MIT CEEPR
Center for Energy and Environmental Policy Research

Working Paper Series

Small Changes, Big Impact: Nudging Employees Toward Sustainable Behaviors

Laura Cappellucci, Lan Ha, Jeremy Honig, Christopher R. Knittel, Amy Vetter and Richard Wilner

OCTOBER 2024

CEEPR WP 2024-17

Working Paper Series.

Small Changes, Big Impact: Nudging Employees Toward Sustainable Behaviors

Laura Cappellucci Lan Ha Jeremy Honig Christopher R. Knittel
Amy Vetter Richard Wilner*

October 29, 2024

Abstract

We designed and conducted three randomized control trials in partnership with a large biopharmaceutical company operating over 160 plasma donation centers, with the aim of promoting sustainable behaviors in a workplace setting. Specifically, we focused on reducing operational errors that led to dropped collection materials, long freezer door open times, and improper recycling practices. To achieve these goals, we employed social norms to nudge employees towards 1) reducing wasted collection materials, 2) minimizing the duration of freezer door openings, and 3) improving recycling practices. We found an average reduction of roughly 70 percent in plastic waste from dropped collection materials and the costs associated with these materials. The frequency of freezer door alarms decreased by over 80 percent, and the duration of alarms decreased by over 45 percent, depending on the empirical specification. We also observed a roughly 40 percent reduction in uncollapsed cardboard, with no statistically significant results for other types of contaminants. Importantly, for each of the interventions, we do not find evidence that the treatment effects waned over time or affected business operations. Our study provides significant implications for promoting sustainable behaviors in a workplace setting, filling an important gap in the literature on the effectiveness of nudges in the workplace.

Keywords: Corporate Sustainability Goals, Behavioral Nudges, Workplace Interventions
JEL Codes: Q40; D90; C90

---

*As is customary in economics, authors are listed alphabetically. Cappellucci, Honig, Vetter, and Wilner worked on this project as part of their day-to-day responsibilities at Takeda. Takeda funded the implementation of the experiments including expenses related to signage and training. Ha and Knittel did not take any funds from Takeda. AEA RCT ID: AEARCTR-0011118. Cappellucci: laura.adelman@takeda.com Ha: MIT, lanha@mit.edu; Honig: jeremy.honig@takeda.com Knittel: George P. Shultz Professor and Associate Dean for Climate and Sustainability Sloan School of Management, Director Center for Energy and Environmental Policy Research, Director MIT Climate Policy Center, and NBER, knittel@mit.edu. Vetter: amy.vetter@takeda.com Wilner: richard.wilner@takeda.com

1 Introduction

There has been a significant increase in firms making aggressive sustainability goals in recent years. These goals often include commitments to reduce greenhouse gas emissions, reduce plastic waste, and increase recycling. One strategy for meeting these targets is to invest in cleaner technologies. The International Energy Agency reports that the world is on track to spend a record USD 1.7 trillion on green energy investment (IEA, 2023). The U.S. Sustainable Investment Forum's Report on U.S. Sustainable Investing Trends identifies a staggering $8.4 trillion in sustainable investment assets under management at the beginning of 2022 (SIF, 2022).

In addition to investing in technology, a second and complementary strategy may be to nudge employee behavior to reduce waste. A long literature has shown that low-cost nudges, such as offering comparative feedback, can be a cost-effective way to reduce household energy consumption and waste. For instance, Wilhite and Ling (1995) conducted a comprehensive three-year study in Norway, revealing a 10% energy savings resultant from more informative billing. Additionally, Staats et al. (2004) demonstrated energy savings in electricity and gas through information exchange among groups of neighbors. The persistence of reduced electricity and natural gas consumption for 7-12 months following information provision was established by Ayres et al. (2013). Further, empirical evidence on home energy reports indicates short-term reductions of 2-2.9% in electricity consumption (Allcott, 2011; Ayres et al., 2013; Allcott and Rogers, 2014; Ferraro and Price, 2013; Jachimowicz et al., 2018; Henry et al., 2019). Recent research in Chinese production departments by Wu and Paluck (2021) illustrates the efficacy of floor decals in encouraging proper trash disposal.

To date, however, the literature on behavioral nudges has tended to focus on nudging household behavior with a few notable exceptions. In particular, Daamen et al. (2001) demonstrates that providing workshop garage managers with information on the environmental correctness of their subordinates' behaviors promoted more pro-environmental behaviors among subordinates. In an experiment conducted within academic buildings, Carrico and Riemer (2011) concluded that peer education and feedback could reduce energy consumption by 4% and 7%, respectively. Through field experiments within a commercial airline, Gosnell et al. (2020) showed that monitoring alone could improve captains' efficient fuel load by almost 8%. Furthermore, providing captains with recent performance information improved their fuel efficiency by almost 10%.

Despite contributing significantly to emissions, the workplace remains an under-explored domain for sustainability interventions. Challenges in fostering sustainable behavior arise from potential disincentives, workforce diversity, and a lack of awareness and motivation. Leveraging nudges in employee decision-making, interventions have proven effective in diverse areas, from motivating recyclable transport in universities (Needleman and Geller, 1992) to enhancing 401(k) participation through default option changes in Fortune 500 companies (Madrian and Shea, 2002). Notably, nudges, such as Save More Tomorrow, where workers pre-commit to saving for retirement (Thaler and Benartzi, 2004), demonstrate enduring impacts on participants' savings rates.

This paper begins to fill the environmental-workplace gap in the literature. We present the results of three randomized control trials (RCTs) that assess nudge interventions' effectiveness in promoting sustainability within the workplace. Collaborating with a major biopharmaceutical company operating over 160 plasma donation centers, our experiments targeted electricity consumption, plastic waste, and recycling, employing a multifaceted approach comprising metrics tracking, education, signage, and communication.

The interventions produced substantial and economically significant behavioral changes, with no evidence of waning treatment effects over time. Plastic waste reduction ranged from 48 to 97%, accompanied by cost reductions of 41 to 85%. Freezer door alarms decreased by 44 to 93%. We also find evidence that uncollapsed cardboard recycling fell by 20 to 41%; however, the event study evidence on this result is not strong. As a whole, these results underscore the potential for sustained impact on workplace behavior through carefully designed nudges.

While we experimentally estimate the benefits of the treatments on behaviors that impact sustainability, an obvious question is whether the interventions negatively affect employee morale. We collected non-experimental data on what employees thought of the interventions through a survey of center managers and assistant managers; the data are non-experimental because managers self-selected to complete the survey. We received responses from 46% of the first experiment treated centers (plastic waste), 27% from the second experiment treated center (freezer doors), and 30% from the third experiment treated centers (recycling), with some centers responses from multiple people. We use all of the responses for what follows, but similar lessons are learned if we restrict to center managers.

In response to the question "How would you rate the ease of implementing the BioLife Green Project in your everyday work?" we find that on a one-to-six scale, one being most difficult and six being most easy, the average response from the interventions were 3.7, 5.0, and 5.25 for the recycling, plastic waste, and freezer door experiments, respectively. This suggests that the interventions for which we found the largest treatment effects were the easiest to implement from the employees' perspective.

This study contributes empirical evidence to the effectiveness of nudging interventions in fostering sustainable behavior within the workplace. The findings hold relevance for corporations seeking cost-effective strategies to reduce environmental impact, emphasizing the pivotal role of behavioral nudges alongside technological investments. As the global business landscape grapples with escalating sustainability imperatives, this research offers a timely and pragmatic blueprint for cultivating lasting environmental stewardship within organizational contexts.

2 Empirical Context

2.1 BioLife's Green Project

The primary subject of this study is BioLife Plasma Services, a subsidiary of Takeda Pharmaceuticals. BioLife is an industry leader in collecting high-quality plasma, which is then processed into plasma-based therapies. The company now operates over 200 plasma collection facilities in the United States (161 during our experiment), with an additional 34 facilities located in Austria, the Czech Republic, and Hungary. Takeda has several sustainability goals, including a 5% reduction in water use by 2025, zero waste to landfill by 2030, net-zero emissions for scopes 1 and 2 before 2035, and net zero for scope 3 emissions by 2040. The operations of these BioLife plasma donation facilities are nearly identical, resulting from the company's focus on achieving cost efficiency and regulatory compliance. As such, BioLife provides an ideal platform for scalable behavioral interventions.

We identified three distinct behaviors that, if improved, would positively impact the company's sustainability objectives. These behaviors are present at different stages of BioLife's operation. The first behavior is in the plasma collection step, in which sterile collection materials known as "softgoods" must be discarded if they are dropped on the floor due to sterility and quality standards.[1] We hypothesized that providing a visible cue would encourage employees to handle these materials with greater care, leading to a reduction in the number of discarded materials and a decrease in plastic waste. To implement our intervention, we

[1] We note that employees are not incentivized to use dropped materials because it is against operating procedure and a form of Environmental Health Safety violation.

placed a clear plastic bin in an employee-only area and instructed employees to discard any dropped softgoods into it. We also tracked outcomes on "Tier 1 KPI" boards within the center. By doing so, the daily count of dropped materials became visible to employees, serving as a subtle yet constant visual reminder to be cautious and prevent softgoods from being discarded due to mishandling.

The second behavior we examined pertains to the freezing process of collected plasma in the center's walk-in freezer. Every plasma donation must be frozen within a defined period of time after collection, which leads the center employees to enter and exit the walk-in freezers within the centers multiple times a day to drop off collections. Leaving the freezer doors open can increase electricity consumption while decreasing its components' lifespan. Excessive door openings can also contribute to wear and tear on the refrigeration system, resulting in substantial costs to the company and can increase scope 1 emissions through the leaking of refrigerants. Our focus was on nudging employees to close the freezer doors within 50 seconds of opening them. We chose 50 seconds because that time had been previously tested as the maximum time necessary for a new employee to complete the task safely without compromising the plasma product. To encourage this behavior, in treated centers, we installed audible alarms that sounded after 30 seconds as a reminder to the staff, then again every 30 seconds with increasing volume until the door is closed. We also tracked outcomes on Tier 1 KPI boards.

The last behavior we aimed to influence involves recycling. If contaminant materials, such as plastic and styrofoam, are inappropriately disposed of in recycling bins, it not only increases landfill waste as these items cannot actually be recycled but also negatively affects BioLife's sustainability goals. Our goal was to encourage employees to recycle materials approved by local waste haulers. To accomplish this, we collaborated with a third-party provider and installed AI-enabled in-dumpster cameras in each recycling bin in control and treated centers. The provider's machine learning tool analyzed snapshots of the bins and generated a weekly count of contaminants by type and an estimated percent fullness of bins. We then relayed these counts to the facility manager, who translated them into Tier 1 KPI board metrics.

To evaluate our hypotheses, we designed three randomized controlled trials. Each experiment focused on a specific behavior involving softgoods drops, freezer door closure, and recycling. Collectively, these efforts are dubbed the "Green Project" within BioLife. Table 1 provides a summary of the treatments.

2.1.1 Pilot Experiment

We conducted two pilot studies, each corresponding to the first two behaviors described above. Our primary goals were to evaluate our approach's feasibility and ensure the smooth execution of the experiments without compromising the plasma donation process or donor experience. In the first pilot study, which focused on softgoods drops, we placed a softgoods drop bin and provided educational resources to employees at two centers. We then compared the number of softgoods dropped at these two centers against the remaining 150 centers (at the time) within Biolife's network. For the second pilot study, which centered on freezer door closures, we randomly selected a group of ten centers, in which we installed audible alarms in three centers, and two centers received educational material on freezer door closures. We monitored freezer door closure activities in all ten centers.

We observed that while messaging alone did not produce measurable desirable outcomes, the impacts of displaying softgoods drops and activating audible freezer door alarms were promising. The average treatment effect (ATE) in the softgoods pilot study was approximately 27%, while that in the freezer door alarm was close to 90%. Table C1 provides the estimated treatment effects derived from the pilot phase. Based on these estimates and the number of eligible BioLife centers, we performed power calculations and determined the duration of the full-scale experiments. We ensured that the treatment and control groups were equally balanced in terms of operational volume and hardware/software capabilities, with each treated center participating in only one experiment. We subsequently omitted the centers participating in the pilot studies from the RCT.

2.2 Experimental Design

The intervention at treated centers comprised five major elements: 1) intervention, 2) metrics tracking, 3) education, 4) signage, and 5) communication. The control group received no intervention, observed metrics tracking, education, signage, or communication. All treated centers received:

1. Observed Metrics tracking: A board tracks four to six "Tier 1 KPIs" that are leading indicators of success for BioLife's key business imperatives and is updated manually on a daily basis. We introduced a new experiment-specific metric to each center's metric board to keep the intervention at the forefront of center operations.

2. Education: A mandatory online training session was assigned to all center employees

two weeks before the experiment launch. The training was designed for the targeted behavior outcome and aimed to introduce employees to corporate environmental targets, program values, and the proposed intervention. It is important to note that the online training was the only component of the treatment that was not present throughout the entire experiment.

3. Signage: Posters were strategically placed in treated centers to increase employees' awareness and remind them of the desired behavioral outcome.

4. Communication: Treated centers are invited to kick-off calls, typically one week before the intervention. The primary purpose of the virtual kick-off calls is to provide center managers with the necessary instructions and resources to access daily data for tier board tracking. Regular check-in calls were scheduled on a weekly/monthly basis to enable center managers to stay informed of program status updates and ask relevant questions.

Examples of the online training materials and signage can be found in Figures D1 through D3.

2.3 Data

We observed center employees' behavior from August 2021 (Experiment 1), January 2022 (Experiment 2), March 2022 (Experiment 3) through November 2022, and the experimental window was from February (Experiment 1), March (Experiment 2) and May (Experiment 3) through September 2022. We tailored our data collection approach toward each experiment's outcome metric(s). In Experiment 1, we obtained a database capturing weekly softgoods drops from all participating centers, which required minimal processing. For Experiment 1, we observed $N = 4416$ center-week observations.

In Experiment 2, we utilized freezer door sensors already installed by the freezer manufacturer in most centers to record and timestamp opening and closing events of any door opening that lasted longer than 50 seconds for both the control and treated centers.[2] With these data, we calculated the frequency and duration of the freezer door openings across all centers. The data provided by the manufacturer were available in multiple formats. However, sporadic sensor malfunctions caused intermittent gaps in the data set, affecting many

[2]The time of 50 seconds was already a standard for employees as this was deemed sufficient time to put plasma in the freezer.

centers. To address these irregularities, we utilized the raw sensor data from the freezer manufacturer as a complementary source to ensure data consistency. Calculating the weekly average for each metric, we observed $N = 714$ center-week observations for the frequency of alarm metric (approximately 1.1% missing) and $N = 674$ (approximately 6.6% missing) center-week observations for the duration of alarm metric.

For Experiment 3, we collaborated with a third-party provider to install smart in-dumpster cameras in both control and treated centers. The third-party technology uses machine learning to identify contaminants in the dumpsters. We collected data on contaminants in recycling dumpsters, such as uncollapsed cardboard and plastic bags. It is important to note that camera installation took place between March 14 and April 20, 2022, due to unforeseen delays. As treatment started on May 1, 2022, pre-treatment data were more limited than initially planned. Moreover, the presence of cameras in both groups may have influenced employee recycling behavior, potentially leading to an underestimation of the treatment effect in Experiment 3. To ensure data quality, we conducted validation checks. During this process, we discovered that bin labels were sometimes misidentified as either "refuse" or "recycle," which could contribute to inaccurate counts of contaminants. We worked closely with the provider to address these issues, recognizing that these discrepancies may have impacted the results' accuracy. Additionally, we experienced instances where cameras occasionally dislodged during hauling or operation, resulting in data collection disruption for various weeks. We observed $N = 636$ center-week observations with 6.45% missing data.

We followed a sequential approach to assign centers into control and treatment groups for each experiment, primarily because of variations in data collection costs and practical feasibility. For Experiments 2 and 3, the installation of data collection equipment necessitated time and effort, whereas for Experiment 1, the data collection mechanism was already in place and straightforward to implement. To qualify for Experiment 2, a center required the necessary measurement equipment to have been installed by the freezer manufacturer. Experiment 3 required new equipment altogether. Therefore, we first randomly assigned control and treatment statuses to centers eligible for Experiment 2, followed by those eligible for Experiment 3. Finally, we allocated the remaining centers to Experiment 1. Throughout each randomization step, we followed a consistent methodology. We generated six distinct candidate assignments for each experiment for control and treatment status, conducted a joint orthogonality test on several observable characteristics, and ultimately chose the assignment candidate associated with the highest comparability between control and treatment groups. The joint orthogonality test allowed us to ensure that the assignment of treatment and control groups results in balanced and statistically comparable groups. Specifically, we considered characteristics on a center-month level, including the number of plasma donations, volume of plasma donations, square meter, electricity consumed per square meter, number of cooling degree days, and number of heating degree days.

Tables B1, B2, and B3 summarize the aforementioned center characteristics and experiment-specific metrics and tests for balance across treatment and control groups in each experiment. Columns 1 and 2 present the mean of each characteristic (with standard deviation in parentheses) for the treatment and control groups, respectively. Column 3 displays the treatment-control difference in means (and standard error in parentheses) for each characteristic as the coefficient from a regression of the particular characteristic on the treatment binary variable, with standard errors clustered at the center level. The tables also report an F-statistic for a joint test of significance. For Experiment 1 (Table B1), we do not find statistically significant differences in characteristics, although two observables, electricity consumption and electricity consumed per square meter, are marginally significant. However, treated centers consumed about 15% less electricity than control centers. The F-statistic, however, is statistically significant. In Experiment 2 (Table B2), the number and duration of freezer door alarms do not differ across treatment and control groups. However, the differences between treated and control centers in plasma donations and volume are marginally significant. These differences motivate including donation variables in the empirical specifications discussed below. The F-statistic is not statistically significant. In Experiment 3 (Table B3), we do not find statistically significant differences in the outcome variables or volumes on an individual basis, but the F-statistic is statistically significant.

3 Empirical Strategy

3.1 Panel Fixed Effects Approach

We estimate the causal impact of the interventions on the outcome metrics by estimating:

$$Y_{it} = \beta_0 + \beta_1 K_{it} + \beta_2 T_{it} + \eta X_{it} + \theta_i + \omega_t + \epsilon_{it} \tag{1}$$

where Y_{it} is the outcome metric at center i at week t. The treatment indicator T_{it} is a binary variable that takes a value of one for treated centers from the intervention start date onward. K_{it} is a binary variable taking a value of one for treated centers between the kick-off call and the intervention start date. We introduce K_{it} to test our hypothesis that treated center managers' awareness of the experiment during the kick-off call might lead to changes in employees' behaviors before the intervention's official launch date. X_{it} is a vector of center i's operation volume at week t. θ_i are center fixed effects. ω_t represents week or month fixed effects. ϵ_{it} is an error term. The coefficient of interest, β_2, reflects the average treatment effect of the intervention in the unit of the outcome metric. The coefficient β_1 measures whether there are any effects from announcing the experiment.

We present results from several specifications. In our baseline specification, we control for center fixed effects, accounting for time-invariant characteristics at each center, but do not include X_{it} or any time fixed effects. Our preferred fixed effects specification includes center fixed effects and month fixed effects, the latter to control for time trends, and X_{it} to control for differences in center operation volume levels. As a result, our identification comes from within-center and within-month differences between the treated and untreated centers. Across all specifications, we cluster standard errors at the center level to account for arbitrary within-center correlation.[3]

To help interpret the point estimates, we report the average value of the dependent variable in the pre-intervention period across all control centers. We then report the treatment effect by dividing β_2 by the baseline value.

To account for the skewed nature of center sizes, we utilized two methods:

1. We performed weighted regressions with weights corresponding to the center's average donation during the pre-period.

2. We included donation-related variables in our control variables.

In Experiment 2, we conducted winsorization on observations above the 99th percentile of the dependent variables. We adopted this approach with the assumption that these outliers are a consequence of sensor malfunctions and data reporting errors.

3.2 Event Study

We begin by showing results from an event study model to investigate the treatment effects over time and confirm balance in the trends of outcomes prior to the launch of the experi-

[3]We note that there are 19 centers in Experiment 2. Inference is robust to clustering the standard errors at a higher level, e.g., at the center-by-month level.

ments. Figures 1 through 6 provide graphical results from event study regressions for each experiment. The x-axis plots weeks before and after the kick-off call, with the week of the kick-off call normalized to zero. The dotted vertical line marks the start of the intervention. We plot the point estimates and 95% confidence intervals from a regression with our preferred set of fixed effects: center fixed effects and month fixed effects. Standard errors are clustered at the center level. The estimating equation is:

$$Y_{it} = \beta_0 + \sum_{w=a}^{b} \tau^w K_{it}^w + \eta X_{it} + \theta_i + \omega_t + \epsilon_{it} \tag{2}$$

where K_{it}^w is a binary variable equaling one if an observation is w weeks before (or after) the kick-off call, where $w \in [a, b]$. We measure treatment effects relative to $w = 0$.

Figure 1 plots the event study results for Experiment 1. The plots do not suggest any pre-existing trend differences across treatment and control. Following the launch of the experiment, softgoods drops began to fall in the treatment centers relative to the control centers. The treatment effect reaches its largest value at 18 weeks after the launch. The point estimate suggests roughly 13 fewer weekly drops in treatment centers. This is larger than the pre-experiment mean of 8.7. This is likely due to the skewness of the distribution of softgoods drops across centers and motivates the weighted regressions discussed above. For softgoods drops, the treatment effects appear to increase somewhat until week 18 but then stabilize at a point estimate of roughly five. Similarly, the cost associated with softgoods drops dropped to its highest level at week 18 after the launch, savings about \$33 or more than 130% of the pre-experiment mean of \$25.6.

Figures 2 and 3 plot the event study for Experiment 2. The first plots data on the duration of alarms, while the second plots the frequency of alarms. Unfortunately, we have only four weeks of data before the on-boarding call and an additional four weeks between the on-boarding call and the launch. With this caveat in mind, we do not find strong evidence of a pre-existing trend for either the duration or frequency of alarms. There is some, albeit weak, evidence that the frequency of alarms in treatment centers began to trend downward between the on-boarding call and the experiment's launch.

The duration and frequency of alarms fall after the launch of the experiment. The average treatment effect reaches its largest value at 30 weeks after the on-boarding call for the duration of alarms and 11 weeks after the on-boarding call for the frequency. At this point, there is some evidence of attenuation, but the point estimates remain negative. The peak negative effects represent roughly 124% and 89% of the mean for the frequency and duration, respectively.

Finally, we turn to Experiment 3. Here, the event studies are less conclusive. Figures 4 through 6 plot the event study results for Total Contaminants, Uncollapsed Cardboard, and Plastic Bag Contaminants. There is some evidence for a non-zero treatment effect for uncollapsed cardboard, but the first three weeks of the pre-period also exhibit fewer instances of uncollapsed cardboard in treatment centers relative to control centers. Therefore, one might worry that the perceived treatment effect is driven by a sudden increase in uncollapsed cardboard among treatment centers in the week just prior to the kick-off meeting.

4 Results

Table 3 displays the results from estimating Equation 1 with the dependent variable being both the count of softgoods drops and their corresponding dollar costs using five different sets of fixed effects. We estimate ten different specifications, increasing the set of control variables as the table from left to right and alternating between not weighting observations and weighting observations based on the pre-treatment center average volumes. In the first two specifications, we include only center fixed effects. We then add month fixed effects in specifications three and four. Specifications five and six replace the month fixed effects with week-of-sample fixed effects. In the final four specifications, we add the number of donations (specifications seven and eight) and the number and volume of donations (specifications nine and ten). In our most conservative estimation, softgoods drops are estimated to decrease by at least 48%, corresponding to a weekly decline of 11 softgoods. When utilizing pre-treatment weights in the specifications, it is anticipated that softgoods drops may decrease by as much as 97%, corresponding to a weekly decrease of 23 softgoods. Similarly, the dollar costs associated with softgoods drops are expected to decrease between 41 and 85%. Consequently, each treated center is estimated to save between \$3.29 and \$6.89 per week in softgoods costs.

Tables 4 and 5 display the results from estimating Equation 1 with the duration of freezer alarms and the frequency of freezer alarms, respectively, as the dependent variables. Each table includes the same ten specifications described above. Installing audible alarms led to an average reduction between 44 and 48% in the duration of daily alarms (about 6.6 to 7.2 minutes of freezer opening times). Furthermore, the intervention also led to an average reduction between 59 and 93% in the number of daily freezer door alarms, with our preferred specification—Column (10)—showing a 93% reduction.

In Experiment 3, the dependent variables consist of the weekly total number of contaminants, uncollapsed cardboard and plastic bags, the number of weekly recycling bins containing contaminants, and the percentage of contaminated pickups. Tables 6 through 10 display the results from estimating Equation 1 on each of the dependent variables. The findings indicate that the intervention did not lead to statistically significant effects, perhaps except for uncollapsed cardboard. Table 7 shows that sharing contaminant data led to a 20 to 41% reduction in uncollapsed cardboard, although the effects are not always statistically significant. In the preferred specification—Column (10)—the reduction of 41% is statistically significant at the 5% level. However, as noted above, one might be less confident in these results given the increase in uncollapsed cardboard within treatment centers just prior to the kick-off meeting.

5 Discussion

The goal of the experiments was to test whether behavioral nudges can increase the sustainability of Biolife's operations. Here, we discuss back-of-the-envelope estimates of the reductions in externalities associated with operations from our interventions. Using the estimated treatment effect of 70% from Experiment 1 and weekly softgoods drops of 4.2 lbs, we anticipate an annual decrease of 14 metric tons in plastic waste across 200 donation centers. Reducing plastic waste can positively impact ecosystems and the livelihood of humans and countless other species. Firstly, it conserves non-renewable energy since plastic production largely depends on fossil fuels (WEF, 2016). Secondly, it lowers energy use and associated greenhouse gas emissions during plastic disposal. Thirdly, it mitigates the risk to wildlife from ingestion or entanglement in plastic, reducing injuries and fatalities. Lastly, it contributes to public health by lessening microplastics in the food chain. Beaumont et al. (2019) estimated that "each tonne of plastic in the ocean has an annual cost in terms of reduced marine natural capital of between $3,300 and $33,000". Based on this estimation, the potential social benefits of saving 14 metric tons of plastic waste can range from $46,200 to $462,000 annually.[4]

From BioLife's previous study, low-temperature refrigeration systems supporting the

[4]This assumes that all of the waste would have made it to the ocean.

plasma freezer use 19% of the total electricity consumed on average and account for 23% of CO_2 emissions. While it is unclear how much electricity opening freezer door consumed, we estimate scaling the experiment to the whole network could save 8,500 hours of freezer door openings and lead to a reduction in energy consumption and carbon footprint.

Following Experiment 3, Biolife implemented a right-sizing effort in which about 42% of participating centers achieved cost savings through reducing waste pick-up frequency. By decreasing pick-ups from three to two times per week, these centers achieved an average cost savings of $110 per month. Scaling up this right-sizing effort to the whole network would result in estimated annual savings exceeding $110,000. We can calculate a back-of-the-envelope estimate of the greenhouse gas savings from the intervention. The US Department of Energy estimates that refuse trucks have an average fuel economy of 2.53 mpg. Additionally, if we assume each pick-up requires 2 miles of travel, the less-frequent dumpster servicing would conserve roughly 8,700 miles of travel per year and 3,500 gallons of fuel.[5] At 20 pounds of CO2 per gallon, this would reduce emissions by roughly 30 tonnes. Using the estimated social cost of carbon of $200 per metric ton from Pindyck (2019), our findings could result in a modest saving of $6,000 annually. Other notable co-benefits of reducing vehicle miles traveled are reductions in other air pollutant emissions, water pollution, and wildlife mortality (Fang and Volker, 2017). In this context, the private benefits of reducing pick-up frequency outweigh the social benefits.

There are a few caveats from the experiments. A limitation of Experiment 2 is the imbalance between the control and treatment groups. Some centers that were initially identified as being eligible for the intervention did not meet the necessary data requirements. As a result, the study had a limited number of control group centers, which, as noted above, could have introduced the potential for selection bias or unobserved confounding variables that may reduce result reliability. For Experiment 3, the presence of the camera could have altered the behaviors of centers in both the control and treatment groups. It is, therefore, important to carefully consider the impact of this potential source of bias when interpreting the results.

A final caveat is that the environmental and sustainability benefits of the program must be weighed versus any costs or burdens faced by employees. To gain some traction on this issue, we sent out a survey to center managers and assistant managers who were treated

[5]The vehicle miles traveled are based on a rough estimated mileage driven to service a dumpster based on industry data provided by our third-party partner.

across the three experiments. We asked the following questions:

- How would you rate the ease of implementing the BioLife Green Project in your everyday work
- How would you rate the ease of maintaining a sustainability KPI at Tier 1 board?
- How would you rate your level of engagement with the Green Project?
- How much impact do you feel you have on BioLife's sustainability goals?

We also asked whether the program increased their awareness of their own sustainability behaviors to gauge whether the interventions might have spillover benefits in other facets of the employees' lives. Given that we do not have responses to these questions from control-group centers, relative statements across the three interventions are most relevant. We received responses from 46% of the centers in the plastic goods experiment, 27% in the freezer door experiment, and 30% of centers in the recycling experiment.[6] The recycling intervention was the least easy to implement. The average ease-of-implementing score for recycling was 3.7, compared to 5.0 and 5.25 for the plastic waste and freezer door programs, respectively.[7] The ranking across programs differed for the ease of maintaining a sustainability KPI at Tier 1 board changes. In this case, the freezer door program had the lowest score (2.8) compared to 4.5 and 3.9 for the plastic waste and recycling programs, respectively.[8] The feeling of engagement was more homogeneous across the programs; the averages are 3.3, 3.9, and 3.4 across the freezer door, plastic waste, and recycling programs, respectively.

In terms of how the different programs affected the employees' feeling of engagement with Biolife's sustainability goals, the plastic waste program had the highest average (3.9) compared to averages of 3.25 and 3.4 for the freezer door and recycling programs, respectively. However, none of these differences are statistically significant.[9]

[6] This is the share of unique center responses. For some centers, we received multiple responses. In what follows, we use all of the response data.

[7] The recycling rating is statistically different from the two others at the 10-percent level.

[8] There is a statistical difference between the freezer door and plastic program at the 5-percent level.

[9] While not the focus of the study, there is some evidence that the program raised employee awareness of their own sustainability habits, at least for those who chose to fill out the survey. This is especially true of the plastic waste and recycling programs. For the plastic waste program, 71% said the program raised awareness of their own landfill habits (29% saying "very much"), 67% said the program raised awareness of their own recycling habits, and 71% said the program raised awareness of their own energy habits (23% saying "very much"). For the recycling program, 71%, 58% (42% saying "very much"), and 86% (14% saying "very much") said the program raised awareness of their own landfill, recycling, and energy habits, respectively. For the freezer door program, 50% said it raised awareness of their own landfill, recycling, and energy habits.

6 Conclusion

This paper addresses the growing importance of sustainability goals in firms. While significant attention has been given to technological investments, this study focuses on the often-overlooked potential of behavioral nudges, particularly in the workplace, as a complementary approach to achieving sustainability objectives.

The literature review establishes a foundation for the paper by highlighting successful nudging strategies in household energy consumption, waste reduction, and recycling. Despite the workplace being a substantial contributor to emissions, there is a dearth of research on nudges in professional settings. The introduction sets the stage for the paper's primary focus on three randomized control trials conducted in collaboration with a biopharmaceutical company operating numerous plasma donation centers.

The experiments, targeting plastic waste reduction, electricity consumption, and recycling behavior, employed a multi-faceted approach, including metrics tracking, education, signage, and communication. The results demonstrate significant and lasting impacts on employee behavior, with plastic waste reduction ranging from 48 to 97%, freezer door alarms decreasing by 44 to 93%, and uncollapsed cardboard recycling seeing a reduction of 20 to 41%.

These findings contribute to the broader literature by providing empirical evidence of the efficacy of nudging interventions in the workplace. The results not only reinforce the potential of behavioral nudges to drive sustainable practices but also emphasize the economic significance of these interventions.

The implications of this research extend beyond the specific context of plasma donation centers, offering valuable insights for businesses aiming to promote sustainability in the workplace. The demonstrated success of nudging interventions suggests a promising avenue for organizations seeking cost-effective strategies to reduce their environmental footprint. Moreover, the paper underscores the need for increased attention to behavioral interventions in the workplace, considering its significant contribution to overall emissions.

In a broader context, as firms continue to set ambitious sustainability goals, this research advocates for a holistic approach that combines technological advancements with targeted behavioral interventions. By acknowledging and addressing the challenges of promoting sustainable behavior in diverse workplaces, organizations can foster a culture of environmental responsibility and contribute meaningfully to global sustainability efforts. As the world witnesses unprecedented levels of investment in green technologies, this study serves as a timely reminder of the untapped potential within the workforce to drive positive change toward a more sustainable future.

References

Allcott, H. (2011): "Social norms and energy conservation," *Journal of Public Economics*, 95, 1082–1095.

Allcott, H. and T. Rogers (2014): "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review*, 104, 3003–3037.

Ayres, I., S. Raseman, and A. Shih (2013): "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage," *Journal of Law, Economics, and Organization*, 29, 992–1022.

Beaumont, N. J., M. Aanesen, M. C. Austen, T. BÃ¶rger, J. R. Clark, M. Cole, T. Hooper, P. K. Lindeque, C. Pascoe, and K. J. Wyles (2019): "Global ecological, social and economic impacts of marine plastic," *Marine Pollution Bulletin*, 142, 189–195.

Carrico, A. R. and M. Riemer (2011): "Motivating energy conservation in the workplace: An evaluation of the use of group-level feedback and peer education," *Journal of Environmental Psychology*, 31, 1–13.

Daamen, D. D. L., H. Staats, H. A. M. Wilke, and M. Engelen (2001): "Improving Environmental Behavior in Companies: The Effectiveness of Tailored Versus Nontailored Interventions," *Environment and Behavior*, 33, 229–248.

Fang, K. and J. Volker (2017): "Cutting Greenhouse Gas Emissions Is Only the Beginning: A Literature Review of the Co-Benefits of Reducing Vehicle Miles Traveled," Tech. rep., National Center for Sustainable Transportation.

Ferraro, P. J. and M. K. Price (2013): "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment," *Review of Economics and Statistics*, 95, 64–73.

Gosnell, G. K., J. A. List, and R. D. Metcalfe (2020): "The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains," *Journal of Political Economy*, 128.

Henry, M. L., P. J. Ferraro, and A. Kontoleon (2019): "The behavioural effect of electronic home energy reports: Evidence from a randomised field trial in the United States," *Energy Policy*, 132, 1256–1261.

IEA (2023): "World Energy Investment 2023," Tech. rep., International Energy Agency, Paris.

Jachimowicz, J. M., O. P. Hauser, J. D. OâBrien, E. Sherman, and A. D. Galinsky (2018): "The critical role of second-order normative beliefs in predicting energy conservation," *Nature Human Behaviour*, 2, 757–764.

Madrian, B. and D. Shea (2002): "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *The Quarterly Journal of Economics*, 117, 377–377.

Needleman, L. D. and E. S. Geller (1992): "Comparing interventions to motivate work-site collection of home-generated recyclables," *American Journal of Community Psychology*, 20, 775–785.

Pindyck, R. S. (2019): "The social cost of carbon revisited," *Journal of Environmental Economics and Management*, 94, 140–160.

SIF (2022): "US SIF "Trends Report" Documents Sustainable Investment Assets of $8.4 Trillion," .

Staats, H., P. Harland, and H. A. M. Wilke (2004): "Effecting Durable Change: A Team Approach to Improve Environmental Behavior in the Household," *Environment and Behavior*, 36, 341–67.

Thaler, R. and S. Benartzi (2004): "Save More Tomorrowâ¢: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, 112, S164–S187.

WEF (2016): "The New Plastics Economy: Rethinking the future of plastics," Tech. rep., World Economic Forum.

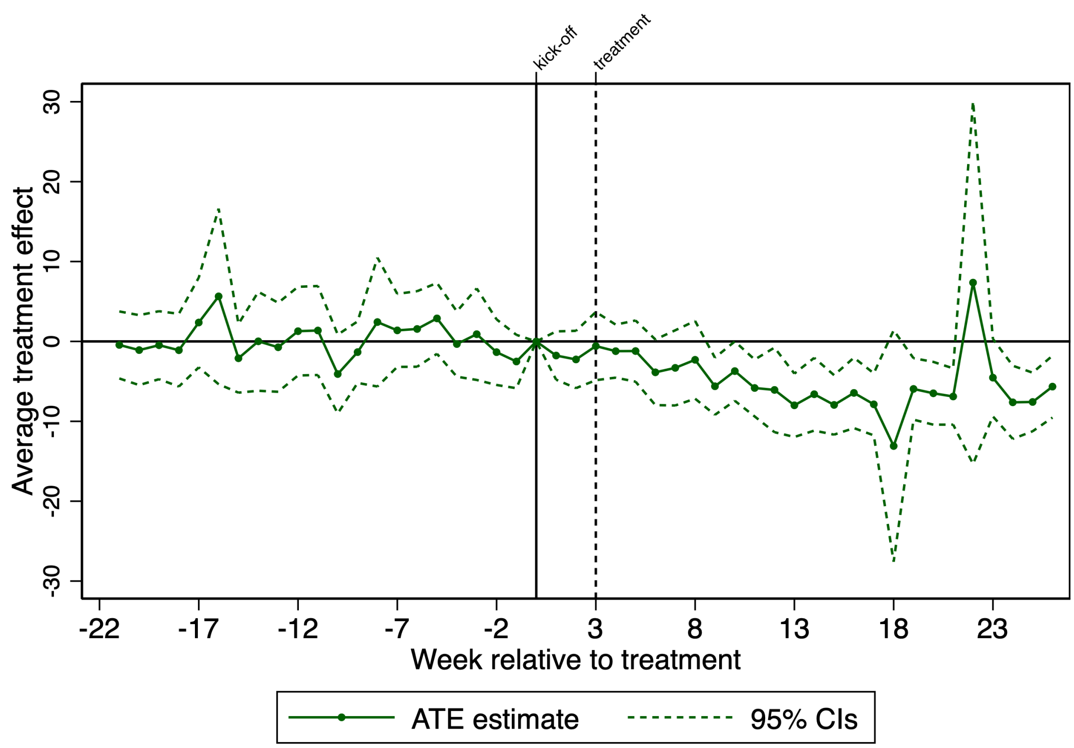
Wilhite, H. and R. Ling (1995): "Measured energy savings from a more informative energy bill," *Energy and Buildings*, 22, 145–155.

Wu, S. J. and E. L. Paluck (2021): "Designing nudges for the context: Golden coin decals nudge workplace behavior in China," *Organizational Behavior and Human Decision Processes*, 163, 43–50.

A Main Figures and Tables

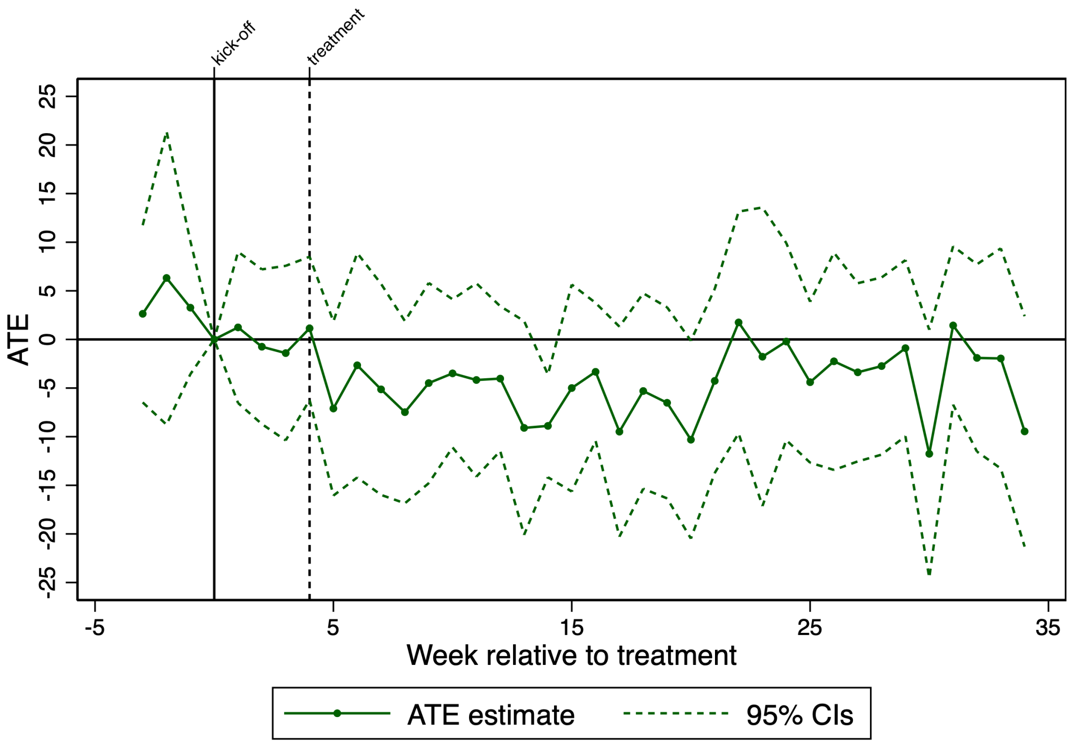
A.1 Figures

Figure 1: Experiment 1 - Softgoods Quantity



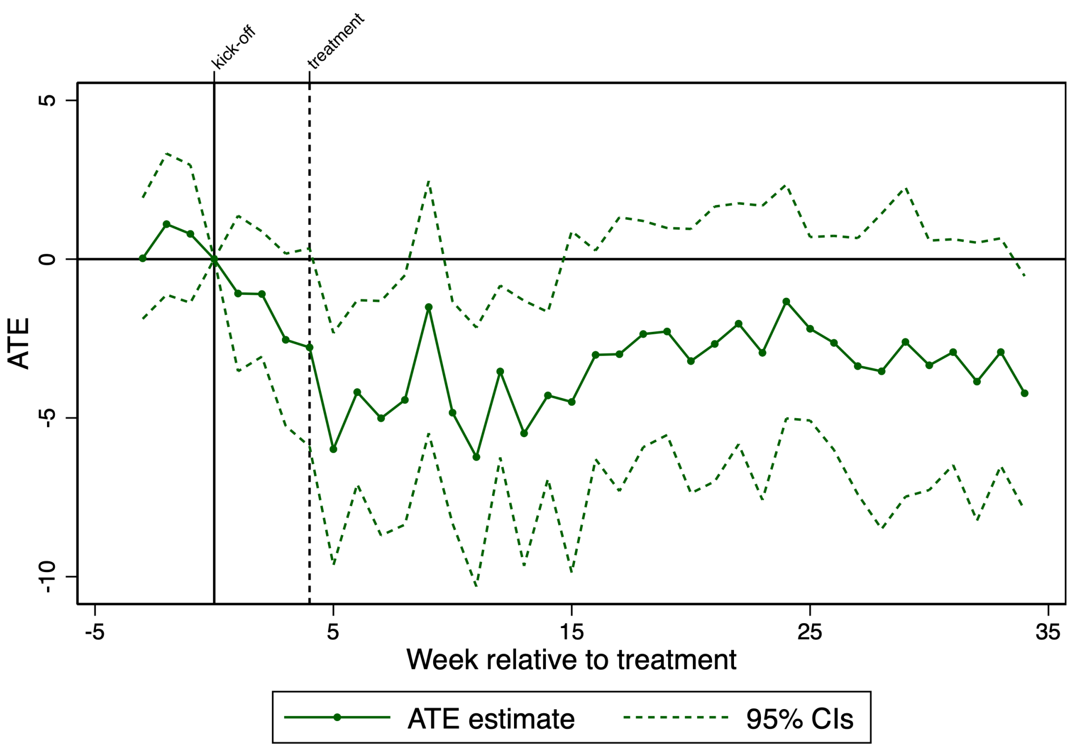
Notes: The solid-line data points are event-study coefficients from the estimation of Equation 2 using a number of weekly softgoods drops as the dependent variable. Dashed lines indicate 95% confidence intervals. The event study binary variable corresponding to the week of the kick-off call is omitted from the regression and thus set to zero in the figure; all other points are interpretable as predictive effects relative to this omitted week.

Figure 2: Experiment 2 - Duration of Alarms



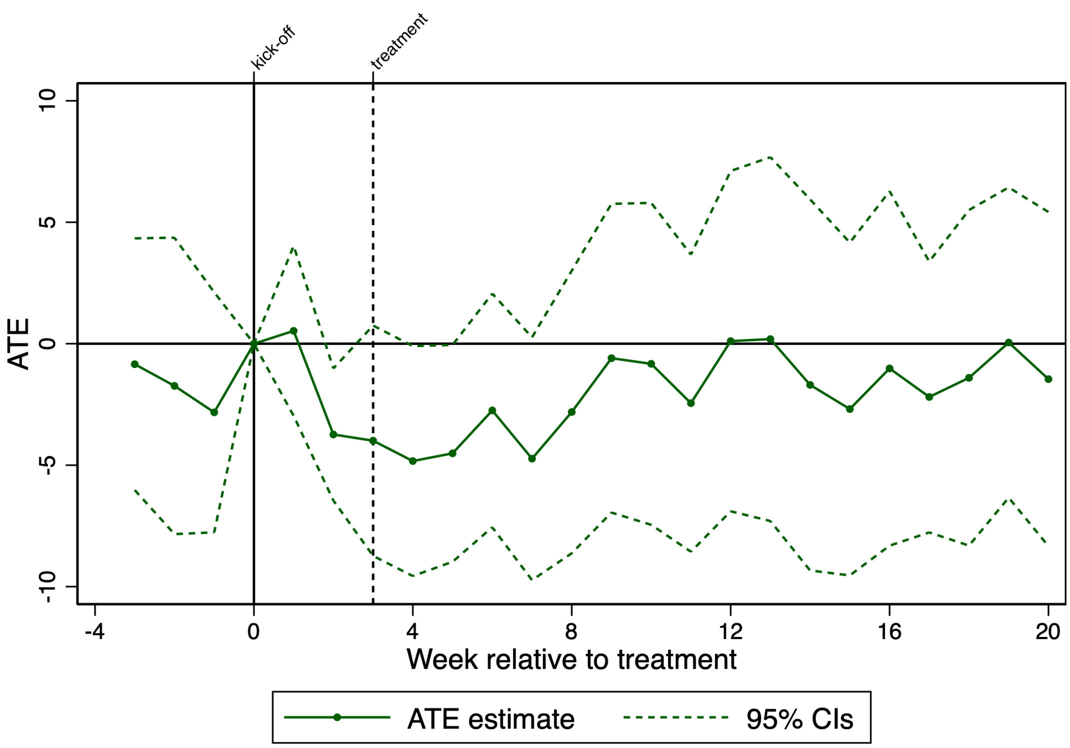
Notes: The solid-line data points are event-study coefficients from the estimation of Equation 2 using the duration of freezer door alarms as the dependent variable. Dashed lines indicate 95% confidence intervals. The event study binary variable corresponding to the week of the kick-off call is omitted from the regression and thus set to zero in the figure; all other points are interpretable as predictive effects relative to this omitted week.

Figure 3: Experiment 2 - Frequency of Alarms



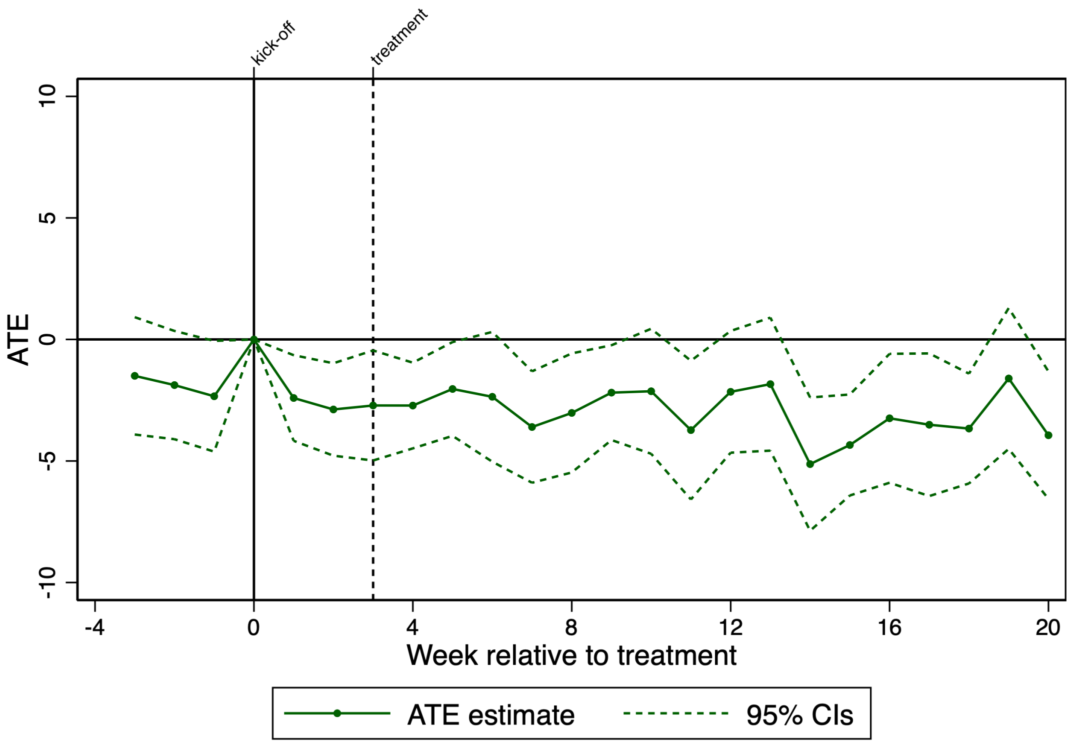
Notes: The solid-line data points are event-study coefficients from the estimation of Equation 2 using the frequency of freezer door alarms as the dependent variable. Dashed lines indicate 95% confidence intervals. The event study binary variable corresponding to the week of the kick-off call is omitted from the regression and thus set to zero in the figure; all other points are interpretable as predictive effects relative to this omitted week.

Figure 4: Experiment 3 - Total Contaminants



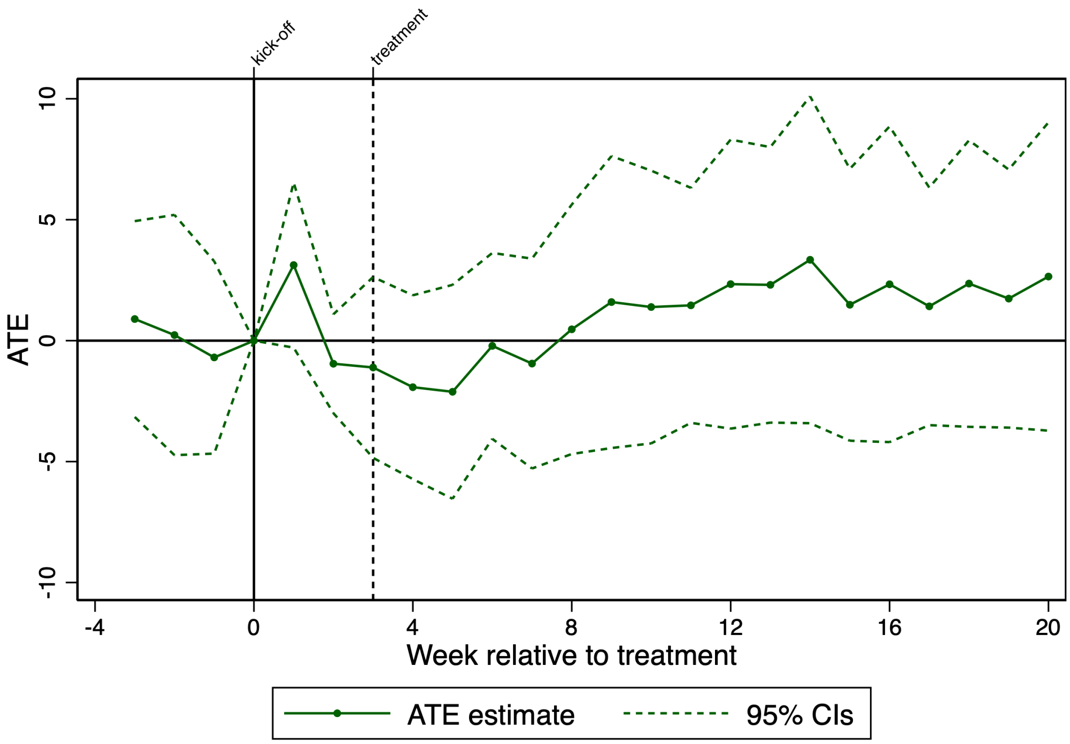
Notes: The solid-line data points are event-study coefficients from estimation of Equation 2 using number of total weekly contaminants as the dependent variable. Dashed lines indicate 95% confidence intervals. The event study binary variable corresponding to the week of the kick-off call is omitted from the regression and thus set to zero in the figure; all other points are interpretable as predictive effects relative to this omitted week.

Figure 5: Experiment 3 - Uncollapsed Cardboard



Notes: The solid-line data points are event-study coefficients from estimation of Equation 2 using number of weekly uncollapsed cardboards as the dependent variable. Dashed lines indicate 95% confidence intervals. The event study binary variable corresponding to the week of the kick-off call is omitted from the regression and thus set to zero in the figure; all other points are interpretable as predictive effects relative to this omitted week.

Figure 6: Experiment 3 - Plastic Bags



Notes: The solid-line data points are event-study coefficients from estimation of Equation 2 using number of weekly plastic bags as the dependent variable. Dashed lines indicate 95% confidence intervals. The event study binary variable corresponding to the week of the kick-off call is omitted from the regression and thus set to zero in the figure; all other points are interpretable as predictive effects relative to this omitted week.

A.2 Tables

Table 1: Experiment Description

Experiment	Intervention	Hardware Installation	Outcome Metric
E1: Softgood drops	Softgood drops display	N/A	Number of weekly softgood drops
E2: Freezer door closures	Audible alarm activation*	Audible alarms	Number of weekly average alarm activated and alarm duration
E3: Recycling	Recycling bin data sharing*	Recycling bin cameras	Number of weekly contaminants in recycling bins

Notes: In Experiment 2, freezer door closure, the control centers did not have audible alarms installed; however, we tracked door closures through the same method as in the treatment centers. For Experiment 3, package recycling, we installed cameras in recycling bins for all centers, but only shared the data with treatment group centers.

Table 2: Experiment 1 Results – Softgood Drops

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-0.090	-5.279	-1.565	-4.811	-5.111	-9.009	-4.117	-7.555	-4.031	-7.303
	(3.907)	(6.483)	(3.922)	(5.830)	(4.272)	(6.703)	(3.964)	(6.152)	(4.069)	(6.486)
Treated × Post	-11.327***	-17.410***	-17.431***	-22.635***	-17.694***	-22.946***	-17.338***	-22.127***	-17.264***	-21.905***
	(3.181)	(5.325)	(3.797)	(5.819)	(3.837)	(5.922)	(3.625)	(5.583)	(3.677)	(5.860)
Plasma Donation Count							0.026***	0.025***	0.020	-0.000
							(0.004)	(0.006)	(0.028)	(0.045)
Plasma Donation Volume									0.009	0.033
									(0.035)	(0.055)
Observations	4,416	4,416	4,416	4,416	4,416	4,416	4,416	4,416	4,416	4,416
N Control	68	68	68	68	68	68	68	68	68	68
N Treatment	24	24	24	24	24	24	24	24	24	24
Adjusted R-squared	0.275	0.270	0.284	0.280	0.291	0.293	0.310	0.304	0.310	0.304
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	23.70	23.70	23.70	23.70	23.70	23.70	23.70	23.70	23.70	23.70
Treatment Effect	-47.79	-73.46	-73.55	-95.51	-74.66	-96.82	-73.15	-93.36	-72.84	-92.43

Notes: This table reports results from estimating Equation 1, with the number of weekly softgood drops as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -11.327, for example, suggests that the intervention led to an average reduction of 47.79% in the number of weekly softgood drops. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 3: Experiment 1 Results – Softgood Drops Cost

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	0.814	-0.615	0.177	-0.707	-0.910	-2.103	-0.611	-1.666	-0.522	-1.510
	(1.133)	(1.850)	(1.115)	(1.645)	(1.254)	(1.933)	(1.142)	(1.787)	(1.181)	(1.916)
Treated × Post	-3.288***	-4.625***	-5.550***	-6.788***	-5.630***	-6.891***	-5.523***	-6.645***	-5.446***	-6.508***
	(0.862)	(1.042)	(1.155)	(1.427)	(1.160)	(1.439)	(1.126)	(1.447)	(1.113)	(1.476)
Plasma Donation Count							0.008***	0.008***	0.001	-0.008
							(0.001)	(0.002)	(0.009)	(0.013)
Plasma Donation Volume									0.009	0.020
									(0.012)	(0.016)
Observations	4,416	4,416	4,416	4,416	4,416	4,416	4,416	4,416	4,416	4,416
N Control	68	68	68	68	68	68	68	68	68	68
N Treatment	24	24	24	24	24	24	24	24	24	24
Adjusted R-squared	0.235	0.225	0.244	0.233	0.250	0.243	0.263	0.251	0.263	0.252
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	8.087	8.087	8.087	8.087	8.087	8.087	8.087	8.087	8.087	8.087
Treatment Effect	-40.66	-57.19	-68.63	-83.93	-69.62	-85.21	-68.30	-82.17	-67.34	-80.47

Notes: This table reports results from estimating Equation 1, with the number of weekly softgood drops as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -3.288, for example, suggests that the intervention led to an average reduction of 40.66% in the cost of weekly softgood drops. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 4: Experiment 2 Results – Duration of Alarms (minutes)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-3.207	-3.071	-2.060	-1.606	-2.033	-1.541	-2.244	-1.768	-2.281	-1.874
	(1.979)	(2.351)	(1.967)	(1.893)	(3.474)	(4.084)	(3.444)	(4.033)	(3.427)	(4.053)
Treated × Post	-6.610***	-6.852***	-7.094**	-7.160**	-7.072**	-7.063*	-7.060**	-7.120*	-7.174**	-7.197**
	(1.778)	(2.219)	(2.725)	(2.740)	(3.177)	(3.522)	(3.167)	(3.477)	(3.124)	(3.411)
Plasma Donation Count							0.014	0.010	-0.228	-0.250
							(0.037)	(0.037)	(0.161)	(0.167)
Plasma Donation Volume									0.298	0.322
									(0.202)	(0.207)
Observations	674	674	674	674	674	674	673	673	673	673
N Control	5	5	5	5	5	5	5	5	5	5
N Treatment	14	14	14	14	14	14	14	14	14	14
Adjusted R-squared	0.390	0.386	0.395	0.392	0.403	0.399	0.404	0.399	0.406	0.402
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	15.10	15.10	15.10	15.10	15.10	15.10	15.10	15.10	15.10	15.10
Treatment Effect	-43.78	-45.39	-46.99	-47.42	-46.84	-46.78	-46.76	-47.16	-47.52	-47.67

Notes: This table reports results from estimating Equation 1, with the daily duration of freezer door alarms as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -6.610, for example, suggests that the intervention led to an average reduction of 43.78% in duration of daily freezer door alarms. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 5: Experiment 2 Results – Frequency of Alarms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-0.828	-1.051	-0.505	-0.535	-1.429*	-1.662*	-1.685**	-1.926**	-1.687**	-1.945**
	(0.691)	(0.887)	(0.606)	(0.671)	(0.729)	(0.895)	(0.701)	(0.773)	(0.707)	(0.800)
Treated × Post	-2.863***	-3.248***	-3.515**	-3.888**	-3.920**	-4.372**	-3.929**	-4.452***	-3.958**	-4.471***
	(0.762)	(1.069)	(1.381)	(1.445)	(1.464)	(1.558)	(1.457)	(1.529)	(1.465)	(1.536)
Plasma Donation Count							0.015	0.011	-0.049	-0.050
							(0.019)	(0.019)	(0.057)	(0.065)
Plasma Donation Volume									0.080	0.076
									(0.073)	(0.081)
Observations	714	714	714	714	714	714	713	713	713	713
N Control	5	5	5	5	5	5	5	5	5	5
N Treatment	14	14	14	14	14	14	14	14	14	14
Adjusted R-squared	0.546	0.524	0.548	0.527	0.545	0.525	0.549	0.527	0.549	0.527
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	4.823	4.823	4.823	4.823	4.823	4.823	4.823	4.823	4.823	4.823
Treatment Effect	-59.35	-67.35	-72.89	-80.61	-81.29	-90.65	-81.46	-92.31	-82.08	-92.70

Notes: This table reports results from estimating Equation 1, with the frequency of daily freezer door alarms as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -2.863, for example, suggests that the intervention led to an average reduction of 59.35% in frequency of daily freezer door alarms. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 6: Experiment 3 Results – Total Contaminants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-0.226	-0.063	-0.252	-0.091	0.791	0.438	0.782	0.355	0.785	0.355
	(1.244)	(1.030)	(1.279)	(1.049)	(1.615)	(1.243)	(1.604)	(1.249)	(1.603)	(1.252)
Treated × Post	-1.669	-1.303	-0.863	-0.010	-0.259	0.312	-0.240	0.319	-0.232	0.319
	(1.846)	(1.423)	(2.006)	(1.812)	(2.073)	(1.841)	(2.188)	(1.755)	(2.192)	(1.756)
Plasma Donation Count							0.000	0.003	-0.003	0.003
							(0.004)	(0.003)	(0.021)	(0.018)
Plasma Donation Volume									0.004	-0.000
									(0.025)	(0.023)
Observations	636	636	636	636	636	636	636	636	636	636
N Control	16	16	16	16	16	16	16	16	16	16
N Treatment	15	15	15	15	15	15	15	15	15	15
Adjusted R-squared	0.362	0.377	0.369	0.385	0.364	0.379	0.363	0.384	0.362	0.383
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	7.233	7.233	7.233	7.233	7.233	7.233	7.233	7.233	7.233	7.233
Treatment Effect	-23.07	-18.01	-11.93	-0.133	-3.577	4.310	-3.312	4.410	-3.206	4.411

Notes: This table reports results from estimating Equation 1, with the number of contaminants in recycling dumpsters as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -1.669 for example, suggests that the intervention led to an average reduction of 23.07% in the number of contaminants. We excluded the first week of data collection after installation for each center to ensure that the cameras and algorithm were adequately calibrated. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 7: Experiment 3 Results – Uncollapsed Cardboard

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-0.248	-0.250	-0.257	-0.255	-0.696	-0.773	-0.685	-0.756	-0.692	-0.761
	(0.365)	(0.408)	(0.358)	(0.394)	(0.726)	(0.717)	(0.719)	(0.716)	(0.734)	(0.729)
Treated × Post	-0.794	-0.814	-1.360**	-1.388**	-1.589**	-1.647**	-1.614**	-1.649**	-1.637**	-1.647**
	(0.565)	(0.588)	(0.641)	(0.661)	(0.704)	(0.717)	(0.700)	(0.716)	(0.723)	(0.755)
Plasma Donation Count							-0.001	-0.001	0.009	0.012
							(0.001)	(0.001)	(0.008)	(0.008)
Plasma Donation Volume									-0.012	-0.015
									(0.010)	(0.010)
Observations	636	636	636	636	636	636	636	636	636	636
N Control	16	16	16	16	16	16	16	16	16	16
N Treatment	15	15	15	15	15	15	15	15	15	15
Adjusted R-squared	0.309	0.295	0.330	0.315	0.322	0.308	0.321	0.308	0.322	0.309
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	4.050	4.050	4.050	4.050	4.050	4.050	4.050	4.050	4.050	4.050
Treatment Effect	-19.61	-20.10	-33.58	-34.27	-39.22	-40.68	-39.85	-40.71	-40.42	-40.65

Notes: This table reports results from estimating Equation 1, with the number of uncollasped cardboard as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -1.669, for example, suggests that the intervention led to an average reduction of 23.07% in number of uncollasped cardboard. We excluded the first week of data collection after installation for each center to ensure that the cameras and algorithm were adequately calibrated. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 8: Experiment 3 Results – Plastic Bags

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	0.097	0.243	0.073	0.209	1.600	1.310	1.580	1.209	1.588	1.214
	(1.253)	(0.983)	(1.293)	(1.019)	(1.383)	(1.094)	(1.385)	(1.143)	(1.371)	(1.135)
Treated × Post	-0.787	-0.391	0.631	1.521	1.491	2.135	1.537	2.144	1.565	2.142
	(1.691)	(1.221)	(1.893)	(1.758)	(1.925)	(1.746)	(2.054)	(1.625)	(2.045)	(1.628)
Plasma Donation Count							0.001	0.004	-0.011	-0.007
							(0.004)	(0.003)	(0.017)	(0.014)
Plasma Donation Volume									0.014	0.013
									(0.020)	(0.018)
Observations	636	636	636	636	636	636	636	636	636	636
N Control	16	16	16	16	16	16	16	16	16	16
N Treatment	15	15	15	15	15	15	15	15	15	15
Adjusted R-squared	0.360	0.368	0.371	0.390	0.373	0.385	0.372	0.398	0.372	0.398
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	3.150	3.150	3.150	3.150	3.150	3.150	3.150	3.150	3.150	3.150
Treatment Effect	-24.97	-12.41	20.02	48.29	47.35	67.79	48.80	68.06	49.67	68

Notes: This table reports results from estimating Equation 1, with the number of plastic bags in recycling dumpsters as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -0.787 for example, suggests that the intervention led to an average reduction of 24.97% in the number of plastic bags. We excluded the first week of data collection after installation for each center to ensure that the cameras and algorithm were adequately calibrated. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 9: Experiment 3 Results – Contaminated Pickups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-0.131	-0.200	-0.094	-0.151	-0.370	-0.475	-0.375	-0.484	-0.377	-0.485
	(0.229)	(0.258)	(0.203)	(0.219)	(0.275)	(0.299)	(0.276)	(0.302)	(0.277)	(0.302)
Treated × Post	-0.420	-0.519	-0.370	-0.407	-0.507	-0.566	-0.495	-0.566	-0.499	-0.565
	(0.310)	(0.367)	(0.304)	(0.358)	(0.326)	(0.378)	(0.324)	(0.375)	(0.325)	(0.375)
Plasma Donation Count							0.000	0.000	0.002	0.001
							(0.000)	(0.000)	(0.002)	(0.002)
Plasma Donation Volume									-0.002	-0.001
									(0.003)	(0.003)
Observations	636	636	636	636	636	636	636	636	636	636
N Control	16	16	16	16	16	16	16	16	16	16
N Treatment	15	15	15	15	15	15	15	15	15	15
Adjusted R-squared	0.292	0.284	0.291	0.284	0.292	0.287	0.291	0.287	0.291	0.286
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	2.250	2.250	2.250	2.250	2.250	2.250	2.250	2.250	2.250	2.250
Treatment Effect	-18.66	-23.05	-16.44	-18.07	-22.52	-25.18	-21.99	-25.14	-22.19	-25.13

Notes: This table reports results from estimating Equation 1, with the number of weekly contaminated pickups as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -0.420 for example, suggests that the intervention led to an average reduction of 18.06% in the number of contaminated pickups. We excluded the first week of data collection after installation for each center to ensure that the cameras and algorithm were adequately calibrated. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

Table 10: Experiment 3 Results – Percentage of Contaminated Pickups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated × Kick-off	-0.031	-0.042	-0.024	-0.033	-0.071	-0.085	-0.073	-0.088	-0.073	-0.088
	(0.069)	(0.066)	(0.068)	(0.064)	(0.087)	(0.084)	(0.087)	(0.085)	(0.088)	(0.085)
Treated × Post	-0.059	-0.068	-0.040	-0.024	-0.062	-0.049	-0.059	-0.049	-0.060	-0.048
	(0.065)	(0.070)	(0.070)	(0.071)	(0.077)	(0.079)	(0.076)	(0.078)	(0.077)	(0.077)
Plasma Donation Count							0.000	0.000	0.001	0.001
							(0.000)	(0.000)	(0.001)	(0.001)
Plasma Donation Volume									-0.001	-0.001
									(0.001)	(0.001)
Observations	636	636	636	636	636	636	636	636	636	636
N Control	16	16	16	16	16	16	16	16	16	16
N Treatment	15	15	15	15	15	15	15	15	15	15
Adjusted R-squared	0.302	0.304	0.303	0.305	0.298	0.300	0.297	0.300	0.297	0.300
Center FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE			Yes	Yes						
Week FE					Yes	Yes	Yes	Yes	Yes	Yes
Pre-Treatment Weights		Yes		Yes		Yes		Yes		Yes
Baseline	0.797	0.797	0.797	0.797	0.797	0.797	0.797	0.797	0.797	0.797
Treatment Effect	-7.437	-8.562	-5.024	-3.054	-7.790	-6.124	-7.368	-6.094	-7.572	-6.082

Notes: This table reports results from estimating Equation 1, with the percentage of weekly contaminated pickups as the dependent variable. The four independent variables are Treated × Kick-off, which is equal to 1 if the center is in the treatment group after the kick-off call but before the treatment period; Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Plasma Donation Count and Volume indicate the weekly number and volume of plasma donation that a center receives. The coefficient of Treated × Post of -0.794 for example, suggests that the intervention led to an average reduction of 19.06% in the percentage of contaminated pickups. We excluded the first week of data collection after installation for each center to ensure that the cameras and algorithm were adequately calibrated. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

B Summary statistics

Table B1: Experiment 1 Balance Table

	Treatment Mean/SD	**Control** Mean/SD	**Balance** Difference/SE
Number of Softgoods Drops			10.195 (8.553)
Softgoods Drops Cost			31.636 (29.083)
Softgoods Drops per Plasma Donation			0.002 (0.001)
Plasma Donation Count			-519.497 (731.681)
Plasma Donation Volume			-392.455 (579.916)
Electricity Consumption (kWh)	31,707.499 (11,511.640)	37,006.669 (18,517.436)	-5,299.170* (2,742.577)
Center Area (sqm)	1,391.708 (155.178)	1,869.287 (3,824.753)	-477.578 (467.293)
Electricity Consumed per sqm	22.631 (7.374)	25.548 (11.086)	-2.916* (1.581)
Number of Cooling Degree Days	63.571 (97.407)	72.771 (108.719)	-9.200 (10.283)
Number of Heating Degree Days	241.522 (260.304)	219.791 (255.952)	21.731 (27.064)
F-test			5.137 (0.000)

Notes: Columns (1) and (2) display the mean of the listed center characteristic for the treatment and control groups, respectively. Standard deviations are listed beneath in parentheses. Column (3) checks for balance between the control and treatment groups with respect to the center characteristic. Results are from an OLS regression with standard errors clustered at the center level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Experiment 2 Balance Table

	Treatment Mean/SD	**Control** Mean/SD	**Balance** Difference/SE
Number of Freezer Door Alarms			1.370
			(3.804)
Duration of Freezer Door Alarms (minutes)			-11.259
			(11.421)
Plasma Donation Count			1,099.343*
			(623.986)
Plasma Donation Volume			940.258*
			(528.710)
Electricity Consumption (kWh)	32,518.656	26,594.750	5,923.906
	(19,568.392)	(8,017.249)	(6,201.615)
Center Area (sqm)	1,511.929	1,320.200	191.729
	(336.935)	(246.027)	(142.489)
Electricity Consumed per sgm	21.443	20.043	1.399
	(12.494)	(5.249)	(4.012)
Number of Cooling Degree Days	10.825	0.325	10.500
	(31.858)	(0.639)	(8.680)
Number of Heating Degree Days	472.839	489.700	-16.861
	(284.854)	(143.353)	(97.892)
F-test			1.099
			(0.401)

Notes: Columns (1) and (2) display the mean of the listed center characteristic for the treatment and control groups, respectively. Standard deviations are listed beneath in parentheses. Column (3) checks for balance between the control and treatment groups with respect to the center characteristic. Results are from an OLS regression with standard errors clustered at the center level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Experiment 3 Balance Table

	Treatment Mean/SD	**Control** Mean/SD	**Balance** Difference/SE
Total Contaminants			-4.300 (5.906)
Uncollapsed Cardboards			-2.528 (2.462)
Plastic Bags			-1.943 (5.301)
Contaminated Pickups			-0.824 (1.166)
Contaminated Pickups (%)			-0.032 (0.085)
Plasma Donation Count			-192.585 (1,015.786)
Plasma Donation Volume			-188.699 (851.545)
Percentage Recycling	0.372 (0.117)	0.341 (0.189)	0.032 (0.056)
Electricity Consumption (kWh)	34,967.442 (12,905.433)	35,371.857 (9,952.734)	-404.415 (4,239.027)
Center Area (sqm)	1,473.913 (134.313)	1,491.955 (85.403)	-18.042 (44.502)
Electricity Consumed per sqm	23.845 (8.845)	23.953 (7.569)	-0.108 (3.043)
Number of Cooling Degree Days	27.545 (51.744)	40.918 (62.074)	-13.373 (19.964)
Number of Heating Degree Days	272.477 (164.752)	208.673 (156.975)	63.805 (59.190)
F-test			2.636 (0.038)

Notes: Columns (1) and (2) display the mean of the listed center characteristic for the treatment and control groups, respectively. Standard deviations are listed beneath in parentheses. Column (3) checks for balance between the control and treatment groups with respect to the center characteristic. Results are from an OLS regression with standard errors clustered at the center level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C Pilot Study

Table C1: Pilot Results

Dependent Variable	Daily Softgoods Drops (1)	Weekly Softgoods Drops (2)	Weekly Freezer Alarm Duration (3)	Weekly Freezer Alarm Frequency (4)
Treated × Post	-0.385**	-2.526***	-110.1**	-44.74**
	(0.170)	(0.842)	(37.23)	(15.22)
Communication			-41.08*	-23.74**
			(21.13)	(10.37)
Plasma Donation Count	0.002**	0.005***	0.133**	0.038*
	(0.001)	(0.002)	(0.048)	(0.020)
Observations	20,368	2,888	170	166
N Treatment	2	2	3	3
N Communication	0	0	2	2
N Control	150	150	5	5
Adjusted R-squared	0.090	0.385	0.771	0.806
Center FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Baseline	1.416	9.534	122.8	49.51
Treatment Effect	-27.20	-26.50	-89.64	-90.36

Notes: This table reports results from estimating Equation 1 on four dependent variables. From Columns 1 through 4, the dependent variables are daily softgoods drops, weekly softgoods drops, weekly freezer alarm duration, and weekly freezer alarm frequency. The independent variables are Treated × Post, which is equal to 1 if the center is in the treatment group during the treatment period; Communication is equal to 1 if the center receives messaging about the desired behavior; Plasma Donation Count indicates the daily/weekly number of plasma donation that a center receives. The coefficient of Treated × Post of -0.405, for example, suggests that the intervention led to an average reduction of 28.60% in the daily number of softgoods drops. Standard errors, clustered at the center level, are in parentheses. Statistically significant coefficients are denoted by asterisks, with * indicating significance at the 10% level, ** indicating significance at the 5% level, and *** indicating significance at the 1% level.

D Signage

Figure D1: Experiment 1 - Softgoods Signage

Figure D2: Experiment 2 - Freezer Door Signage

Figure D3: Experiment 3 - Recycling Signage

Contact.

MIT CEEPR Working Paper Series
is published by the MIT Center for Energy and Environmental Policy Research from submissions by affiliated researchers.
For inquiries and/or for permission to reproduce material in this working paper, please contact:

General inquiries: ceepr@mit.edu
Media inquiries: ceepr-media@mit.edu

MIT CEEPR
Center for Energy and
Environmental Policy Research

MIT Center for Energy and
Environmental Policy Research
Massachusetts Institute of Technology
77 Massachusetts Avenue, E19-411
Cambridge, MA 02139-4307
USA

ceepr.mit.edu

MASSACHUSETTS INSTITUTE OF TECHNOLOGY