Working Paper in Economics No. 783

Does teaching school children about recycling reduce household waste?

Claes Ek and Magnus Söderberg

Department of Economics, February 2020

ISSN 1403-2473 (Print)
ISSN 1403-2465 (Online)

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Does teaching school children about recycling reduce household waste?

Claes Ek[a], Magnus Söderberg[b,c]

[a]*Department of Economics, University of Gothenburg, P.O. Box 640, SE-405 30 Gothenburg, Sweden*
[b]*Department of Sociology, Environmental and Business Economics, University of Southern Denmark, Niels Bohrs Vej 9-10, DK-6700 Esbjerg, Denmark*
[c]*Ratio Institute, P.O. Box 3203, SE-103 64 Stockholm, Sweden*

Abstract

Reduced waste generation is a prioritized environmental policy objective in the EU as well as worldwide. We perform a randomized controlled trial in Sweden with school children aged 10-16 to evaluate an intervention designed to reduce household waste, Environmental Education Programs (EEP). Crucially, we are able to examine the causal effect of a waste-themed EEP on the actual waste generated in households where a child was treated. This is done by coupling the addresses of participating students with high-resolution address-level panel data on collected waste amounts, supplied by municipal waste authorities. Our design allows identification of the differential effect of the EEP on waste generation in treated versus control households. We find no significant evidence that the intervention had any effect on waste generation. There is also no indication that this null result is due to interference between treated and control students.

Keywords: Field experiments, Environmental Education Programs, household waste, intergenerational learning
JEL classification: D13, I21, Q53

1 Introduction

For at least 40 years, social scientists have considered the problem of promoting pro-environmental behavior. One widely used approach for stimulating behavioral change is information provision, consistent with knowledge-deficit models that assume people lack knowledge about why and how to reduce their environmental impacts (Abrahamse and Matthies, 2012). Given that this assumption is particularly likely to hold for children, providing in-

☆This version: 30 January 2020.
Email address: claes.ek@economics.gu.se (Claes Ek)

formation to school pupils in the classroom may be especially promising. Interest in such Environmental Education Programs (EEP) arose in the 1990s (e.g. Hungerford and Volk, 1990), e.g. with Campbell et al. (1999) finding that EEPs increase children's knowledge about the environment and lead to stronger pro-environmental attitudes. In addition, the literature on 'intergenerational learning' (Ballantyne et al., 2001a,b; Vaughan et al., 2003; Duvall and Zint, 2007; Gentina and Muratore, 2012; Lawson et al., 2018) shows that children share new knowledge about the environment with other family members and moreover suggests that they can influence their parents' attitudes and behavior. This implies that policy makers may be able to use EEPs as a 'back door' toward influencing the behavior of entire households.

There are, however, several weaknesses with the existing literature on EEPs. First, with only a few exceptions (Boudet et al., 2016; Lawson et al., 2019), studies are of a non-experimental nature, with results that cannot necessarily be given a causal interpretation; identification typically relies either on comparisons across non-randomly allocated treatment and control groups, or on uncontrolled before-after comparisons. Second, while many studies examine attitudes and knowledge (e.g. Leeming et al., 1997; Lawson et al., 2019) or self-reported behaviors (Grodzinska-Jurczak et al., 2003; Damerell et al., 2013; Boudet et al., 2016), measuring the effect of EEPs in terms of verified third-party measurements of pro-environmental behavior is very rare. As far as we are aware, no previous study has addressed both of these potential shortcomings. The aim of the present paper is to do so: we use a Randomized Controlled Trial (RCT) to investigate whether EEPs can reduce actual household waste.[1]

We develop and implement a novel Environmental Education Program for school pupils aged 10-16 that is meant to encourage the entire household to reduce their waste. The program covers 32 classes in two municipalities in western Sweden, Varberg and Falkenberg; we randomly allocate students either to the EEP or to a placebo intervention on meteorology. Besides a classroom segment, we include a home assignment where students weigh the waste generated in their own household. The latter aspect of the EEP follows Bulkeley and Gregson (2009)'s recommendation that "municipal waste policy needs a far closer engagement with the household" and that, in terms of time and context, interventions should be positioned as close to waste-generating activities as possible. In addition, many students likely need to involve other household members to complete the assignment, and we believe our program

[1]We preregistered the study at the AEA RCT registry, ID R-0003300.

is therefore relatively well placed to affect the waste behavior of the entire household.

Addresses reported by the students are then coupled with high-resolution data on actual waste amounts collected from each address by VIVAB, the municipal waste company in Varberg and Falkenberg. This highly unique data set, together with our RCT approach, allows us to investigate if households with children that were exposed to the EEP have changed their waste behaviour relative to households with children that were not exposed, thus identifying the causal impact of the EEP.

The only other study we are aware of that evaluates the effect of a waste-themed EEP on actual waste amounts is that of Maddox et al. (2011); indeed, we are aware of no other similar study that uses verified (as opposed to self-reported) actions as its outcome variable, for any pro-environmental behavior. Nevertheless, in their paper, waste amounts are not measured at the address level, but are aggregated by collection zones roughly corresponding to school catchment areas. Furthermore, the authors do not compare the before-after changes measures in treated areas to some control group which was not exposed. Thus, we believe we are able to provide a significant contribution in addition to their study.

Reducing waste is a prioritized objective for policy since this lowers human exposure to harmful substances as well as the need for raw materials and resource extraction (Eurostat, 2015). Similar arguments appear in discussions about 'circular economy systems' (European Commission, 2014). The UN 2030 Agenda for Sustainable Development thus includes targets to 'substantially reduce waste generation through prevention, reduction, recycling and reuse' as well as to halve global food waste (Goal 12). Similarly, the original 2008 EU Waste Framework Directive included a target that 50% of household waste should be recycled by 2020 in each member state; more stringent targets, including one of 65% recycling by 2035, was added in a 2018 revision. A recent report (European Commission, 2018) found that, while the EU-wide recycling rate stood at 46% in 2016, 14 member states were at risk of missing the 2020 target.

In the EU, most policies aiming to reduce household waste are municipality-specific or otherwise local in scope, involving mainly (i) less frequent collection of household waste to increase pressure on recycling efforts, (ii) offering curbside collection of additional waste streams, (iii) introduction of 'pay-as-you-throw' marginal-cost incentives, e.g. through weight-based waste fees (Bucciol et al., 2015), (iv) provision of clearer and better information, (v) establishment of markets for recycled products such as second-hand clothing, (vi) establishment of waste recycling centres.[2] However, given that even the EU targets for 2020 are in

[2]Policy makers have also made more general efforts to raise public awareness on waste issues. Examples

danger of not being met, there is both need and scope for additional, innovative ways to reduce household waste.

The structure of the paper is as follows. In the next section, we describe the design of our EEP intervention. Section 3 outlines our data sets, identification strategy, and statistical model. Section 4 presents our main results as well as various extensions, while section 5 concludes the paper.

2 Experimental design

The study includes 13 schools and 32 classes averaging 20-25 student per class, summing to roughly 700-750 students. Schools were recruited into the study by outreach with the Education Administrations in Varberg and Falkenberg, as well as emailing teachers directly. The schools and/or teachers included in the study are those that volunteered to participate. The study is limited to the municipalities Varberg and Falkenberg because these have implemented a two-part waste tariff, where the fee paid by households depends in part on the amount thrown (in kilograms); this is why address-level data on waste is available.

Each participating class was visited twice. On the first visit, all present students within a class were randomized into a treatment and a placebo control condition. Randomization was done by manually shuffling assignment cards marked A or B. We stratified treatment assignment by class: for even numbers of students in a class, equal numbers of A and B cards were drawn; for odd numbers, one additional A or B card was added, in an alternating pattern across classes. All school visits were conducted by the same two experimenters, who were randomly assigned to either group A or B (treatment or control) by means of a coin flip.

The treatment condition (i.e. the EEP) was waste-oriented, and since our visits were viewed by both teachers and students as part of the natural science curriculum, we chose to focus the placebo condition on meteorology. Within each condition, students were first given a home assignment. Treated students were asked to measure, each day over a period of one week, the amount of waste generated in their household or the household they were visiting on that day. Experiments lent each student a hand scales for this purpose. Control students faced a similar task of measuring the outdoor temperature and other weather factors for a consecutive seven days.

include the EU-wide 'European Week for Waste Reduction', supported by the LIFE+ Programme and several national authorities; the 'Pre-waste' project, supported by the European Regional Development Fund; and web sites like `lovefoodhatewaste.com` and `thinkeatsave.org`.

Between one and three weeks after the conclusion of these assignments, each class was revisited by the experimenters. The second visit typically engaged all students present, regardless of whether or not they participated during the initial visit. In this second session, treated students listened to a brief lecture on waste and the environment, participated in a subsequent experimenter-led group discussion, and finally played an educational game which involved answering quiz questions on waste and sorting custom-made playing cards representing different waste fractions. Control students instead listened to a lecture on geographical variation in temperature and rainfall, and participated in a group discussion and quiz on those themes.

As part of the home assignment, each student filled in a form provided by the experimenters, which formed the main data source from the intervention itself.[3] One of the fields specified the address where, for each day, the assignment was carried out. We combine these addresses with household-level waste data from VIVAB, the municipal authority in charge of waste management in both Varberg and Falkenberg. This allows estimation of the differential effect of treatment on waste amounts in the households where a student was treated. The form also collects self-reported information on social networks within classes, allowing us to control for social interaction as a mechanism for spillovers in behavior between treatment and control.[4]

3 Data and empirical strategy

The VIVAB data includes collected waste weights from all addresses stated in the forms collected from students. The raw data contains one line per bin-specific collection event, with waste bins falling into three categories: food, household, and unsorted waste, where a household typically either has food and household bins, or a single unsorted-waste bin.[5] We recode weight variables associated with household and unsorted waste as a single residual-waste variable. For each address in the data set, we then sum waste weights (in kilograms) separately for residual and food waste across two-week periods, reflecting the fact that waste collection is biweekly for most households. Finally, we divide all weights by the number of

[3] Experimental materials (scripts, form templates, and game rules) are provided in Appendix A.

[4] Students also provided information on a large number of other variables, including the age of members in the household, whether they owned a pet, what they had for dinner during the assignment week, etc. This information will not be used in our analysis.

[5] We drop all collection events with negative weights or non-identifiable waste types. VIVAB also flags all events where e.g. a bin was not placed curbside and thus could not be collected. We will take such incident flags into account only if no strictly positive weight is reported on that data line; Table B.1 in Appendix B lists how different incidents are coded, if applicable.

household members as given by the Swedish Tax Authority.

The result is a panel data structure with one pair of weight observations (residual and food waste) per participating address and two-week interval. Each two-week period runs from Monday to Sunday the following week, covering the period between 7 May 2018 and 17 March 2019, with the first intervention occurring on 10 September 2018 (Table 1). This implies 9 baseline (untreated) periods, 6 post-treatment periods, and 7 periods where, at least for part of the period, some classes have been exposed to the intervention and some have not.

We exclude a number of addresses from the study. First, we drop all addresses that are not reported more than once (out of seven) on any form, implying that all forms where no address is reported at least twice are disregarded.[6] Second, we exclude addresses of apartment buildings, since treated households will generate only a minor fraction of the total waste collected from such addresses. Third, we are unable to match a total of 42 stated addresses with waste data and household-size data. Fourth, we exclude households that have more than 80% missing or zero observations (across all two-week periods in Table 1) for both residual and food waste.[7] Fifth, we also drop outlier households with a mean residual- or food-waste weight above 15 kg/person. Finally, single outlier observations where residual- or food-waste weight is above 60 kgs/person are likewise excluded. After these operations, the total number of addresses included in our data set is 351, with 181 in the treatment group and 170 in the control group.[8] Testing for differential attrition between treatment and control, the share of addresses in the treatment group (51.6%) is not significantly different from 1/2 ($p = 0.557$).[9]

[6] There is one exception to this rule: in some cases, a form contains a weight but no address for some date(s). We disregard these fields unless only one address appears on the form in question. All measurements on that form are then assumed to have occurred at that address, even if the address is only stated once.

[7] In our pre-analysis plan, this proportion was 90%. However, after data collection we found that 20 adresses have no observations during the intervention period, and revised the cutoff to exclude these households.

[8] After the conclusion of the intervention, teachers sent the forms collected from students to a third party for purposes of ensuring anonymity. However, forms from a single school (Söderskolan) were lost by the postal service and was never received by the research team. Instead, we went back to the school in question and re-sampled the treatment status, home addresses, and social networks of participating students. These addresses are included in our main analysis, but the results are robust to excluding them.

[9] In particular, the treatment share within excluded apartment-block addresses is also not significantly different from 1/2 ($p = 0.571$). A separate point is that detailed information on questionnaire nonresponse rates (as opposed to exclusion rates) is limited to the latter 21 classes, where we noted down the division of present students into groups A and B. However, the split among present students is also very close to 50-50 (239 vs. 241 subjects), and the same method of treatment allocation was used in the first 11 classes. Running the equality-of-proportions test on the final data set and addresses only in the latter 21 classes also fails to reject the null hypothesis ($p = 0.200$).

Table 1: *Study timeline*

Period	Start date	End date	Classes visited	Treated addresses	Control addresses
2018					
1-9	7 May	9 September	0	-	-
10	10 September	23 September	3	17	18
11	24 September	7 October	1	6	7
12	8 October	21 October	5	26 (25)	33
13	22 October	4 November	6	19	17 (16)
14	5 November	18 November	3	20	17 (14)
15	19 November	2 December	12	63 (58)	47 (46)
16	3 December	16 December	2	30 (28)	31 (29)
17	17 December	30 December	0	-	-
2019					
18-22	31 December 2018	17 March 2019	0	-	-

Table lists the time periods included in the experiment. The column 'Classes visited' lists the number of classes visited for the first time within each period, and the final two columns report the number of addresses associated with those classes, arranged by treatment arm. Numbers of modal addresses are given in parentheses.

Our main regression specification has the following standard difference-in-differences structure

$$y_{it} = \alpha_i + \lambda_t + \beta T_{it} + \epsilon_{it} \tag{1}$$

which includes address and two-week period fixed effects, respectively. For inference, we use robust standard errors clustered at the address level.[10] The treatment variable T_{it} is always equal to zero for untreated households (see below). For treated households, $T_{it} = 1$ in all periods subsequent to the period of the first class visit by the experimenters. In the two-week period including the first visit, T_{it} is equal to the share of week days in the period occurring after the visit. Thus, for instance, if a school was visited on Thursday of the second week,

[10] All regressions calculate robust standard errors clustered at the address level; we do not cluster at the class level because the number of classes (32) is low. We also do not include fixed effects for class, teacher, school or experimenter, because these are all typically invariant within addresses and thus captured by the address fixed effects.

$T_{it} = 0.9$. In all other periods, $T_{it} = 0$.[11]

Most of the regressions reported below restrict the sample to the set of modal addresses given on each participant's form. We take the conservative approach of determining modality before excluding addresses. For instance, a participating student may report an apartment-block address four times and a single-family housing address twice; thus, after removing the former address, the latter would be considered nonmodal. There is only one tie, where a student reported two addresses three times each, and we use both addresses in this case. All in all, the number of modal addresses in the final data set is 336. Addresses that are modal for more than one student are coded as treated as long as at least one participant is treated there, and the timing of treatment is taken as that of the first student treated at that address. In addition, friends are pooled by address.[12]

In Appendix C, we report the results of a power calculation based on the difference-in-difference estimator and using a within-household variance parameter estimated from historical data. The fact that we are able to repeatedly measure accurate waste weights over a long time period improves precision greatly, and we find that our minimum detectable effect at 80% power is about 8.7% of a standard deviation. This implies that our main analysis should be able to detect even quite subtle treatment effects.

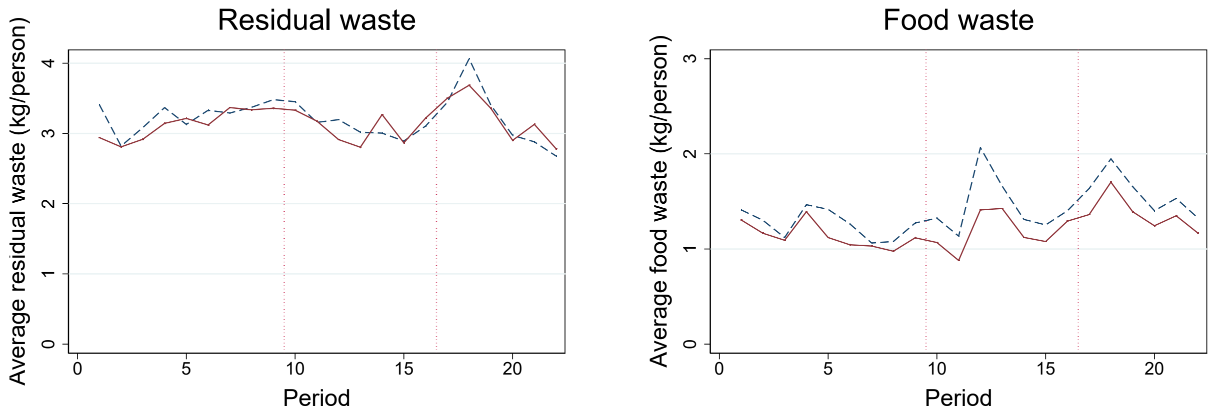
4 Results

We begin by checking for sample balance in our outcome variables across all baseline periods 1-9 prior to the first class visit.[13] As this effectively checks for coinciding pre-treatment trends, it is somewhat similar to the 'parallel-trend' assumption employed in non-experimental difference-in-difference studies. Figure 1 displays treatment-arm averages across the entire study period. If randomization into treatment was successful, graphs for treatment and control should coincide, or nearly so, prior to the intervention period; if there is also a substantial treatment effect, they should begin to diverge in subsequent periods. In the left panel, which graphs residual-waste weights, we see that trends are indeed very close before the intervention period. Differences are also minor in later periods, suggesting a zero or quite small treatment effect.

[11]Our main results are practically unchanged if each period-class observation where a class was visited for the first time is excluded from the data.

[12]In exactly one case, an address is modal for one student but not for another; however, neither of them report any friends, and the student for which the address is nonmodal belongs to the control group.

[13]Testing across all such periods, the sample is balanced with respect to the number of residents at each address ($\beta = 0.067$, $p = 0.525$).



Note: This figure depicts household average weights per person of residual and food waste, respectively. The solid (dashed) line depicts treatment (control) group averages. The vertical dotted lines define the intervention period, which lasted from two-week period 10 to 16: the lines are thus placed between periods 9 and 10, and between periods 16 and 17.

Figure 1: Trends in waste weights before, during, and after the intervention period

In the right hand panel, treatment-group averages (solid line) are generally below those of the control group (dashed line), but differences seem to appear prior to the intervention period, suggesting randomization was unsuccessful in eliminating all average differences between treatment and control.[14] Nevertheless, given the panel structure of the data, this is problematic for identification only if the treatment-control differences not erased through randomization are time-variant. Recall that difference-in-differences effectively conditions on baseline outcome differences across treatment arms, so any discrepancy that is constant over time (and does not interact with treatment) will cancel out in estimation. When we test whether the pre-treatment (period-by-period) differences in food waste across treatment arms reported in table D.1 are all equally large, the test statistic is nonsignificant ($F = 0.381$, $p = 0.931$). Thus, pretreatment trends appear largely parallel, suggesting that difference-in-difference will still yield valid estimates.

Our main treatment-effect estimates are reported in table 2. All pairs of regressions (residual and food waste, respectively) use critical values subjected to an approximate adjustment for multiple hypothesis testing, the D/AP method described in Sankoh et al (1997), with $K = 2$. The correlation between address-level demeaned waste weights (corresponding to outcome variables in our fixed-effect regressions) is 0.170, which gives new critical values

[14] In table D.1 of Appendix C, we perform statistical tests of balance by regressing pre-treatment waste amounts on (eventual) treatment status. For food waste, differences between treatment and control addresses are only marginally significant when such regressions are run separately for each period 1-9. However, all coefficients are negative, and when we pool all baseline periods, treatment-control differences are highly significant.

Table 2: *Treatment effect estimates*

	Residual waste		Food waste	
Variable	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.
Treatment	0.074	0.458	-0.021	0.725
	(0.097)		(0.062)	
Address FE	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes
Observations	7,231	7,231	7,236	7,236
Addresses	336	336	336	336
R-squared (within)	0.018	–	0.030	–

Table presents our main regression estimates for the effect of treatment. Robust standard errors clustered at the address level reported in parentheses. Columns 'p, rand. inf.' give randomization-t p values, i.e. the share of re-randomized treatment vectors (out of 1000) that yield larger t statistics than those implied by the regression results of the preceding column. From our adjustment for multiple hypothesis testing, approximately * $p < 0.058$, ** $p < 0.028$, *** $p < 0.006$.

as $\alpha_k = 1 - (1 - \alpha)^{2^{-0.830}}$, where α is the original critical value and α_k is the adjusted one. Thus, in these regressions, approximately * $p < 0.058$, ** $p < 0.028$, and *** $p < 0.006$.

For robustness purposes, our tables supplement standard regression output by also reporting exact two-tailed p values from a randomization inference procedure involving the following steps. First, we re-randomize the treatment vector 1,000 times for the set of relevant addresses, with both the number of treated addresses and the distribution of initial class visits identical to our actual data set.[15] Thus, for instance, if x addresses were first visited in period t, x addresses will be coded as such at t in each re-randomization as well; though these addresses are, of course, very unlikely to be the same ones that were visited in actuality. Next, we run the regressions reported in the table using each re-randomized T_{it} vector. The randomization p value is then equal to the share of regression t statistics (out of 1,000) with an absolute value larger than that of the t statistic obtained in our actual

[15] This approach is suggested by MacKinnon and Webb (2019). Our re-randomization ignores the details of addresses potentially being modal for multiple students, etc. We also do not stratify re-randomized treatment by class, despite doing so during the initial class visits; this is because subsequent data restrictions (e.g. dropping all apartment-block addresses) imply that treatment is in any case no longer 50-50 within each class in our final data set.

regression(s). Young (2019) calls this the 'randomization-t' approach, as the comparison between actual and re-randomized regression output is made for test statistics rather than coefficient estimates.

In table 2, which uses only modal addresses, we find that both standard regression inference and randomization inference fail to reject the null of no effect, both for residual and food waste.[16] In table D.2, we show that the same is true when we drop all addresses that are modal for more than one student, thus restricting ourselves to one-to-one correspondences between students and addresses (with the single exception of the tie mentioned in section 3).

4.1. Extensions

It is possible that the estimates in table 2 mask treatment heterogeneity in the sense that not all students may have engaged to the same extent with the home assignment and other aspects of the intervention. Furthermore, those living at addresses stated by students were likely exposed to the intervention to differing degrees. For example, suppose that a treated student stayed long enough at a relative's house during the assignment week for it to be considered modal for that student. This address will then be flagged as treated despite not being the home address of the student.

To examine the intensity of the intervention across students, we moderate the treatment variable by interacting it with a variable capturing 'general engagement': the number of days out of seven where the (treated) subject associated with an address reported weighing waste at *any* address. For all addresses that are associated with at least two treated students, we use the largest of the number of days reported by the students. In a second analysis, we interact our treatment variable with 'specific engagement': the number of days out of seven where the (treated) subject associated with an address reported weighing waste at *that* address. Again, for addresses associated with more than one treated student, we use the maximum of these numbers across the relevant set of students. In both analyses, the moderated treatment variable is used in place of the uninteracted variable T_{it}.

The results are reported in table 3; note that these regressions include all feasible addresses, not just modal ones.[17] Again we report p values for randomization inference. Note

[16]Note that randomization inference is not strictly comparable with regression-based inference, as it tests the sharp null that each address-level treatment effect is zero rather than the non-sharp null that the average treatment effect is zero.

[17]Both regressions exclude a single address that occurs for one control student more than once and is reported exactly once by a treated student. In our pre-analysis plan, we proposed running the general-engagement regression only on modal addresses and with this address included; doing so produces virtually identical results. Note also that neither regression in table 3 includes Söderskolan, where no information

Table 3: *Treatment effects by engagement with the task*

	Residual waste		Food waste		Residual waste		Food waste	
Variable	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.
General engagement	0.015	0.397	-0.006	0.484				
	(0.017)		(0.009)					
Specific engagement					0.011	0.467	-0.006	0.487
					(0.017)		(0.010)	
Address FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,048	7,048	7,055	7,055	7,048	7,048	7,055	7,055
Addresses	327	327	327	327	327	327	327	327
R-squared (within)	0.020	–	0.035	–	0.021	–	0.035	–

Table presents regression estimates for the effect of treatment interacted with two measures of subject engagement with the assignment: (i) *Specific engagement*, the maximum number of days (out of seven) where some treated subject associated with an address reported weighing waste at the address in question, and (ii) *General engagement*, the maximum number of days (out of seven) where some treated subject associated with an address reported weighing waste at any address. Robust standard errors clustered at the address level reported in parentheses. Columns 'p, rand. inf.' give randomization-t p values, i.e. the share of re-randomized treatment vectors (out of 1000) that yield larger t statistics than those implied by the regression results of the preceding column. From our adjustment for multiple hypothesis testing, approximately * $p < 0.058$, ** $p < 0.028$, *** $p < 0.006$.

that re-randomization applies only to T_{it}, thus holding the interacted engagement variables constant across draws; this amounts to a possibly tenuous assumption that engagement is independent of treatment. When, for simplicity, we restrict the sample to the 315 addresses associated with exactly one student and regress the total number of days reported on treatment-group status, we find no significant differences ($\beta = -0.076$, regression $p = 0.612$); repeating the exercise for the number of days a particular address is reported, differences are marginally significant ($\beta = -0.163$, regression $p = 0.064$). In any case, both regression and randomization inference imply nonsignficant results in table 3: we find no evidence that the effect of the intervention is moderated by either measure of its address-level intensity.[18]

Next, we perform a simple test to check for potential dynamics in the impact of the intervention. It is common to see a pattern of 'action and backsliding' as a result of behavioral interventions, with larger immediate than long-term effects (e.g. Allcott and Rogers, 2014). It is likewise plausible that any effect of our intervention would attenuate over time, as treated students progress to other parts of the curriculum. We check for such dynamics by splitting our treatment variable into two components: (i) an effect active during the first three periods of treatment, including the period of the initial class visit (where the variable is equal to the corresponding value of T_{it}; it is equal to one in the two following periods, and again zero afterwards), and (ii) an effect that is active in all subsequent periods, as captured by a binary variable.

The results are given in the first four columns of table 4. Although the coefficient for residual waste is larger in absolute value in the initial three periods than in later ones, the opposite is true for food waste, and both regression and randomization p values are again quite large. Thus, there is little to suggest that our EEP had either a short-term or a long-term effect on waste.

Another potentially interesting subgroup analysis is to check whether treatment effects differed by the average amount of waste generated prior to the intervention period. For example, it is possible that only households where waste generation was initially high responded to the intervention. Our test is simple, involving only a binary split according to whether a given address was above or below the across-address median for waste generated throughout all baseline periods 1-9. We then regress residual (food) waste weights on T_{it} as well as T_{it} interacted with a dummy for whether an address had above-median residual

beyond treatment status and subejct home addresses is known; recall footnote 8.

[18] In a non-preregistered analysis, we also attempt to proxy for the household negotiation power of treated children by adding an interaction between T_{it} and household size to regression (1). The coefficient for this variable is insignificant, both for residual waste (regression $p = 0.530$) and food waste (regression $p = 0.353$).

Table 4: *Dynamics and heterogeneity of treatment effects*

	Dynamics				Baseline weights			
	Residual waste		Food waste		Residual waste		Food waste	
Variable	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.
First three periods	0.093	0.400	-0.016	0.857				
	(0.109)		(0.075)					
Later periods	0.064	0.548	-0.024	0.689				
	(0.110)		(0.067)					
Below median					0.197*	0.313	-0.077	0.468
					(0.100)		(0.081)	
Above median					-0.276*	0.568	0.091	0.653
					(0.144)		(0.090)	
Address FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,231	7,231	7,236	7,236	6,912	6,912	6,917	6,917
Addresses	336	336	336	336	321	321	321	321
R-squared (within)	0.018	–	0.030	–	0.020	–	0.031	–

Columns labeled 'Dynamics' divide the treatment variable into (i) effects present during the first three two-week periods of the intervention, including the period of the initial class visit; and (ii) effects of treatment in all later periods. Columns 'Baseline weights' instead interact treatment with a dummy for whether a given address had above-median or below-median average residual or food waste weights across the set of baseline periods 1-10. Robust standard errors clustered at the address level reported in parentheses. Columns 'p, rand. inf.' give randomization-t p values, i.e. the share of re-randomized treatment vectors (out of 1000) that yield larger t statistics than those implied by the regression results of the preceding column. From our adjustment for multiple hypothesis testing, approximately * $p < 0.058$, ** $p < 0.028$, *** $p < 0.006$.

(food) waste weights in the baseline. The uninteracted treatment effect will thus implicitly represent treatment effects among below-median addresses. For simplicity, we include only addresses that are modal for exactly one student in these regressions.

The final four columns of table 4 presents the results of this exercise, seemingly providing weak evidence that above-median residual waste weights have dropped while below-median weights have increased as a result of treatment; both of these effects are marginally signficant. However, caution is warranted in interpreting these estimates, since they are also qualitatively what one would expect to see if treatment effects were exactly zero for all addresses but there was a tendency for waste weights to regress to the mean over time.[19] Indeed, the large randomization-inference p values we find lend strong support to this idea. For example, among the set of above-median addresses, the distribution of above-median interacted-treatment t statistics induced by the randomization distribution is centred not around zero but around roughly $t = -2$, and below-median t statistics similarly average approximately $t = 1.5$. This implies that our marginally significant regression outcomes are likely driven by regression to the mean rather than treatment assignment. A similar but less pronounced pattern appears for food waste.

4.2. Interference between treatment and control

A final concern is that, perhaps especially since randomization was performed within rather than between classes, there may have been interference between treatment and control. For example, it is highly likely that students discussed their respective assignments with each other, with the result that control students were aware that treated students were weighing household waste. If such knowledge led to different waste behavior at control addresses, the 'stable unit treatment value assumption' (Rubin, 1980) that potential outcomes do not depend on the treatment status of others will have been violated, and our estimates will be biased. Typically the bias is thought to drive estimates toward zero, i.e. in the direction of smaller differences between the two groups; thus, it is possible that such spillovers are the cause of our nonsignificant main results.

Under the assumption that any interference operates between friends, we may perform a test of the null hypothesis that there were no spillovers. We follow the procedure outlined by Athey et al. (2018). This is essentially a randomization-inference approach which constructs a sharp null of no interference by partioning the set of experimental units (addresses) into

[19] In an OLS setting, one might control for such patterns by also including an uninteracted above-median dummy, capturing regression to the mean at control addresses. Such a dummy will be constant within addresses, however, and so cannot be estimated in our fixed-effects regressions.

two groups, which the authors term the focal and the auxiliary groups. Treatment status is then re-randomized for the auxiliary units while held constant for the focal units; for each new treatment vector, we estimate a regression with an interference variable (number of treated friends), only using the focal units. Thus, for the focal group, treatment status is held constant across randomizations while the number of treated friends will vary. This permits calculation of exact randomization p values for the interference variable.

As the first step in the procedure, we construct a network matrix $\mathbf{G}$. In this matrix, element G_{ij} is initially equal to one if a student associated with address i reports a friendship link with a student at address j; it is zero otherwise. Taking the data as is, $\mathbf{G}$ is not symmetric, which the method requires: on the contrary, it is quite common for friendship statements to not be reciprocated. Thus, we make $\mathbf{G}$ symmetric by a 'maximal' approach where element G_{ij} is recoded as one whenever $G_{ji} = 1$. As a robustness test, we also take a 'minimal' approach where G_{ij} is recoded as zero whenever $G_{ji} = 0$. We then trim $\mathbf{G}$, retaining only columns and rows associated with addresses that have at least one friendship link to another address. The resulting dimension of the network matrix, equal to the number of such addresses, is 258 (158) using the maximal (minimal) approach.

Since, as described above, estimation exploits links between focal and auxiliary units, we next select focal addresses using a greedy algorithm to maximize the number of such links. This algorithm, outlined in Athey et al. (2018), uses the following procedure. Initially, all addresses in the (maximal or minimal) network are allocated to the auxiliary group. Then, for each address, we calculate the number of focal-auxiliary links that would be added if that address were to be moved to the focal group. These numbers are given by $\Delta_{N,i} = \mathbf{G}\left(\mathbf{1} - 2\mathbf{F}\right)$, where $\mathbf{1}$ is a vector of ones, and vector $\mathbf{F}$ denotes focal-group inclusion: element $F_i = 1$ if address i has already been added to the focal group, and is zero otherwise. Intuitively, as an address is added to the focal group, all prior links from that address to the focal group are lost, and thus the total number of links involving the address need to be at least twice as many for the number of focal-auxiliary links in the population to increase. The address associated with the largest element of $\Delta_{N,i}$ is then added to the focal group, and the process is repeated until the number of focal-auxiliary links in the population cannot be increased by adding more addresses to the focal group.[20]

After the greedy algorithm has concluded, we perform the re-randomization. For each

[20] While this algorithm is described in Athey et al. (2018), the authors use a somewhat more complex weighted approach in actual simulations. Using the simpler algorithm instead increases the total number of focal-auxiliary links found by over 50%.

draw, we run a version of the difference-in-differences regression (1) where we also add a variable summing T_{it} across all friendship-linked addresses, thus capturing approximately the number of treated friends. To produce more precise estimates of fixed effects as well as the main treatment effect, we include all addresses that lack friendship links in the focal group in these regressions. Note that the minimal network matrix leaves more addresses without friendship links, which has the net effect of increasing the number of focal addresses.

Table 5 gives the results, presented separately for the maximal and the minimal network matrices. Each randomization p value is based on 1,000 draws from the auxiliary-specific randomization distribution. Since focal addresses see randomization-induced variation only in their number of treated friends and not in their own treatment status, the table does not report randomization p values for the treatment variable. All p values relating to the interference variable, however, are again quite large. Thus, we conclude that there is no significant evidence of biasing treatment-control spillovers operating through student friendships.

5 Concluding remarks

Does teaching school children about recycling reduce household waste? For the environmental education program examined in this paper, the answer is apparently in the negative: we find no evidence that the intervention had any effect on the actual amounts of residual or food waste generated in households with participating children. This is an unusual result in the empirical literature on intergenerational learning (Lawson et al., 2018), and is especially notable given the methodological contributions of our study, i.e. that we are able to both implement a panel RCT design and to measure verified rather than self-reported pro-environmental behavior. The null result also arises despite the fact that an *ex-ante* power analysis indicated that our sample was sufficient to detect even quite minor effects. Finally, given that treatment-arm allocation was randomized within, rather than between classes (as in a cluster-randomized trial), it is notable that we also find no evidence of treatment-control spillovers operating across student social networks.

Of course, these findings do not imply that EEPs never affect actual behavior. First, it is possible that the students in our sample had already been ‘treated’ to some extent through conventional school activities. Thus, our EEP might have been ineffectual because it was added on top of teaching that was already affecting waste-related household behavior. The national Swedish curriculum for elementary and high schools does state that, for example, students should “be given opportunities to... form personal attitudes in large-scale and global environmental issues”, and schools also have an explicit objective to make students

Table 5: *Treatment-control interference estimates*

	Maximal number of links				Minimal number of links			
	Residual waste		Food waste		Residual waste		Food waste	
Variable	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.	Coefficient	p, rand. inf.
Treatment	0.085		-0.029		0.095		-0.022	
	(0.139)		(0.083)		(0.117)		(0.075)	
No. treated friends	0.074	0.390	-0.012	0.759	0.029	0.526	-0.020	0.377
	(0.098)		(0.045)		(0.091)		(0.046)	
Address FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,907	3,907	3,916	3,916	5,289	5,289	5,295	5,295
Addresses (focal)	182	182	182	182	246	246	246	246
R-squared (within)	0.020	–	0.039	–	0.020	–	0.032	–

Robust standard errors clustered at the address level reported in parentheses. Columns 'p, rand. inf.' give randomization-t p values, i.e. the share of re-randomized treatment vectors (out of 1000) that yield larger t statistics than those implied by the regression results of the preceding column. From our adjustment for multiple hypothesis testing, approximately * $p < 0.058$, ** $p < 0.028$, *** $p < 0.006$.

"respect and care for both the local and the wider environment". Nevertheless, we would have expected that the intervention would make waste issues more salient to students during the intervention period, driving at least a short-lived effect on behavior; no such effect was found.

Second, it is possible that our EEP was too limited in scope to provide a significant effect on behavior; we note that the interventions that have been evaluated in previous studies generally tend to be more extensive (e.g. Grodzinska-Jurczak et al., 2003). Relatedly, participating students were certainly aware that the program was external to their usual activities and would have no bearing on their grades, which may have lowered their engagement and hence the impact on household behavior. However, we also find no moderating effect of engagement within our sample, which seems inconsistent with null results being caused by such issues.

In any case, future research should strive to overcome these various issues, boosting the intensity of the intervention and integrating it more closely with normal school curricula without compromising the robustness of the identification strategy. In the meantime, this paper has provided initial field-experimental evidence on the impact of an EEP on the actual, focally pro-environmental behaviors of recycling and waste reduction.

References

W. Abrahamse and E. Matthies. Informational strategies to promote pro-environmental behaviour: Changing knowledge, awareness and attitudes. In L. Steg, A.E. van den Berg, and J.I.M. de Groot, editors, *Environmental psychology: An introduction*, pages 223–232. John Wiley & Sons, Oxford, UK, 2012.

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

S. Athey, D. Eckles, and G.W. Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113:230–240, 2018.

R. Ballantyne, J. Fien, and J. Packer. Program effectiveness in facilitating intergenerational influence in environmental education: Lessons from the field. *Journal of Environmental Education*, 32(4):8–15, 2001a.

R. Ballantyne, J. Fien, and J. Packer. School environmental education programme impacts upon student and family learning: A case study analysis. *Environmental Education Research*, 7(1):23–37, 2001b.

H. Boudet, N.M. Ardoin, J. Flora, K.C. Armel, M. Desai, and T.N. Robinson. Effects of a behavior change intervention for Girl Scouts on child and parent energy-saving behaviours. *Nature Energy*, 1:1–10, 2016.

A. Bucciol, N. Montinari, and M. Piovesan. Do not trash the incentive! Monetary incentives and waste sorting. *Scandinavian Journal of Economics*, 117(4):1204–1229, 2015.

H. Bulkeley and N. Gregson. Crossing the threshold: Municipal waste policy and household waste generation. *Environment and Planning A*, 41(4):929–945, 2009.

J. Campbell, T.M. Waliczek, and J.M. Zajicek. Relationship between environmental knowledge and environmental attitude of high school students. *The Journal of Environmental Education*, 30(3):17–21, 1999.

P. Damerell, C. Howe, and E.J. Milner-Gulland. Child-oriented environmental education influences adult knowledge and household behavior. *Environmental Research Letters*, 8 (1), 2013.

J. Duvall and M. Zint. A review of research on the effectiveness of environmental education in promoting intergenerational learning. *The Journal of Environmental Education*, 38(4): 14–24, 2007.

C. Ek. A formula for power calculation in cluster-randomized experiments with panel data, 2019. Working paper.

European Commission. Towards a circular economy: A zero waste programme for Europe, 2014. COM(2014) 398 final/2.

European Commission. Report on the implementation of EU waste legislation, including the early warning report for Member States at risk of missing the 2020 preparation for re-use/recycling target on municipal waste, 2018. COM/2018/656 final.

Eurostat. Sustainable development in the European Union. 2015 monitoring report of the EU sustainable development strategy, 2015. Luxembourg: Publications Office of the European Union.

E. Gentina and I. Muratore. Environmentalism at home: The process of ecological resocialization by teenagers. *Journal of Consumer Behavior*, 11(2):162–169, 2012.

M. Grodzinska-Jurczak, A. Bartosiewicz, and A. Twardowska. Evaluating the impact of school waste education programme upon students', parents' and teachers' environmental knowledge, attitudes and behaviour. *International Research in Geographical and Environmental Education*, 12(2):106–122, 2003.

H.R. Hungerford and T.L. Volk. Changing learner behavior through environmental education. *Journal of Environmental Education*, 21(3):8–21, 1990.

D.F. Lawson, K.T. Stevenson, M.N. Peterson, S.J. Carrier, R. Strnad, and E. Seekamp. Intergenerational learning: Are children key in spurring climate action? *Global Environmental Change*, 53:204–208, 2018.

D.F. Lawson, K.T. Stevenson, M.N. Peterson, S.J. Carrier, R.L. Strnad, and E. Seekamp. Children can foster climate change concern among their parents. *Nature Climate Change*, 9:458–462, 2019.

F.C. Leeming, B.E. Porter, W.O. Dwyer, M.K. Cobern, and D.P. Oliver. Effects of participation in class activities on children's environmental attitudes and knowledge. *The Journal of Environmental Education*, 28(2):33–42, 1997.

J.G. MacKinnon and M.D. Webb. Randomization inference for difference-in-differences with few treated clusters, 2019. Queen's Economics Department Working Paper No. 1355.

P. Maddox, C. Doran, I.D. Williams, and M. Kus. The role of intergenerational influence in waste education programmes: The THAW project. *Waste management*, 31(12):2590–2600, 2011.

D. McKenzie. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2), 2012.

D.B. Rubin. Comment on: "Randomization analysis of experimental data: The Fisher randomization test". *Journal of the American Statistical Association*, 75(371):575–582, 1980.

C. Vaughan, J. Gack, H. Solorazano, and R. Ray. The effect of environmental education on school children, their parents, and community members: A study of intergenerational and intercommunity learning. *Journal of Environmental Education*, 34(3):12–21, 2003.

A. Young. Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics*, 134(2):557–598, 2019.

Online Appendices for article “Can children’s engagement in recycling processes reduce household waste?”

Appendix A. Experimental materials

A.1 Scripts for first class visit (translated from Swedish)

In front of the entire group:

My name is X and this is Y. We are from University of Gothenburg and we are doing research on questions related to science and the environment. We will visit you twice. Today we will give you a home assignment. We will return in a week or two to discuss your experiences and what you thought about the task. At that occasion we will also play a game.

We will now divide you into two groups. Y will give you one note each, and on this you will see an A or a B written. Those that receive an A will stay in this room with me, whereas those who got a B will go to another class room together with Y.

Control group:

You will be given the task of recording the weather once per day for a week. Each of you will receive a form. [Hold up the form.]

As you can see, there are seven fields in the lower part of the form. They have the headings “Monday”, “Tuesday”, down to “Sunday”. Today is [weekday] and it is the first day of the assignment, so you start by filling in that field today. Then you continue to fill one day at a time until all the boxes are filled in a week. In each field you should write down the outdoor temperature and precipitation, as well as some other information. [Hand out the form].

First look at your form. As you can see, every day there are several questions about the weather — you should fill in: (i) what the outdoor temperature is right now, (ii) whether it has rained or not during the day, and (iii) what the sky looks like right now. You can check the outdoor temperature on a thermometer in your home. When you do, write the value on this line. [Show where to write]

Next, you should indicate whether it has rained or not. Here you only tick one of the boxes “yes” or “no”. To describe the sky, choose one of following three options: (i) clear, (ii) partially clear, and (iii) cloudy. If the sky is completely blue, or almost entirely blue, select “clear”. If the sky is roughly half-half cloudy and clear, choose “partially clear”. If the sky is completely cloudy, or almost completely cloudy, select “cloudy”.

Any questions so far?

You should also write at which address you measured the outdoor temperature, among other things. For example, one day during the week you might be at a friend’s or a relative’s home. Then you write their addresses. If you are at home, you write your own address, of course.

Finally, there is a question on what time it is. It is a good idea to adopt the habit of recording the weather at about the same time each day, so that the figures are comparable. You could, for example, try to do the recording every night before brushing your teeth and getting ready for bed. But you could also do it immediately when you get home, before you do anything else.

Now, let’s answer the questions at the top of the form together. At the very top it says that you should write your name, class and school, so please start with that. We will eventually collect your forms as part and use them as part of our research, but we will treat them such that we never find out who has given what answers. This applies to everything on the form, including your addresses. Once your information is added in the computer, your names will be replaced by a number. [Walk around and see that everyone understands and knows where on the form to write]

On the next line, you should write how many people live in your household. As you can see, you should include yourselves. If you live in two or more places, use the place where you live most days this week. Next, you write the age of the other people in your household; it is enough that you write the ages with commas in between. After that, write whether you have pets or not, and if you do have pets, please specify what type.

The next question is about which classmates you spend the most time with in your spare time. Note that you should only consider those who are in the same class as you. If you spend time with fiends that are not in the same class as you, please do not write their names.

The last question is how far from a lake/the ocean you live. Write down the shortest distance — if it is closer to a lake, write that distance.

You can now start filling in the top part of the form. Raise your hand if you have a question, and I will come and help you. [Walk around and answer questions until everyone has filled in the top part of the form]

That is all for today. I will come back in a few weeks, and we will then talk some more about temperature and precipitation.

Treatment group: You will be given the task of weighing the waste at home once a day for a week. You will get a box with the materials and equipment you need to do this. Please do not open the boxes until I have explained what the task entails. [Hand out the boxes.]

In each box, you will find a form. [Hold up the form.]

As you can see, there are seven fields in the lower part of the form. They have the headings "Monday", "Tuesday", down to "Sunday". Today is [weekday] and it is the first day of the assignment, so you start by filling in that field today. Then you continue to fill one day at a time until all the boxes are filled in, a week from now. In each field you should write down the outdoor temperature and precipitation, as well as some other information. [Hand out the form].

Each of you will also be given a scale like this one. [Hold up the scale for the class.]

It is easy to use. Simply press the "on" button, then attach the bag you want to weigh on the hook. Wait for the scale to settle, and soon you will be able to see the weight in kilograms in the display. Note that the weight (in kilograms) is displayed with three decimals, so you should use the comma and all decimals when you write the weight in your form.

You can now open the box in front of you. The form is at the bottom. There are also seven plastic bags in the box — I will explain how you should use them.

First, look at your form. As you can see, there are two lines for waste weights each day: one for food waste, and one for residual, unsorted waste. That's because most of you have two different bins at home where you put food wastes in a separate bin. All the other waste goes in the residual waste bin. You should weigh both the food and residual waste bags separately. This is how you do that: [Hold up the scale]

Take the bag for food waste, put it in one of the plastic bags you have received in the box, and hang the plastic bag on the scales to record its weight. After that, take out the food waste bag from the plastic bag and replace it with the residual waste bag, and finally weigh that. This completes the weighing for that day.

If you only have a single bin for all waste at home, you just weigh the total waste and write it on the "residual waste" line. Leave the food waste line empty.

Any questions so far?

You should also write at which address you weighed the waste. For example, one day during the week you might be at a friend's or a relative's home. Then you write their addresses. If you are at home, you write your own address, of course. You should also write if someone visited you during the day (and if so, by whom) and what you had for dinner.

We suggest that you change both your food waste bag and general waste bag every day, so that the waste you weigh is just what you generate each day. It is also a good idea to adopt the habit of weighing the waste at about the same time each day, so that all the weights are comparable. You could, for example, try to do it every night before brushing your teeth and getting ready for bed. But you could also do it immediately when you get home, before you do anything else.

Now, let's answer the questions at the top of the form together. At the very top it says that you should write your name, class and school, so please start with that. We will eventually collect your forms as part and use them as part of our research, but we will treat them such that we never find out who has given what answers. This applies to everything on the form, including your addresses. Once your information is added in the computer, your names will be replaced by a number. [Walk around and see that everyone understands and knows where on the form to write]

On the next line, you should write how many people live in your household. As you can see, you should include yourselves. If you live in two or more places, use the place where you live most days this week. Next, you write the age of the other people in your household; it is enough that you write the ages with commas in between. After that, write whether you have pets or not, and if you do have pets, please specify what type.

The next question is about which classmates you spend the most time with in your spare time. Note that you should only consider those who are in the same class as you. If you spend time with fiends that are not in the same class as you, please do not write their names.

The last question is about the type of waste bins you have at home. The question is what waste bins you have standing outside your house or in your garbage room. As we've said, most of you sort out food waste, and then you can write "Food and residual waste bins" on the form. If you do not, but rather mix food and other waste in the same bin, you can write "Mixed waste" on the form. Please raise your hand if you know that you have some other arrangement at home, and I will come help you. And if you don't know what bins you have, you can wait to fill in this line until you have asked your parents about it.

You can now start filling in the top part of the form. Raise your hand if you have a question, and I will come and help you. [Walk around and answer questions until everyone has filled in the top part of the form]

That is all for today. I will come back in a few weeks, and we will then talk some more about waste and recycling.

A.2 Student forms (translated from Swedish)

Waste diary

Your name (first and family names), class and school: ....

How many people live in the same house as you? Include yourself. ....

How old are the others in your family?........ ....

Do you have pet(s) in your family? If 'yes', what kind of pet(s)? ....

Which class mates do you spend most time with during your spare time? Write first and family names...

....

What kind of (outdoor) waste bins do you have at your house?...

Monday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Tuesday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Wednesday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Thursday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Friday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Saturday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Sunday

Residual waste: .... kg

Food waste: .... kg

What is the address of the house where you weighed the waste today? Write street name and house number:

....

Did someone visit you today? If 'yes', whom? ....

What did you have for dinner? ? ....

Figure A.1: Form filled out by treated participants (translated from Swedish)

Weather diary

Your name (first and family names), class and school: ……………………………………………………

How many people live in the same house as you? Include yourself. ……………………………………………………

How old are the others in your family?…… ……………………………………………………

Do you have pet(s) in your family? If 'yes', what kind of pet(s)? ……………………………………………………

Which class mates do you spend most time with during your spare time? Write first and family names……………………………………………………

……………………………………………………

How far from a lake/the ocean do you live?……………………………………………………

Monday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Tuesday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Wednesday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Thursday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Friday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Saturday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Sunday

Temperature: …………………… ° C

Has it rained today?

☐ Yes ☐ No

At what address did you register the weather today? Give street name and number:

……………………………………………………

How would you describe the sky now? ☐ Clear ☐ Partly cloudy ☐ Cloudy

What is the time now? …………………………

Figure A.2: Form filled out by control participants (translated from Swedish)

A.3 Waste game rules (translated from Swedish)

In this game, players acquire new knowledge about waste management, and practise their hands-on recycling skills by placing different types of waste in the right recycling category. To be given the opportunity to recycle the waste at hand, a player needs to give the right answer to the question that is read to her/him. If s/he gives the wrong answer, s/he will not be allowed to put down any of the waste cards in her/his hand. At the end of the game, each player sums up her/his points. The one with the most points is the winner.

The following cards are used in the game:

Packaging Cards
Each player gets 10 cards from the deck of packaging cards. Some examples of packaging cards:

Recycling Cards
Each player places this set of cards on the table in front of her/him, visible to everyone. Recycling categories are (from left to right): plastic, metal, paper, glass, and organic.

Question cards
Each card has a question and three possible answers. The correct answer is typed at the bottom, along with a reference for those who wants to read more about the topic.

How to Play the Game
The player to the right of the active player takes the top question card and reads the question on that card aloud. When a player has read a question, s/he places that card at the bottom of the deck.

<u>If the active player answers correctly</u>: The active player draws a packaging card and may choose to put any packaging card in her/his hand on one of the five recycling cards in front of her/him. The game then continues: the player that just answered a question draws the next question card and reads it to the new active player sitting on her/his left.

<u>If the answer is incorrect</u>: The player that gave the wrong answer does not draw a packaging card and is not allowed to recycle any of her/his packaging cards. The game continues: the player that just answered a question draws the next question card and reads it to the new active player sitting on her/his left.

How the Game Ends
The game ends whenever a player has put down all her/his packing cards on the recycling cards. When this happens, do the following:
1. All players sum up their points on the packaging cards that they have placed on their recycling cards.
2. Note that no points are awarded for cards that have been placed on the wrong recycling category.
3. The player that first put down all her/his packaging cards gets a bonus of 50 points.

Appendix B. Anomaly report coding

Table B.1: *Coding of anomaly incidents*

Description	Report data code	Action taken
Bin not curbside	010	Code as zero
Blocked, car	020	Code as zero
Blocked, snow	030	Code as zero
Blocked, other	040	Code as zero
Locked door/gate	050	Code as zero
Not shoveled	060	Code as zero
Not plowed	070	Code as zero
Not gritted	080	Code as zero
Incorrect bin contents, not collected	090	Code as zero
Incorrect bin contents, collected	095	Ignore incident
Overfull	100	Ignore incident
Heavy bin	105	Ignore incident
Other	110	Ignore incident
Broken bin	120	Ignore incident
Bar code missing	130	Code as missing
Label missing	135	Ignore incident
Empty bin	140	Code as zero
Sacks collected	150	Code as missing
Broken wheel	160	Ignore incident
Food waste bag	165	Ignore incident
Food waste bags often	166	Ignore incident
Broken lid	170	Ignore incident
Cannot find bin	180	Code as zero
Bar code broken	190	Code as missing
Manual collection	195	Code as missing

Table lists possible anomaly incidents, their coding in the raw data, and how we handle these incidents within the study.

Appendix C. Power calculation

Our power calculation assumes that the data is generated by the process

$$y_{it} = \lambda_t + \delta T_{it} + u_i + u_{it} \tag{C.1}$$

where i indexes addresses, t indexes time, and y_{it} are per-person residual-waste weights. Random errors u_i and u_{it} operate between and within addresses, respectively, and are assumed i.i.d. with mean zero and constant variances σ_p^2 and σ_{pt}^2. Under such conditions (McKenzie, 2012), the difference-in-difference estimator

$$\hat{\delta}^{DD} = \bar{y}_T^{POST} - \bar{y}_T^{PRE} - \left(\bar{y}_C^{POST} - \bar{y}_C^{PRE}\right)$$

which we estimate by regression methods, has the variance

$$Var\left(\hat{\delta}^{DD}\right) = \left(\frac{1}{N_T} + \frac{1}{N_C}\right)\left(\frac{1}{m} + \frac{1}{r}\right)\sigma_{pt}^2$$

and this formula forms the basis of our power calculation. Here N_T and N_C are the numbers of treated and control addresses in the sample; we will use the actually realized figures for modal addresses, namely $N_T = 173$ and $N_C = 163$. Furthermore, this formula makes the simplifying assumption that all addresses are treated at the same time: m is the number of periods prior to treatment, and r is the number of periods after treatment has been activated. Supposing that all addresses are treated at the start of period 13, which is halfway through the actual intervention period, we have $m = 12$ and $r = 10$. Choosing other (common) starting points within our intervention period typically has a very minor effect on power, at most increasing our 80% minimum detectable effect (MDE) by about 7%.

Finally, for σ_{pt}^2, we use an estimate calculated for another project (AEA RCT Registry ID 0003301) on a historical residual-waste dataset for Varberg covering nearly all single-family homes in that municipality and the period from 25 October 2017 to 28 August 2018, a total of 22 two-week periods. We use a mixed-effects model (the `mixed` command in Stata) to estimate C.1 with $T_{it} = 0$ everywhere; for details, see Ek (2019). Our point estimates are $\sigma_p^2 = 9.301$ and $\sigma_{pt}^2 = 7.259$.

Taken together, these values imply that $Var(\hat{\delta}^{DD}) \approx 0.016$, and thus the MDE at 80% power is 0.353 kg of residual waste per person. This is about 9.3% of the data average, or 8.7% of a standard deviation. Since 20% of a SD is usually considered a 'small' effect, we conclude that the study is adequately powered to estimate even quite small effects.

Appendix D. Regression tables

Table D.1: *Balance tests*

	Difference	p	N
Residual waste			
Period 1	-0.469*	0.094	343
Period 2	-0.012	0.963	348
Period 3	-0.166	0.508	348
Period 4	-0.224	0.415	350
Period 5	0.085	0.755	343
Period 6	-0.209	0.521	348
Period 7	0.077	0.786	347
Period 8	-0.035	0.901	348
Period 9	-1.226	0.689	347
Period 1-9	-0.120	0.203	3,122
Food waste			
Period 1	-0.106	0.439	343
Period 2	-0.136	0.304	348
Period 3	-0.029	0.822	349
Period 4	-0.074	0.639	350
Period 5	-0.296*	0.075	342
Period 6	-0.220	0.134	348
Period 7	-0.032	0.811	348
Period 8	-0.102	0.385	349
Period 9	-0.155	0.285	347
Period 1-9	-0.128***	0.007	3,124

Table tests for pre-treatment covariate balance across treatment and control groups. Each row corresponds to a single regression of either residual or food waste in some baseline period(s) on a dummy that is equal to one for all addresses that are eventually treated. N is the number of addresses included in each regression (time period). Waste variables measured in kgs/person. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.2: *Treatment effect estimates: excluding ambiguous cases*

	Residual waste		Food waste	
Variable	Coefficient	p (rand. inf.)	Coefficient	p (rand. inf.)
Treatment	0.554	0.577	-0.030	
	(0.098)		(0.064)	
Address FE	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes
Observations	6,912	6,912	6,917	6,917
Addresses	321	321	321	321
R-squared (within)		–		–

Table presents regression estimates for the effect of treatment, excluding all adresses that are modal for more than one student. Robust standard errors clustered at the address level reported in parentheses. Columns 'p (rand. inf.)' give randomization-t p values, i.e. the share of re-randomized treatment vectors (out of 1000) that yield larger t statistics than those implied by the regression results of the preceding column. From our adjustment for multiple hypothesis testing, approximately * $p < 0.058$, ** $p < 0.028$, *** $p < 0.006$.