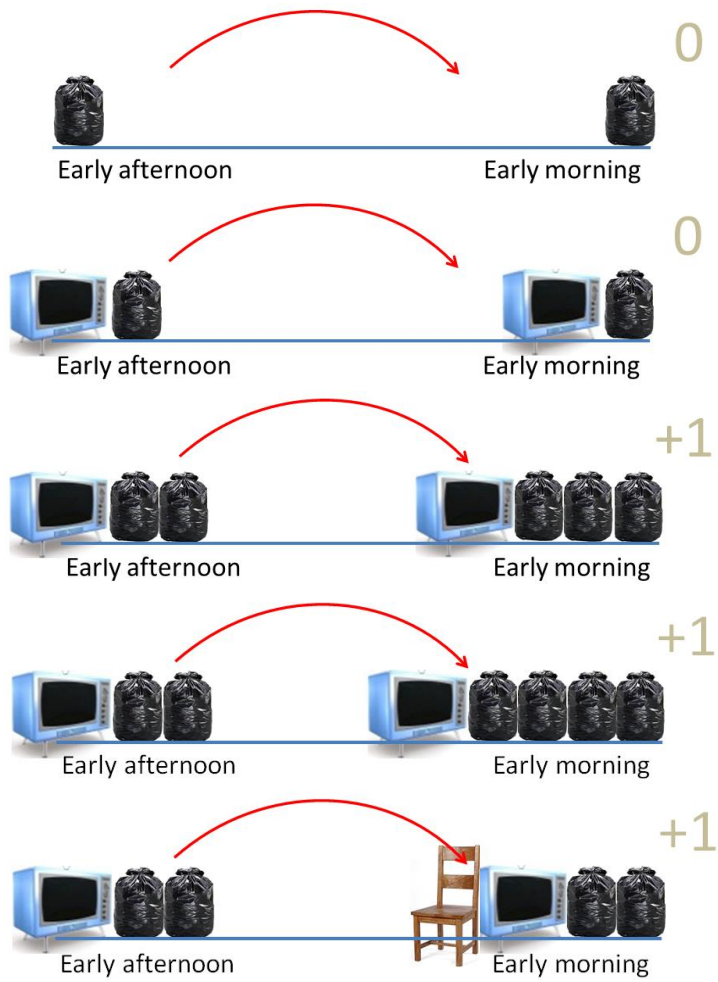
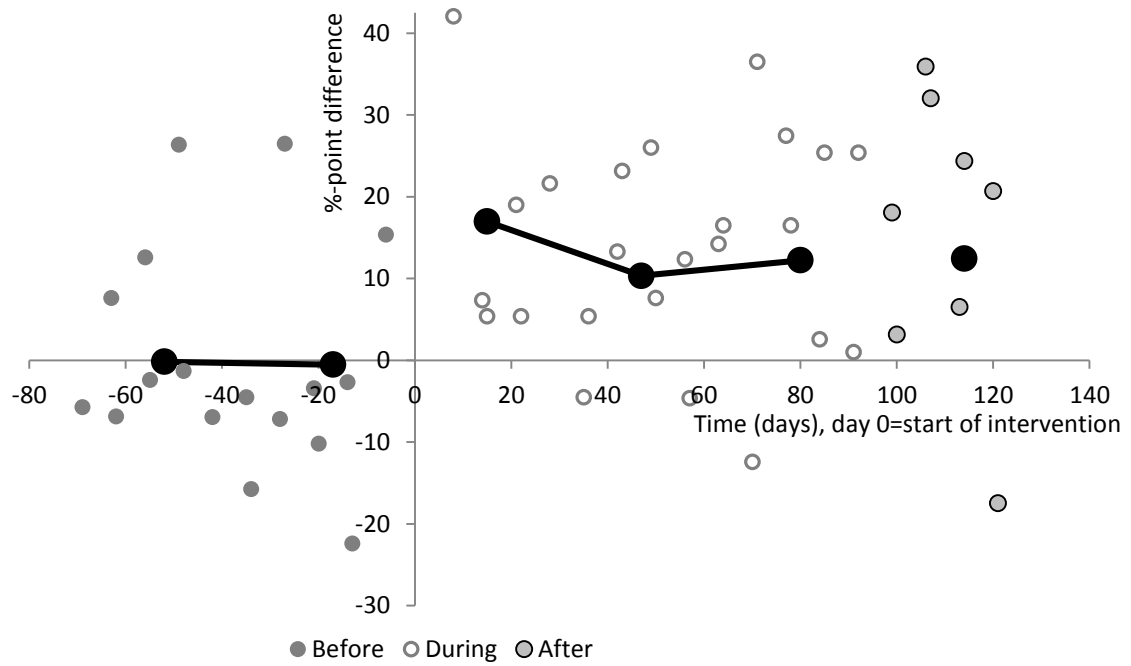
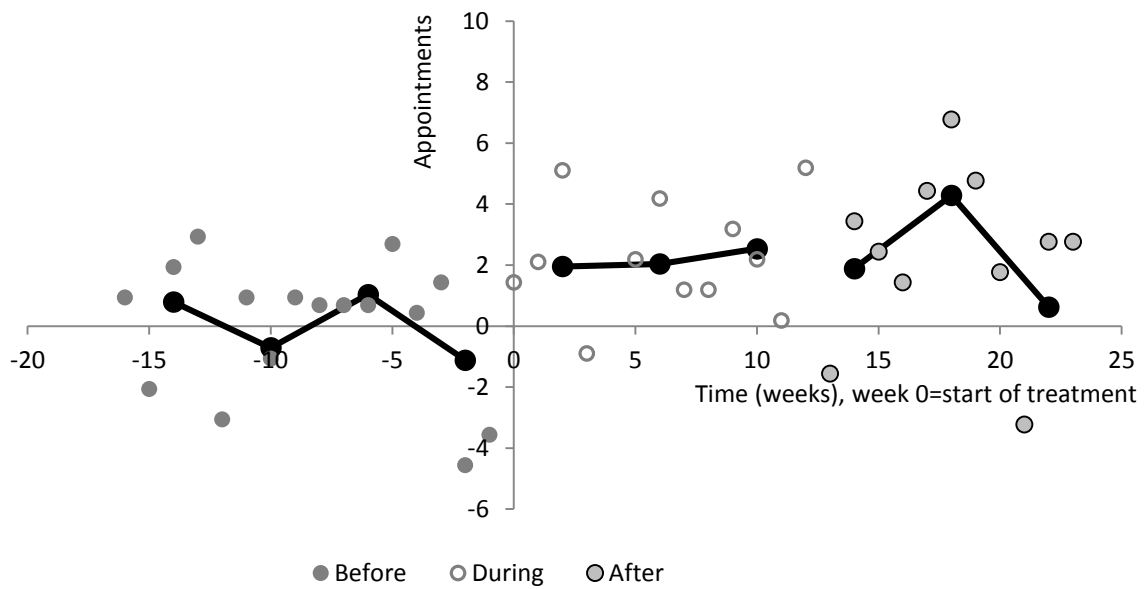


Figure 4. *Difference in accumulation of garbage between treatment and control area, by day and by month*



Note. Graph plots coefficients α_t from estimation equation (1). Daily coefficients are in light gray; separately estimated monthly coefficients are in black.

Figure 5. *Difference in number of appointments for pick-up of discarded household items between treatment and control area, by week and by month*



Note. Graph plots coefficients α_t from estimation equation (3). Weekly coefficients are in light gray; separately estimated monthly coefficients are in black.

Appendix

The cleanliness of the 41 container locations that are part of our study differs greatly. As an additional exercise, we analyzed the relationship between characteristics of container locations and accumulation of garbage. Due to the non-experimental nature of the variation that we use, the evidence is correlational rather than causal. Dependent variable is the average occurrence of accumulation of garbage next to the container from early afternoon until early morning next day during the baseline period.

Environmental characteristics that were suggested by service workers included the visibility of a container location and the presence of shops. Lack of visibility allows residents not to be seen when illegally disposing garbage next to the container. Lack of visibility is coded as one if a location is at a street corner rather than in the middle of a street and if it faces a blind wall rather than windows of homes and zero otherwise. The presence of shops was thought to be relevant since shop owners often have relatively big garbage items and may be using the containers rather than their own garbage trolleys. In addition to these environmental characteristics, we also make a distinction between types of containers. First, some locations have only fully underground containers while other locations have containers that are only half underground. The lid of a half underground container was larger and easier to open than the lid of a fully underground container, which may affect the likelihood of illegal disposal. Second, some locations had both types of containers during the period that we observed them. This was the result of a policy of slowly turning all containers into fully underground containers. We label these locations as having ‘combi-type’ containers.

Table A1 presents the results. We use three alternative outcome measures: accumulation of all garbage (column 1); accumulation of garbage bags (column 2); accumulation of household

items (column 3). We find a positive coefficient for a location being at a street corner or facing a blind wall for all outcome measures, but the coefficient is never statistically significantly different from zero. The presence of shops at a container location is strongly related to greater accumulation of all types of garbage. The coefficient is statistically significant at either the 90 or 95 percent level. Given an average accumulation of all garbage of 0.224 across the 41 container locations, accumulation is more than 50 percent higher if shops are nearby. We also find a strong relation between the type of containers and accumulation of garbage, particularly if a location combined the two types of containers.

Table A1. *Relation between illegal garbage disposal and characteristics of container locations*

Dependent variable: average occurrence of accumulation of garbage next to container from early afternoon until early morning next day	(1) All garbage	(2) Garbage bags	(3) Household items
Street corner and/or blind wall	0.048 (0.052)	0.034 (0.033)	0.048 (0.052)
Shops	0.144 (0.077)*	0.102 (0.049)**	0.143 (0.077)*
Underground container	0.109 (0.065)*	0.071 (0.041)*	0.109 (0.065)*
Combi-type container	0.346 (0.126)***	0.222 (0.079)***	0.373 (0.126)***
Number of observations	41	41	41

Note. Based on data during baseline period, excluding days at which a container location was emptied. Observations by container location. Standard errors between parentheses. * $p < .10$; ** $p < .05$; *** $p < .01$.