recycling

MDPI

Article

The Impact of Social Norm Versus Educational Message Framing on Residential Recycling: A Randomized Field Trial

Suraksha Baral and Brian E. Roe *

Department of Agricultural, Environmental, & Development Economics, Ohio State University, Columbus, OH 43210, USA; baral.20@osu.edu
* Correspondence: roe.30@osu.edu

Abstract

National recycling rates have stagnated at approximately 32 percent in the United States, suggesting that existing programs are not achieving desired levels of diversion. A growing body of field experiments has examined whether informational and norm-based messaging interventions can shift recycling behavior, though most studies rely on self-reported outcomes or stated preferences rather than objective behavioral measures. Using bin-level audit data from a randomized field trial among 200 voluntarily enrolled households in Columbus, Ohio, we find that norm-based and educational message framing produce statistically indistinguishable average effects on recycling fill level, participation, and contamination over a 29-week program. However, the analysis identifies that norm-based messaging outperforms educational messaging when adverse weather raises participation costs, while educational messaging holds up better during high-distraction community events. These patterns point toward adaptive messaging designs that condition content on observable contextual triggers as a promising direction for extracting greater behavioral returns from recycling outreach investment.

Keywords: recycling; consumer behavior; randomized field trial; norm-based messaging; contamination rate

1. Introduction

Recycling reduces the social costs of waste disposal by diverting materials from landfills, lowering greenhouse gas emissions, and conserving natural resources, yet households do not fully internalize these returns, providing a rationale for public intervention when households fall short of compliance with their community's recycling mandates [1,2]. In the United States, the national recycling rate has stagnated at approximately 32 percent, well below the EPA's 50 percent target, suggesting that existing programs are not achieving socially optimal levels of diversion [3]. Understanding why households fail to comply with recycling mandates and what interventions can cost-effectively close this gap is therefore a question of both environmental and economic significance. Even households with convenient access to recycling infrastructure routinely under-recycle, which points to behavioral frictions, including information deficits, weak motivation, and inattention to social expectations, as the binding constraints [4–6]. Addressing these frictions through low-cost messaging programs, rather than expensive infrastructure expansion, offers an attractive policy lever, but its effectiveness depends critically on which messages change behavior.

A growing body of field experiments has examined whether informational and norm-based messaging interventions can shift recycling behavior [7–10], though most studies rely on self-reported outcomes or stated preferences rather than objective behavioral

Academic Editor: Giovanni De Feo

Received: 23 June 2026
Revised: 19 July 2026
Accepted: 25 July 2026
Published: 30 July 2026

measures [11–13], a gap reinforced by direct evidence that self-reported recycling intention has limited predictive power for actual behavior [14]. Schultz [7] provided early experimental evidence that normative feedback increased the frequency of curbside recycling participation among community residents, measured through direct observation over 17 weeks. In the energy and water conservation domains, large-scale field experiments have demonstrated that norm-based messages produce modest but meaningful reductions in consumption, measured through utility records rather than self-reports [4,5]. In the recycling context, however, evidence from field experiments using objective bin-level data remains limited. Czajkowski, Zagórska, and Hanley [15] found that social norm nudges shift stated willingness to pay for recycling contracts among Polish households, with effects varying by the level of the communicated norm and households' prior recycling behavior. Sorkun [16] similarly drew on survey data to show that social norms influence recycling through perceived convenience rather than knowledge acquisition alone. Wichmann et al. [17] conducted a field trial of pro-environmental messaging on the observed incidence of paper recycling and reuse in University computer laboratories and found that compliance with best practice was significantly higher in the treated laboratory than in the control laboratory. The broader meta-analytic literature confirms that personal internalized norms are the strongest predictors of pro-environmental behavior, but this evidence base is overwhelmingly correlational rather than experimental [18]. Wang et al. [11] similarly relied on self-reported behavioral intentions, finding in a controlled experiment that the effect of gain versus loss message framing on stated waste-separation intent depended on whether the frame was paired with an environmental or a monetary appeal, an early signal that framing effects on recycling-related behavior may be conditional rather than uniform. Most recently, Aravind, Mishra, and Das [12] tested three positive message frames, namely economic gain, behavioral impact, and anthropomorphic appeal, on stated Reduce–Reuse–Recycle intentions in a randomized online experiment. Using a self-reported Likert rating, they found that behavioral-impact and anthropomorphic framing outperformed economic framing. However, experimental evidence using objective measures of recycling behavior to identify the causal effects of message framing on household recycling remains limited.

Two gaps stand out in the current literature. First, most field experiments testing messaging interventions on recycling behavior rely on self-reported outcomes or stated preferences rather than objective, audit-based measures, limiting the ability to draw credible causal inferences about effects on actual behavior [15,18] (though [17] is an exception). Second, existing studies predominantly compare norm-based messaging against a no-treatment control, leaving unanswered whether social norm framing outperforms equally well-designed educational messaging that targets the same information deficit through a different motivational channel [5,7]. Early research by Burn and Oskamp [9] found that persuasive communication and public commitment each outperformed a no-treatment control but did not differ significantly from one another. This is an early instance of the pattern this study revisits with modern message-framing content and objective bin-level measurement. This gap mirrors a pattern identified in the broader message-framing literature on sustainable consumer behavior, where a systematic review found that single message frames rarely produce consistent effects on their own and that comparative or combined framing designs remain comparatively underexplored [19]. These gaps collectively leave municipalities with limited guidance when choosing among specific program designs.

These gaps have direct consequences for municipal recycling policy. Municipalities across the United States allocate substantial resources to recycling outreach and education, yet allocation decisions rarely rest on robust evidence about which messages change observed behavior [3]. Without evidence from objective behavioral measures, program administrators cannot tell whether a given campaign shifts attitudes, behavior, or neither, nor can they judge whether the added cost of social norm content justifies its behavioral returns. The theoretical distinction between educational and norm-based messaging also matters. If educational messages work by reducing information deficits while norm-based messages work by activating social conformity motivation, the two approaches target fundamentally different behavioral mechanisms and may be differentially effective across household types, neighborhood contexts, and program stages [16,18]. Establishing which mechanism dominates, and under what conditions each does, advances both program design and behavioral theory.

This paper addresses these gaps by evaluating differential messaging provided to participants in the Can Fairy Curbside Recycling Program. This randomized field experiment was conducted among the 200 households in the University District, a residential neighborhood in Columbus, Ohio, adjacent to but administratively separate from The Ohio State University, who voluntarily enrolled in the Can Fairy program. Households were randomly assigned to one of two weekly email message conditions: educational messages providing factual guidance on correct recycling practices, or norm-based messages delivering the same educational content framed around community recycling behavior. We measure differences in recycling behavior using objective household-level bin data collected by program staff throughout the intervention period, including weekly participation rates, recycling fill levels, and contamination level. We then analyze data using a difference-in-differences framework that exploits the randomized assignment to identify the causal effect of message framing on recycling behavior. Unlike most existing studies that compare a single intervention to a no-treatment baseline, our two-arm randomization isolates the marginal effect of social norm framing relative to equivalent educational content, offering a more policy-relevant test of whether framing alone drives behavioral change. This also provides actionable evidence for municipalities seeking to optimize the design of low-cost recycling outreach programs.

2. Results

Figure 1 plots the weekly means, with 95 percent confidence intervals, for recycling fill level (panel A), participation rate (panel B), and contamination rate (panel C) by treatment arm over the full study period. Both arms track fill level and participation closely during the pre-intervention period, providing visual support for the parallel trends assumption underlying the difference-in-differences design. Following the November 1 intervention start, both arms exhibit a seasonal decline in fill level and participation, with no systematic divergence between the two arms on average over the post-period. Contamination rate also tracks quite closely during the pre-treatment period, with the four initial weeks yielding identical values with no contaminated bins. Contamination rates remain low throughout most of the study period, with no week exceeding 20 percent of bins contaminated.

2.1. Fill Level

Table 1 presents difference-in-differences estimates of the effect of norm-based versus educational messaging on a key indicator of participation intensity (fill level). Across Models 1 and 2, the coefficient on Treated × Post is statistically insignificant, indicating no average treatment effect of message framing on fill level. The large negative post-period coefficient reflects a seasonal decline in recycling volume affecting both groups equally. Model 2 identifies that several contextual factors are associated with fill level. Namely, fill levels are positively correlated with weekly average temperatures and home football games, though negatively correlated with higher levels of precipitation ('bad weather'). Model 3 reveals significant heterogeneity that households randomly assigned to the norm-based messaging recycled more during bad weather weeks (0.116, $p < 0.01$) but less during warmer weeks (−0.003, $p < 0.10$), while the OSU home game interaction is negative but marginally insignificant.

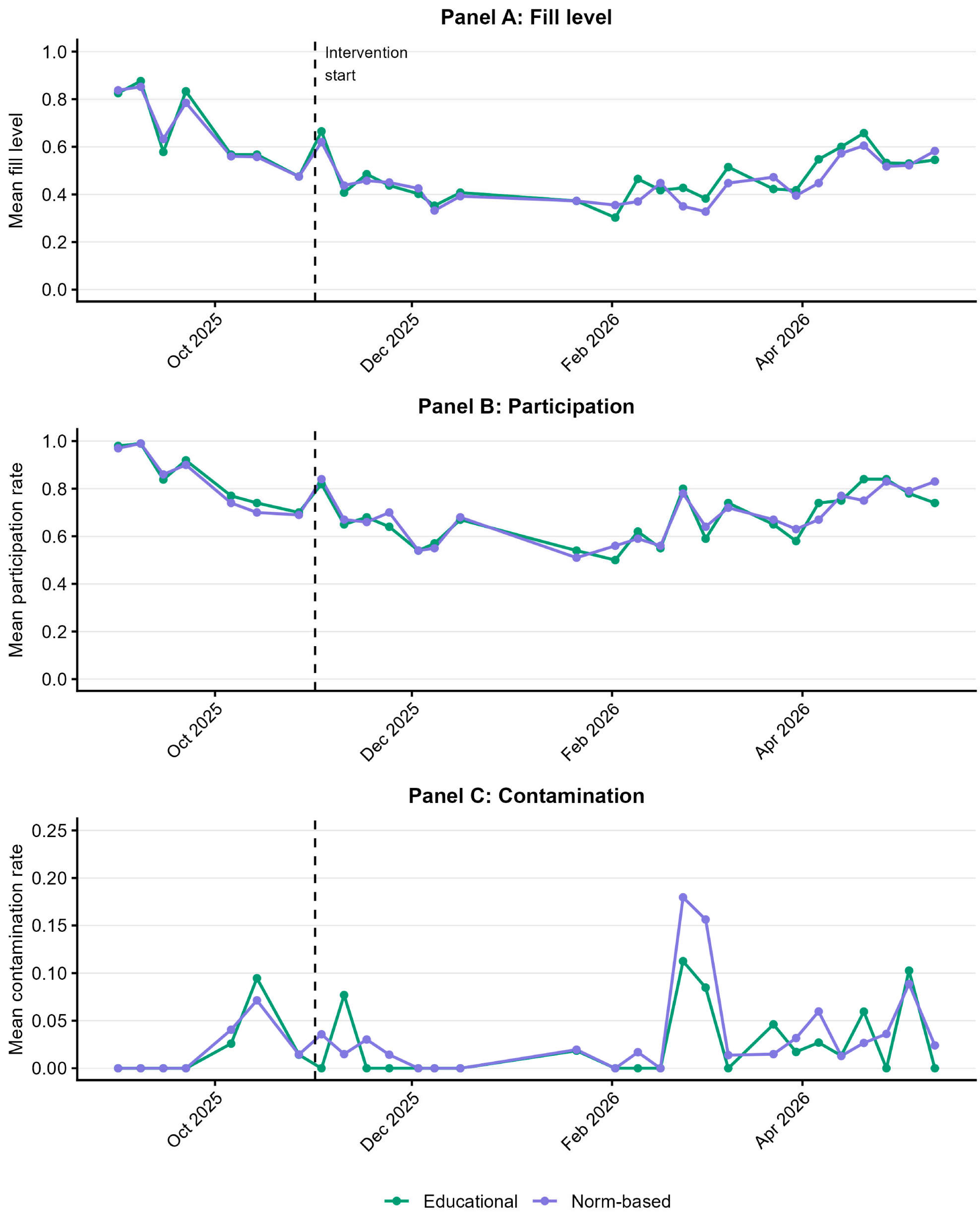

Figure 1. Weekly means for recycling fill, participation rate, and contamination rate by treatment arm.

Table 1. Effects of norm-based versus educational messaging on recycling fill level.

	(1)	(2)	(3)
Dependent variable:	Fill level (0–1 scale)		
Treated × Post (β)	−0.0151	−0.0148	−0.0911 **
	(0.0243)	(0.0243)	(0.0418)
Treated	−0.0028	−0.0031	−0.0031
	(0.0225)	(0.0225)	(0.0225)
Post	−0.2064 ***	−0.0720 ***	−0.0477 **
	(0.0179)	(0.0204)	(0.0216)
Constant	0.6742 ***	0.5207 ***	0.5078 ***
	(0.0157)	(0.0258)	(0.0313)

Table 1. *Cont.*

	(1)	(2)	(3)
Controls			
Avg. temperature (°C)		0.0092 ***	0.0104 ***
		(0.0008)	(0.0011)
Bad weather		−0.0527 ***	−0.0756 ***
		(0.0149)	(0.0175)
OSU home game		0.0466 ***	0.0559 ***
		(0.0139)	(0.0157)
Holiday week		0.0109	0.0134
		(0.0128)	(0.0147)
Treatment × Controls			
Treated × Post × Avg. temp.			−0.0031 *
			(0.0017)
Treated × Post × Bad weather			0.1156 ***
			(0.0326)
Treated × Post × OSU game			−0.0575
			(0.0349)
Treated × Post × Holiday			−0.0216
			(0.0296)
Observations	5796	5796	5796
R^2	0.048	0.076	0.079

Notes. Standard errors clustered at the household level in parentheses. Model (1): simple DiD. Model (2): DiD with time-varying controls. Model (3): DiD with controls interacted with treatment. Bad weather = 1 if weekly total precipitation ≥ 25.4 mm or ≥3 precipitation days. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

2.2. Participation Rate

Table 2 reports estimates for weekly bin setout participation. The Treated × Post coefficient is positive but statistically insignificant across all three specifications, indicating no average effect of message framing on participation rates. Model 2 mirrors the results from the parallel model for fill level in terms of correlated contextual factors: higher average temperatures and home football games are positively correlated with participation while bad weather is negatively correlated. Model 3 also mirrors the results for fill level in that norm-based households show higher participation during bad weather weeks (0.120, $p < 0.01$) and lower participation during OSU home game weeks (−0.072, $p < 0.10$).

Table 2. Effects of norm-based versus educational messaging on recycling participation rate.

	(1)	(2)	(3)
Dependent variable:		Participation (0/1)	
Treated × Post (β)	0.0170	0.0173	−0.0755
	(0.0256)	(0.0257)	(0.0477)
Treated	−0.0120	−0.0123	−0.0123
	(0.0231)	(0.0232)	(0.0232)
Post	−0.1736 ***	−0.0443 **	−0.0284
	(0.0189)	(0.0216)	(0.0231)
Constant	0.8477 ***	0.6910 ***	0.6953 ***
	(0.0154)	(0.0280)	(0.0363)
Controls			
Avg. temperature (°C)		0.0095 ***	0.0098 ***
		(0.0009)	(0.0013)
Bad weather		−0.0420 ***	−0.0678 ***
		(0.0155)	(0.0193)
OSU home game		0.0432 ***	0.0545 ***
		(0.0155)	(0.0185)
Holiday week		−0.0126	−0.0111
		(0.0138)	(0.0153)

Table 2. *Cont.*

	(1)	(2)	(3)
Treatment × Controls			
Treated × Post × Avg. temp.			−0.0011
			(0.0018)
Treated × Post × Bad weather			0.1204 ***
			(0.0401)
Treated × Post × OSU game			−0.0724 *
			(0.0396)
Treated × Post × Holiday			−0.0208
			(0.0325)
Observations	5796	5796	5796
R^2	0.025	0.049	0.051

Notes. Standard errors clustered at the household level in parentheses. Model (1): simple DiD. Model (2): DiD with time-varying controls. Model (3): DiD with controls interacted with treatment. Bad weather = 1 if weekly total precipitation ≥ 25.4 mm or ≥3 precipitation days. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

2.3. Contamination Rate

Table 3 presents contamination estimates, conditional on households setting out their bin with material. Baseline contamination rates are low, averaging 2.4% in the educational arm and 3.2% in the norm-based arm, which limits statistical power. Because treatment may itself affect the probability of setting out material, these results should be interpreted as conditional associations among participating household–weeks rather than as unconditional treatment effects on contamination across the full sample. The Treated × Post coefficient is insignificant in Models 1 and 2. Model 2 reveals a positive correlation between contamination and both average temperature and bad weather, but no significant association with home football games or holidays. In Model 3, norm-based households show significantly higher contamination rates in the reference condition, non-bad-weather, non-game, non-holiday weeks (0.036, $p < 0.10$), but during holiday weeks this is essentially offset with lower contamination (−0.038, $p < 0.01$).

Table 3. Effects of norm-based versus educational messaging on recycling contamination rate.

	(1)	(2)	(3)
Dependent variable:		Contamination (0/1)	
Treated × Post (β)	0.0121	0.0118	0.0361 *
	(0.0101)	(0.0101)	(0.0201)
Treated	−0.0016	−0.0014	−0.0013
	(0.0074)	(0.0074)	(0.0074)
Post	0.0100	0.0049	0.0045
	(0.0073)	(0.0095)	(0.0100)
Constant	0.0169 ***	0.0019	−0.0051
	(0.0056)	(0.0087)	(0.0102)
Controls			
Avg. temperature (°C)		0.0006 **	0.0005
		(0.0003)	(0.0004)
Bad weather		0.0192 ***	0.0263 ***
		(0.0064)	(0.0061)
OSU home game		−0.0079	−0.0036
		(0.0063)	(0.0075)
Holiday week		−0.0067	0.0047
		(0.0050)	(0.0067)
Treatment × Controls			
Treated × Post × Avg. temp.			0.0004
			(0.0006)

Table 3. *Cont.*

	(1)	(2)	(3)
Treated × Post × Bad weather			−0.0231
			(0.0195)
Treated × Post × OSU game			−0.0164
			(0.0163)
Treated × Post × Holiday			−0.0382 ***
			(0.0128)
Observations	4152	4152	4152
R^2	0.003	0.006	0.008

Notes. Standard errors clustered at the household level in parentheses. Model (1): simple DiD. Model (2): DiD with time-varying controls. Model (3): DiD with controls interacted with treatment. Bad weather = 1 if weekly total precipitation ≥ 25.4 mm or ≥3 precipitation days. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

2.4. Robustness Check

The bad weather interaction on fill level is robust across five alternative threshold definitions, remaining positive throughout, with the strongest and most stable estimates observed under thresholds that retain the three-day precipitation rule (0.063 to 0.116) (Appendix A Figure A1). The leave-one-week-out diagnostic further supports this result, with the bad weather interaction remaining positive across all twenty-nine iterations regardless of which week is excluded (Appendix A Figure A2). The holiday interaction on contamination is similarly stable across week exclusions, consistent with its role as the second robust finding of this study (Appendix A Figure A4). The OSU home game interaction is considerably more sensitive to which week is excluded from the leave-one-week-out diagnostic, with point estimates shifting substantially and confidence intervals frequently spanning zero, and it should be interpreted with more caution than the bad weather and holiday findings (Appendix A Figures A2–A4). All figures for both diagnostics are reported in the Appendix A.

3. Discussion

This study provides experimental evidence on the relative effectiveness of norm-based versus educational message framing in a voluntary community curbside recycling program and reveals how key contextual factors such as weather and community events are related to program compliance and participation intensity. Both treatment arms receive active weekly messaging. The educational arm receives factual guidance on correct recycling practices, while the norm-based arm receives the same content supplemented with social comparison information about community recycling behavior. The study does not include a no-message control group, and the findings therefore speak to the relative effectiveness of two active messaging strategies rather than to the absolute effect of either strategy relative to no intervention.

Across three objective outcomes measured through physical audit of household recycling bins over 29 weeks, Models 1 and 2 find no statistically significant difference between the two messaging strategies for two indicators of program compliance (participation, contamination) and an indicator of participation intensity (fill level). This finding is consistent across the simple difference-in-differences and covariate-adjusted specifications. On average, neither strategy is better than the other, and the choice between them based on average behavioral returns alone provides no empirical basis for preferring one over the other. This pattern of equivalence among active treatments has precedent. Burn and Oskamp [9] similarly found that persuasive communication and public commitment did not differ significantly from one another, though in their case both treatments were shown to significantly outperform a no-treatment control, a comparison the present two-arm design cannot speak to given the absence of a control condition.

Model 3 reveals that this average equivalence masks systematic variation in relative behavior across contextual conditions, and these patterns constitute the substantively most important findings of the research. For both fill level and participation, the interaction between the treatment indicator and the bad weather variable is large, positive, and statistically significant at the one percent level. The net effect of norm-based messaging relative to educational messaging on bad weather weeks, combining the baseline treatment interaction coefficient with the bad weather interaction coefficient, is positive for both fill level (net effect approximately 2.5 percentage points) and participation (net effect approximately 4.5 percentage points). This finding is robust to five alternative bad weather threshold definitions and to sequential exclusion of each collection week (Appendix A Figures A1–A3). This means that during weeks when adverse weather raises the physical cost of setting out a recycling bin, norm-based households recycle significantly more than educational households. A plausible interpretation is that social comparison information activates conformity motivation that offsets the higher participation cost on bad weather weeks, while educational reminders about correct sorting practices operate through an informational channel that is less responsive to changes in participation costs. This interpretation is consistent with [16], who finds that perceived convenience mediates the relationship between social norms and recycling behavior, and with the broader argument that norm-based interventions are more effective precisely when behavior requires effort that individuals might otherwise forgo. Manipulation checks, message engagement data, and survey evidence directly linking message receipt to perceived norms, motivation, or convenience were not collected as part of this study. The norm-based treatment combined descriptive, injunctive, and community-identity content rather than isolating a single normative channel. The interaction pattern reported here is therefore consistent with, but does not establish, a norm-salience or conformity mechanism, and alternative explanations involving general differences in message content or tone across arms cannot be ruled out with the present data.

The reverse pattern emerges for participation during Ohio State University home football game weeks, where the treatment interaction is negative and significant at the ten percent level, indicating that educational households maintain higher participation rates than norm-based households during high-distraction campus events. One interpretation is that during periods when social attention is concentrated on shared community events, households are less responsive to social comparison information about recycling, while the practical informational content of educational messages retains behavioral relevance regardless of the broader social context. This finding is consistent with attentional theories of norm salience that predict social norm messages are most effective when norms are made salient in the relevant decision context [20] and least effective when competing social stimuli redirect attention away from the behavior in question.

For contamination, the significant negative holiday week interaction in Model 3 indicates that norm-based households are less likely to contaminate their bins during holiday weeks relative to educational households. This interaction is stable across the leave-one-week-out diagnostic, with the coefficient remaining close to its full-sample estimate regardless of which week is excluded (Appendix A Figure A4). This finding is difficult to interpret mechanistically, given the low overall contamination rates in the sample, averaging 2 to 3 percent across both arms compared to a national average of 15 to 25 percent for single-stream programs [21], and should be treated as exploratory rather than confirmatory, given the low baseline rates and the multiple comparisons involved in the interactive analysis.

The results suggest that the relative effectiveness of norm-based versus educational messaging may not be fixed but could depend on the situational context in which messages are received. Educational messaging performs as well as norm-based messaging on average and shows a directional advantage during high-distraction periods, though this pattern rests on considerably less stable estimates and should not yet be treated with the same confidence as the weather finding discussed below. Norm-based messaging outperforms educational messaging during high-cost periods defined by adverse weather conditions, a finding supported across multiple robustness checks. This asymmetry has a practical implication for program design, though a tentative one given the sensitivity of the campus-event finding. Rather than choosing between the two strategies uniformly, recycling programs might consider conditioning message content on observable contextual triggers such as adverse weather, where the evidence for norm-based framing is strongest, while treating context-conditioning around high-distraction events as a hypothesis for future confirmatory testing rather than an established design recommendation. Such an adaptive design, if confirmed by further research, could recover behavioral gains from each strategy in the conditions where it performs best, at low marginal cost relative to a fixed single-strategy approach.

R^2 values are low across Tables 1–3 (0.003–0.079) because our controls capture only time-varying, household-invariant factors and omit household-level covariates such as size, tenure, or homeownership. This does not bias β given random assignment but leaves cross-household variance unexplained. Future work should collect household-level data at intake to improve model fit. The low contamination rates in this sample, averaging 2 to 3 percent compared to a national average of 15 to 25 percent for single-stream programs, may reflect several factors including the self-selected nature of program volunteers and the prize incentive structure rewarding clean setouts. These factors may limit the statistical power available to detect treatment effects on contamination. Future studies targeting programs with higher baseline contamination rates would provide a stronger test of whether message framing can shift contamination behavior. The Ohio State University home game interaction in particular should be interpreted with caution, as only two post-treatment weeks contain a home game and the leave-one-week-out diagnostic (Appendix A Figure A3) shows the coefficient is sensitive to which week is excluded. Point-of-disposal signage has recently been shown to reduce recycling contamination directly at the moment of disposal [22], suggesting that message framing delivered through email may need to be paired with in-the-moment interventions to meaningfully move contamination rates in low-baseline settings like this one. The external validity of the findings is also constrained by the study population and setting. The sample consists of households that voluntarily enrolled in a dedicated recycling program in a university neighborhood, a population that likely differs systematically from the broader population of households with passive access to curbside recycling infrastructure in terms of recycling motivation, residential stability, and responsiveness to social comparison information. Because enrollment in Can Fairy requires active sign-up, most participating households likely share some baseline level of recycling engagement prior to treatment assignment, which may partly explain why neither messaging strategy produced a detectable average effect relative to the other. This interpretation aligns with evidence that norm-based nudges tend to be most effective among individuals with weaker prior pro-environmental motivation [23], suggesting that a sample already selected for engagement may leave less room for norm-based framing to outperform educational content on average. Whether the equivalence result and the contextual interaction patterns documented here generalize to less engaged populations, more stable residential neighborhoods, or larger-scale municipal programs remains an open question. More broadly, future research employing three-arm designs that include no-message control alongside educational and norm-based treatment conditions would allow identification of the absolute and relative effects of each strategy and would provide the stronger empirical foundation that municipal recycling program administrators require when allocating outreach resources across competing messaging approaches. This need is underscored by a recent systematic review of experimental waste-reduction interventions, which finds that the evidence base remains fragmented across intervention types, behaviors, and measurement approaches, limiting the field's ability to draw generalizable conclusions about which message design choices work best in which settings [24].

4. Data and Methods

4.1. Program Setting and Study Population

The Can Fairy Curbside Recycling Program is a voluntary household recycling initiative administered by the University District Organization (UDO) in Columbus, Ohio, and funded by the Ohio Environmental Protection Agency with a matching grant from The Ohio State University. The program operates during the Ohio State University academic year, enrolling up to 200 households residing in the University District annually. Upon joining, each household receives a dedicated recycling bin collected most Mondays by paid UDO staff. Collected materials are transported to a recycling drop-off facility through a partnership with the City of Columbus and the Solid Waste Authority of Central Ohio. Households that set out clean loads (regardless of group assignment) are eligible for occasional prize drawings, providing a modest material incentive for participation. The study population consists of 200 households enrolled in the 2025–2026 program year. While these households are in an area adjacent to the University, these residences are not official university housing, and household members need not be university students.

4.2. Experimental Design and Randomization

The Can Fairy program constitutes a randomized field experiment in which enrolled households were randomly assigned by UDO staff using a computer to one of two weekly email message conditions, with 100 households per arm. From the start of the program in September 2025, the designated contact for all enrolled households received weekly emails containing educational content about correct recycling practices along with recycling-themed memes, as well as a Sunday reminder about Monday collection. Email messaging (versus the use of social media posts or other modes) decreases the probability that treatment messages were shared with control households. The protocol was reviewed and determined exempt by the Ohio State University Institutional Review Board (Study20252222, 7 October 2025) with informed consent obtained from all survey participants (survey data to be reported elsewhere). The analysis plan was publicly registered prior to deployment of differentiated messaging (https://www.socialscienceregistry.org/trials/17040 (accessed on 23 June 2026)).

Bin data collection also began in September 2025, providing approximately seven weeks of pre-intervention data during which both arms received identical educational messaging. On 1 November 2025, the treatment was introduced by separating households into two distinct messaging conditions. The educational arm received weekly emails with strictly educational framing, providing factual guidance on what items are recyclable, how to avoid contamination, and how to properly prepare materials for collection. The norm-based arm received weekly emails with strictly norm-based framing, delivering the same informational content alongside social comparison information highlighting neighborhood participation rates and emphasizing shared expectations and community pride in recycling. The neighborhood participation figures shown to the norm-based arm were updated weekly, calculated from the preceding week's bin audit data collected by UDO staff, rather than held constant across the intervention period or presented only qualitatively. The norm-based treatment therefore combined descriptive norm content, the weekly participation figures, with injunctive and community-identity elements, the shared expectation and community pride language, rather than isolating a single normative channel, and we treat it as a composite manipulation throughout this paper. Because both arms received identical messaging prior to November 1, any observed differences in recycling outcomes between groups after the intervention started are attributable solely to the differential message framing introduced at that point. Emails were sent weekly through April 2026, covering the full Ohio State University academic year. All households in both arms received identical recycling services, including weekly Monday collection, a dedicated recycling receptacle, and eligibility for prize drawings. Figure 2 summarizes the study timeline and experimental design while Figure 3 provides an example of the contrasting email messages received between the two treatment arms.

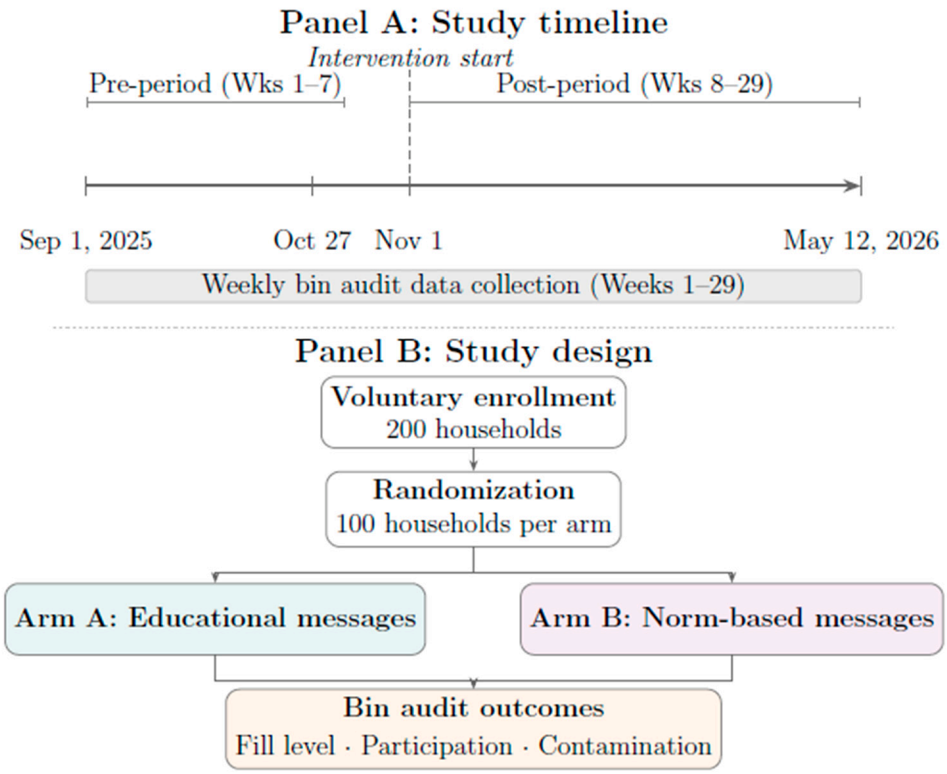

Figure 2. Study timeline and experimental design of the Can Fairy Curbside Recycling Program.

4.3. Outcome Measures

Recycling behavior is measured using three objective outcomes constructed from weekly bin-level observations recorded by UDO staff during curbside collection. For each household in each week, enumerators recorded one of six observations: clean bins at 25, 50, 75, or 100 percent fill capacity; contaminated (fill level not denoted); or not set out. The first outcome, fill level, captures the fraction of the bin filled with recyclable material on a zero to one scale, where clean fill levels are coded at their observed quartile values and bins that were not set out, set out empty, or contaminated are coded as zero. This provides a measure of participation intensity. The second outcome, participation, is a binary indicator equal to one if the household set out their bin with any material in a given week, whether clean or contaminated, and zero otherwise, including weeks in which the bin was not set out. This provides a critical indicator of program compliance. The third outcome, contamination, is a binary indicator equal to one if the household set out a contaminated bin in a given week, conditional on the bin being set out with material. This provides another critical indicator of program compliance.

Educational Frame	Social Norm Frame
Subject: Quick tip: recycle only empty containers Hi @firstname, Clean collections are automatically entered into weekly drawings and monthly prize packs. **This week's recycling tip:** Make sure containers are completely empty before putting them in the bin. Even small amounts of food or liquid can contaminate an entire batch of recycling. **Accepted items:** clean bottles, jars, cans, tubs, cardboard, and paper. **Not accepted:** food scraps, liquids, plastic bags, or Styrofoam. If you have a missed pickup or a question, please use this form: [Submit feedback here] For more details on what can be recycled in Columbus, visit:Recycle Right—Recycle Right Thanks for helping keep collections clean and making recycling easier to manage. Recycle On, Your Can Fairy team	**Subject:** Your neighbors are recycling smart—are you doing the same? Hi @firstname, Clean collections are essential for our program's operation. **This week's recycling tip:** Make sure containers are completely empty before putting them in the bin. Even small amounts of food or liquid can contaminate an entire batch of recycling. Your neighbors have been doing a great job emptying their containers before putting them in their box. Are you doing the same? **Accepted items:** clean bottles, jars, cans, tubs, cardboard, and paper. **Not accepted:** food scraps, liquids, plastic bags, or Styrofoam. Your neighbors are counting on you for clean, accepted materials in your recycling box every week. Recycle On, Your Can Fairy team

Figure 3. Example of educational (**left**) and social-norm (**right**) message framing for weekly email.

Observations where the bin was not set out or was set out empty are coded as missing for the contamination outcome and excluded from contamination regressions via listwise deletion. Across 200 households and 29 collection weeks, the panel comprises 5800 possible household-week observations. One household is missing bin-level observations for weeks 1 through 4, a data collection gap confined to the pre-treatment period, reducing the analytic sample for fill level and participation to 5796 observations. The panel comprises 4152 observations for contamination, reflecting the additional exclusion of weeks in which a bin was not set out or was set out empty. Because treatment assignment may itself affect the probability that a household sets out material in a given week, conditioning the contamination outcome on set-out behavior means the composition of the estimation sample may differ systematically by treatment status in ways not accounted for by random assignment alone. We therefore interpret the contamination results reported in Section 2.3 as conditional associations among participating household-weeks rather than as unconditional treatment effects on contamination across the full sample.

4.4. Analytical Strategy

We estimate the causal effect of norm-based versus educational message framing on recycling behavior using a difference-in-differences design that exploits the randomized assignment of households to treatment conditions and the availability of pre-intervention bin data. The estimating equation is

$$Y_{it} = \alpha + \beta(Treated_i \times Post_t) + \gamma Treated_i + \delta Post_t + X_{it}\theta + \varepsilon_{it}$$

where Y_{it} is the outcome for household (i) in week (t), $Treated_i$ is an indicator equal to one for households assigned to the norm-based messaging treatment, and $Post_t$ is an indicator equal to one for weeks 8 through 29 following the 1 November 2025 intervention start. The interaction term, ($Treated_i \times Post_t$), represents the difference in differences estimator, and its coefficient, β, is the parameter of interest. The vector X_{it} includes time varying covariates, specifically average weekly temperature, a bad weather indicator equal to one when weekly precipitation exceeded 25.4 mm or precipitation occurred on three or more days during the week, an indicator for weeks containing an Ohio State University home football game, and an indicator for weeks containing a major holiday or academic break. Weather data were obtained from National Oceanic and Atmospheric Administration (NOAA) Automated Surface Observing System (ASOS) records for Columbus Airport (station CMH) [25]. Average weekly temperature is a continuous measure with no threshold, and its coefficient reflects the marginal association between a one-degree increase in weekly average temperature and the outcome. θ are parameters to be estimated in relation to the vector **X** while ε_{it} represents unobservable factors driving the dependent variable. Standard errors are clustered at the household level in all specifications. We treat precipitation, rather than temperature, as the primary proxy for participation cost, since heavier or more frequent precipitation directly increases the physical burden and discomfort of carrying a recycling bin to the curb, while temperature is better understood as a general seasonal control rather than a discrete cost shifter. Because participation and contamination are binary outcomes, all specifications for these outcomes are estimated as linear probability models rather than nonlinear models such as logit or probit, consistent with our use of fixed effects and interaction terms, which are more directly interpretable in a linear framework.

We do not include household fixed effects in our primary specifications. Because treatment assignment is fixed for each household across the full panel, household fixed effects would absorb the main effect of treatment, $Treated_i$, which is collinear with time-invariant household identity. The coefficient of interest, β, on the interaction term $Treated_i \times Post_t$, remains identified under household fixed effects, since treatment status combined with the post period indicator still varies within household over time. We nonetheless retain our primary specifications without household fixed effects, both for consistency across models that report separately estimated treatment and post-period main effects, and because household fixed effects require the assumption that unobserved household-specific heterogeneity is time-invariant across the full 29-week panel, an assumption that becomes harder to sustain as panel length grows [26].

We estimate three specifications for each outcome. Model 1 includes only the treated, post, and interaction terms with no additional controls. Model 2 adds the time-varying covariates described above, allowing the effect of each contextual factor to be directly interpreted. Model 3 extends Model 2 by making each covariate interact with the treatment-period indicator, allowing the treatment effect to vary across weather conditions, campus events, and holiday periods. The pre-period balance checks confirm that fill level (p = 0.89), participation (p = 0.54), and contamination rates (p = 0.83) were statistically indistinguishable across treatment arms prior to the intervention, supporting the parallel trends assumption underlying the difference-in-differences design.

4.5. Robustness Check

We conduct two additional diagnostics to assess the robustness of the contextual interaction results from Model 3. The first, which we refer to as the alternative threshold specification check, tests whether the treatment by bad weather interaction depends on the specific cutoff used to define bad weather in the main specification. The main specification codes a week as bad weather if total precipitation reached 25.4 mm or precipitation occurred on at least three days that week. We re-estimate Model 3 under five alternative definitions, independently varying the precipitation amount threshold (19.05, 25.4, and 38.1 mm) and the minimum precipitation day count (three or four days), while holding all other aspects of the specification fixed. For each alternative definition, we recompute the bad weather indicator directly from the underlying weekly precipitation totals and day counts, then refit Model 3 and record the coefficient, standard error, and confidence interval for the treatment by bad weather interaction term. This allows us to determine whether the finding is being driven by a genuine contextual pattern in the data or by the particular threshold chosen for the main analysis.

The second, which we refer to as the leave-one-week-out diagnostic, tests whether any single collection week disproportionately drives the treatment interaction estimates from Model 3. We sequentially re-estimate Model 3 twenty-nine times, each time excluding one of the twenty-nine collection weeks from the analytic sample and refitting the full model on the remaining weeks. We then compare each of the resulting interaction coefficients to the corresponding coefficient from the full sample, expressing the difference in units of the full sample standard error. This diagnostic is particularly informative in this setting because several contextual conditions occur in only a small number of weeks within the post-treatment period. Only two post-treatment weeks fall below the bad weather threshold under the main specification, and only two post-treatment weeks contain an Ohio State University home game, meaning that each of these interaction terms is effectively identified from a small number of contextual weeks. Excluding any one of these weeks can therefore produce a disproportionately large change in the corresponding coefficient, which the leave-one-week-out diagnostic is designed to detect. Full results and figures for both diagnostics are reported in the Appendix A, with a summary interpretation provided in Section 2.

5. Conclusions

Using bin-level audit data from a randomized field experiment among 200 voluntarily enrolled households in Columbus, Ohio, we find that norm-based and educational message framing produce statistically indistinguishable average effects on recycling fill level, participation, and contamination over a 29-week program. The interactive specifications add an important qualification: norm-based messaging outperforms educational messaging when adverse weather raises participation costs, while educational messaging holds up better during high-distraction campus events. These patterns point toward adaptive messaging designs that condition content on observable contextual triggers, as the most promising direction for extracting greater behavioral returns from recycling outreach investment.

Author Contributions: Conceptualization, S.B. and B.E.R.; methodology, S.B. and B.E.R.; investigation, S.B. and B.E.R.; resources, B.E.R.; data curation, S.B.; writing—original draft preparation, S.B.; writing—review and editing, S.B. and B.E.R.; visualization, S.B.; supervision, B.E.R.; project administration, S.B. and B.E.R.; funding acquisition, B.E.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Van Buren program of Ohio State University's Van Buren fund (no grant number).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ohio State University Institutional Review Board (protocol code Study20252222; date of approval 7 October 2025).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are publicly available in GitHub at is available at https://tinyurl.com/39r3hk98 (accessed on 23 June 2026).

Acknowledgments: The authors thank Nora Gerber and Dominique Hadad for providing the anonymized data and for collaborating on message creation. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

https://doi.org/10.3390/recycling11080136

Appendix A

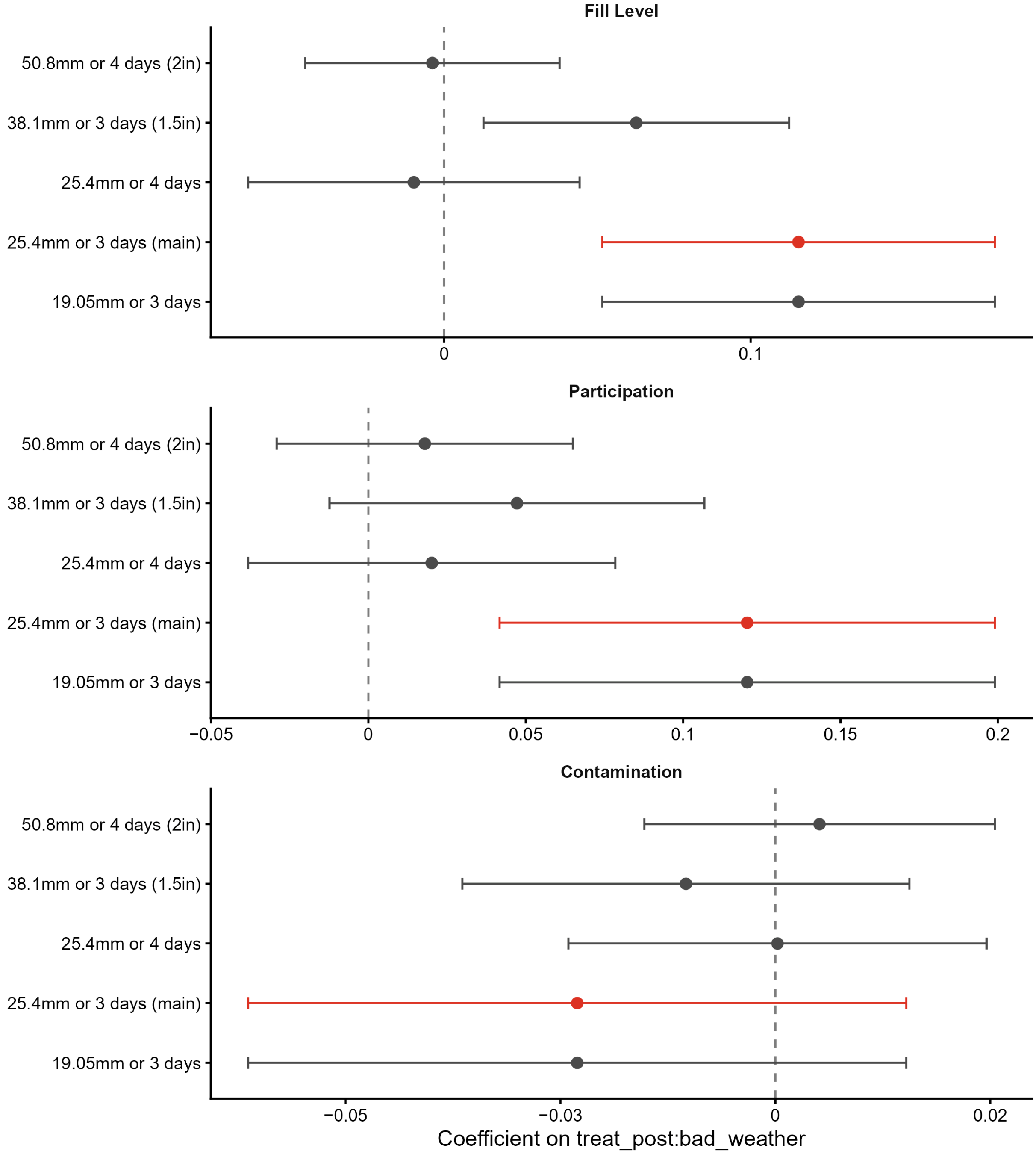

Figure A1. Sensitivity of the Treated × Post × Bad Weather Interaction to Alternative Bad Weather Threshold Definitions. Note: Coefficient estimates (with 95% confidence intervals) on the Treated × Post × Bad Weather interaction term from Model 3, reestimated under five alternative definitions of bad weather, independently varying the precipitation amount threshold (19.05, 25.4, 38.1 mm) and the minimum precipitation day count (3 or 4 days). The main specification (25.4mm or 3 days) is highlighted in red. Panels correspond to fill level (**top**), participation (**middle**), and contamination (**bottom**).

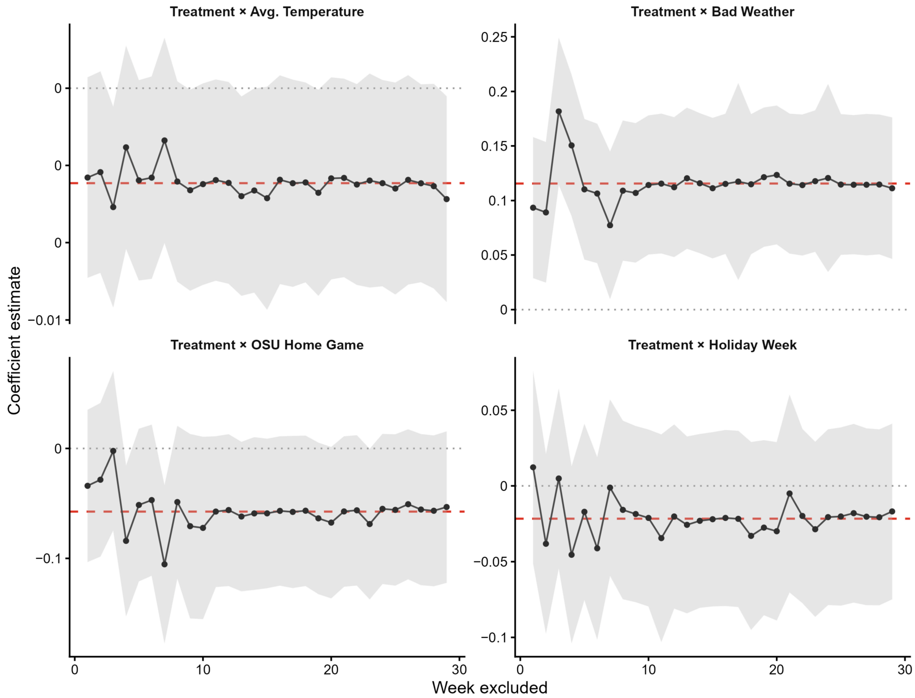

Figure A2. Leave-One-Week-Out Sensitivity of Treatment × Contextual Covariate Interactions, Fill Level (Model 3). Note: Leave-one-week-out sensitivity of the four Treated × Post × [covariate] interaction coefficients from Model 3, fill level outcome (Table 1). Each point shows the reestimated coefficient after sequentially excluding one of the twenty-nine collection weeks; shaded bands show 95% confidence intervals from each reestimation. The red dashed line marks the full-sample coefficient from Table 1, Model 3.

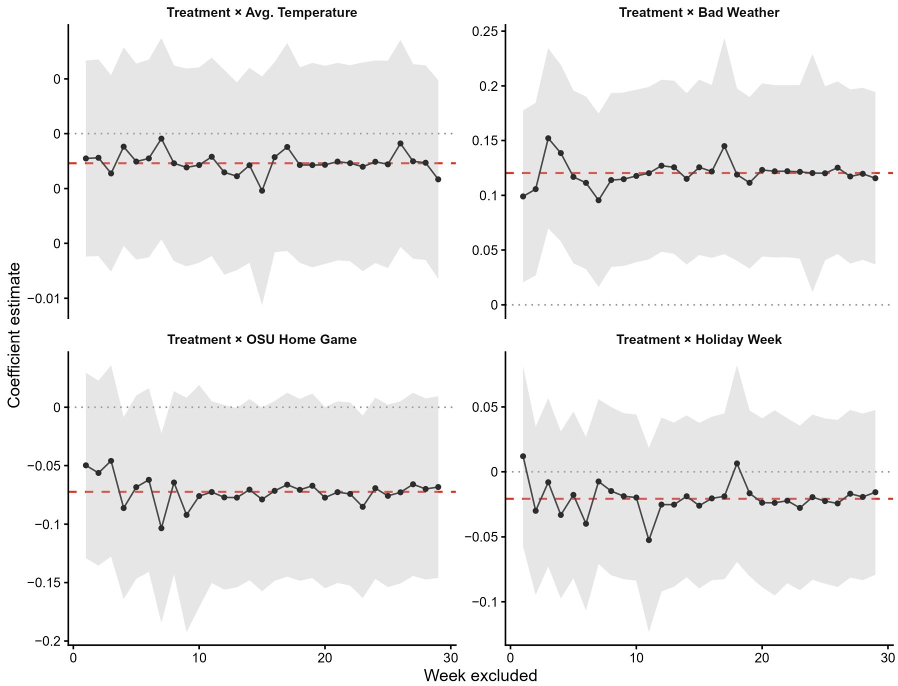

Figure A3. Leave-One-Week-Out Sensitivity of Treatment × Contextual Covariate Interactions, Participation (Model 3). Note: Leave-one-week-out sensitivity of the four Treated × Post × [covariate] interaction coefficients from Model 3, participation outcome (Table 2). Each point shows the reestimated coefficient after sequentially excluding one of the twenty-nine collection weeks; shaded bands show 95% confidence intervals from each reestimation. The red dashed line marks the full-sample coefficient from Table 2, Model 3.

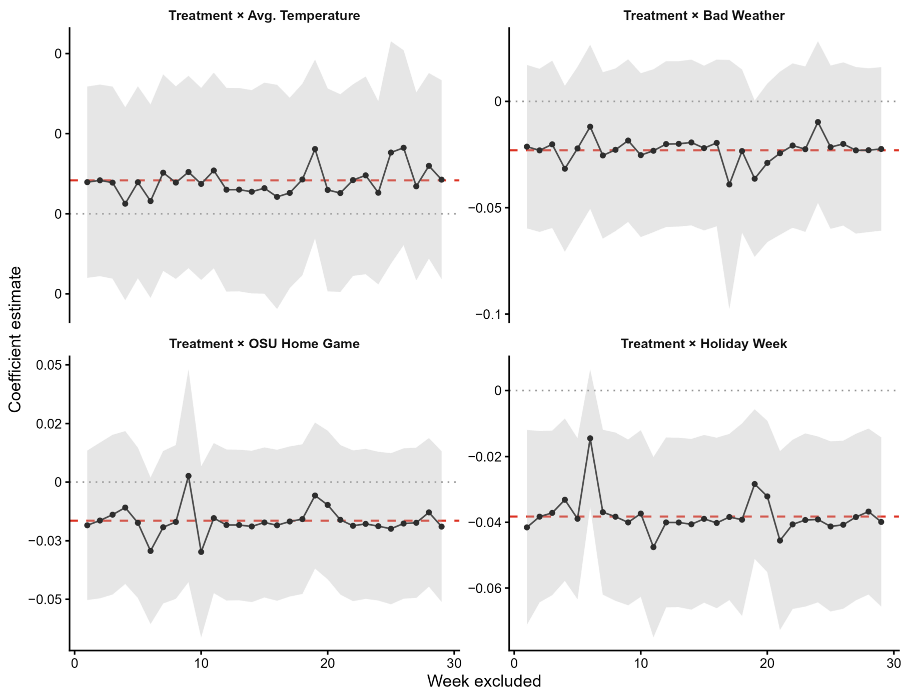

Figure A4. Leave-One-Week-Out Sensitivity of Treatment × Contextual Covariate Interactions, Contamination (Model 3). Note: Leave-one-week-out sensitivity of the four Treated × Post × [covariate] interaction coefficients from Model 3, contamination outcome (Table 3). Each point shows the reestimated coefficient after sequentially excluding one of the twenty-nine collection weeks; shaded bands show 95% confidence intervals from each reestimation. The red dashed line marks the full-sample coefficient from Table 3, Model 3.

References

1. Kinnaman, T.C. Understanding the economics of waste: Drivers, policies, and external costs. *Int. Rev. Environ. Res. Econ.* **2016**, *8*, 281–320. [CrossRef]
2. Turner, D.A.; Williams, I.D.; Kemp, S. Greenhouse gas emission factors for recycling of source-segregated waste materials. *Resour. Conserv. Recycl.* **2015**, *105*, 186–197. [CrossRef]
3. U.S. Environmental Protection Agency. “America Recycles Day”. 2025. Available online: https://www.epa.gov/circulareconomy/america-recycles-day (accessed on 18 July 2026).
4. Allcott, H. Social norms and energy conservation. *J. Public Econ.* **2011**, *95*, 1082–1095. [CrossRef]
5. Ferraro, P.J.; Price, M.K. Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Rev. Econ. Stat.* **2013**, *95*, 64–73. [CrossRef]
6. Liu, Z.; Yang, J.Z. Predicting recycling behavior in New York State: An integrated model. *Environ. Manag.* **2022**, *70*, 1023–1037. [CrossRef] [PubMed]
7. Schultz, P.W. Changing behavior with normative feedback interventions: A field experiment on curbside recycling. *Basic Appl. Soc. Psychol.* **1999**, *21*, 25–36. [CrossRef] [PubMed]
8. Dannenberg, A.; Gutsche, G.; Batzke, M.C.L.; Christens, S.; Engler, D.; Mankat, F.; Möller, S.; Weingärtner, E.; Ernst, A.; Lumkowsky, M.; et al. The Effects of Norms on Environmental Behavior. *Rev. Environ. Econ. Policy* **2024**, *18*, 165–187. [CrossRef] [PubMed]
9. Burn, S.M.; Oskamp, S. Increasing community recycling with persuasive communication and public commitment. *J. Appl. Soc. Psychol.* **1986**, *16*, 29–41. [CrossRef]
10. Burn, S.M. Social psychology and the stimulation of recycling behaviors: The block leader approach. *J. Appl. Soc. Psychol.* **1991**, *21*, 611–629. [CrossRef]
11. Wang, J.; Gu, Y.; Xin, H.; Wang, X. Influence of Appeal Type and Message Framing on Residents’ Intent to Engage in Pro-Environmental Behavior. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15431. [CrossRef] [PubMed]
12. Aravind, A.; Mishra, S.; Das, S.R. Message framing and sustainability awareness: Experimental evidence for advancing reduce, reuse, and recycle practices. *Ecol. Econ.* **2026**, *248*, 109061. [CrossRef]
13. Li, Y.; Yang, D.; Liu, Y. The effect of message framing on consumers’ intentions to purchase recycling-aiding products in China. *Sustainability* **2021**, *13*, 6966. [CrossRef]

14. Zhang, S.; Xia, Z.; Zhang, C.; Tian, X.; Xie, J. Green illusions in self-reporting? Reassessing the intention-behavior gap in waste recycling behaviors. *Waste Manag.* **2023**, *166*, 171–180. [CrossRef] [PubMed]
15. Czajkowski, M.; Zagórska, K.; Hanley, N. Social norm nudging and preferences for household recycling. *Resour. Energy Econ.* **2019**, *58*, 101110. [CrossRef]
16. Sorkun, M.F. How do social norms influence recycling behavior in a collectivistic society? A case study from Turkey. *Waste Manag.* **2018**, *80*, 359–370. [CrossRef] [PubMed]
17. Wichmann, B.; Luckert, M.; Bissonnette, K.; Cumberland, A.; Doll, C.; Gupta, T.; Shi, Y. Using pro-environmental information to modify conservation behavior: Paper recycling and reuse. *Recycling* **2017**, *2*, 5. [CrossRef]
18. Helferich, M.; Thøgersen, J.; Bergquist, M. Direct and mediated impacts of social norms on pro-environmental behavior. *Glob. Environ. Change* **2023**, *80*, 102680. [CrossRef]
19. Florence, E.S.; Fleischman, D.; Mulcahy, R.; Wynder, M. Message framing effects on sustainable consumer behaviour: A systematic review and future research directions for social marketing. *J. Soc. Mark.* **2022**, *12*, 623–652. [CrossRef]
20. Cialdini, R.B.; Reno, R.R.; Kallgren, C.A. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* **1990**, *58*, 1015–1026. [CrossRef]
21. The Recycling Partnership. State of Curbside Recycling Report. 2020. Available online: https://recyclingpartnership.org/wp-content/uploads/dlm_uploads/2020/02/2020-State-of-Curbside-Recycling.pdf (accessed on 18 July 2026).
22. Cakanlar, A.; Hunter, M.; Nenkov, G.Y. Recycle right: How to decrease recycling contamination with informational point-of-disposal signage. *J. Acad. Mark. Sci.* **2025**, *53*, 781–803. [CrossRef]
23. Ling, M.; Xu, L.; Yang, H. Direct and Spillover Effects of Social Norm Nudges for Household Recycling: A Longitudinal Field Experiment. *Sustain. Prod. Consum.* **2023**, *42*, 423–433. [CrossRef]
24. Wilson, B.M.; Delmas, M.A.; Rajagopal, D. Behavioral interventions for waste reduction: A systematic review of experimental studies. *Front. Psychol.* **2025**, *16*, 1561467. [CrossRef] [PubMed]
25. National Oceanic and Atmospheric Administration. Automated Surface Observing System (ASOS) Data, Columbus Airport Station (CMH). Available online: https://www.ncei.noaa.gov/products/land-based-station/automated-surface-weather-observing-systems (accessed on 28 May 2026).
26. Millimet, D.L.; Bellemare, M.F. On the (Mis) Use of the Fixed Effects Estimator. *Oxf. Bull. Econ. Stat.* **2026**, *88*, 484–505. [CrossRef]

https://doi.org/10.3390/recycling11080136