Social Experiment on Household Waste Management

Experiment report




Contents

Executive summary

Waste management has become an issue of national concern in Bhutan; recovery and recycling are minimal, while landfills are overflowing. Bhutanese households are responsible for storing their waste and segregating it into two categories (dry and wet). These types of waste are collected each week by waste service providers, but there is currently no collection service for household (HH) hazardous waste.

To address waste management issues, the Government has taken up waste management as a flagship programme under the 12th Five Year Plan which is being implemented by the National Environment Commission Secretariat (NECS), Royal Government of Bhutan. In support of the programme, NECS, Thimphu Thromde (municipality) and the Gross National Happiness Commission (GNHC) in partnership with the United Nations Development Programme (UNDP) and the Behavioural Insights Team (BIT), ran a randomised controlled trial (RCT) to test the effectiveness of interventions to improve the quality of household segregation of waste into dry, wet and household (HH) hazardous waste.

The RCT randomly assigned approximately 600 households in Changzamtog area of Thimphu, from the data received from Ministry of Works and Human Settlement, to three 'conditions' with differing interventions:

- **Control:** Households received no interventions
- **Treatment 1:** Households received bins plus information (bin stickers)
- **Treatment 2:** Household received bins plus information (bin stickers), plus feedback messages.

To measure impact, 60 enumerators (DeSsups) were hired and supervised by environment inspectors, environment officers and programme coordinators from Thimphu Thromde, NECS and UNDP's Accelerator Lab. Over a period of three weeks, enumerators carried out weekly collections of dry, wet and HH hazardous waste from each household. Waste collection was planned for three consecutive weeks, however, a weeklong partial lockdown prevented enumerators from commuting pushing the third collection by a week.

After collection, enumerators weighed the full bag of each type of waste, then segregated waste into different components: correct waste and incorrect waste (for all three waste bags) and HH hazardous waste (for dry and wet waste bags). The waste, after proper segregation, was weighed again.

There were a large number of discrepancies in the trial data, where the weight of the total waste bag did not match the summed total of the weight of each of the components of the waste. The data was analysed using three different error thresholds: 1%, 5% and 10%.

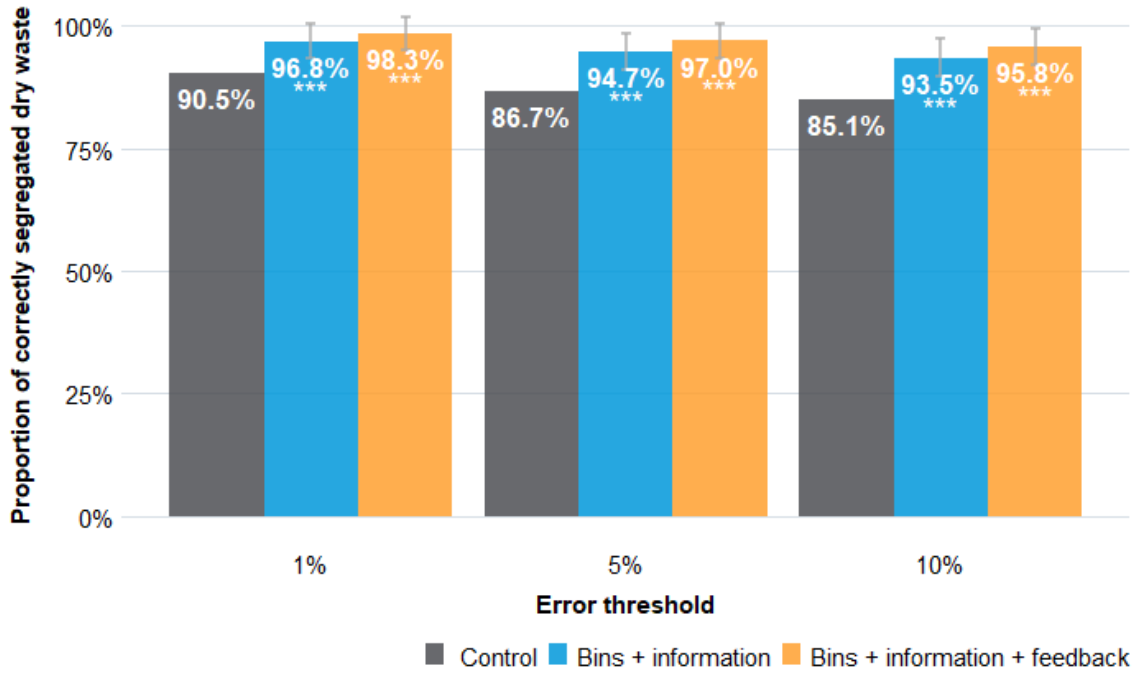
Trial results

Primary outcome I: Segregation of dry waste

Both treatments statistically significantly[1] improved the proportion of correctly segregated dry waste. Households that were assigned 'bins plus information' intervention increased quality

[1] Refer to Annex A for a definition.

of their dry waste segregation by between 6.3 and 8.4 percentage points, compared to households who received no intervention. Households that were assigned 'bins plus information plus feedback messages' increased quality of their dry waste segregation by between 7.2 and 10.7 percentage points, compared to households who received no intervention. Furthermore, both treatments appear to have increased the proportion of observations with 100% correct segregation of dry waste.



Primary outcome II: Segregation of wet waste

Neither of the treatments had a statistically significant effect on proportion of correctly segregated wet waste. This may have been because most households were already segregating their wet waste correctly. It may be easier for households to distinguish what is 'wet' waste from other types of waste, than to distinguish 'dry' waste from other types of waste.

Secondary outcome: Segregation of HH hazardous waste

Both treatments appeared to have significantly reduced the amount of HH hazardous waste incorrectly placed in dry and wet bins.

Exploratory outcomes

It appears that the treatments resulted in slightly higher amounts of waste being disposed, but a significant portion may have been HH hazardous waste that was not disposed correctly by households in the control condition. It also appears that segregation improved by approximately four percentage points across all three trial arms from the first to the last week of the trial.

1) Introduction

According to the National Waste Inventory Survey 2019 conducted by National Statistics Bureau, households in Bhutan generate 80.91 tonnes of solid waste every day. Currently, 80 percent of the waste is disposed in landfills. Most landfills are overflowing, deteriorating health of the pristine environment, including streams and rivers in its vicinity. To reduce the amount of waste being sent to landfill and to maximise recovery rate, waste needs to be correctly segregated at source.

Currently, Bhutanese households are responsible for storing their waste and segregating it into two categories: dry and wet. It is unknown how well households correctly segregate their waste. Dry and wet waste are collected on separate days each week by waste service providers. There is currently no collection service specifically for HH hazardous waste.

To address waste management issues, the Waste Management Flagship Programme was launched on 23 January 2020 by the Government as one of the priorities and national concerns under the 12th Five Year Plan. The programme aims to achieve a Zero Waste Bhutan by 2030, reducing waste being disposed at landfill from 80 percent to less than 20 percent. The programme will provide:

- **Sorting and segregation facilities**, including bins to households for dry, wet, and HH hazardous waste, common residential bins, and drop-off centres;
- **Transportation facilities**, including vehicles to collect HH hazardous waste from households;
- **Recovery facilities**; and
- **Treatment and disposal facilities**.

To initiate the programme's roll-out, NECS, Thimphu Thromde and GNHC in partnership with UNDP and BIT ran a randomised controlled trial (RCT) to test the effectiveness of interventions to improve the quality of household segregation of waste into dry, wet and HH hazardous waste. The RCT was conducted with households in the Changzamtog area of Thimphu, Bhutan.

The purpose of this report is to provide an overview of the RCT that was conducted, including:

- The design of the interventions - **see Section 2;**
- The design of the RCT - **see Section 3; and,**
- The results of the RCT - **see Section 4.**

A recommendation, based on the results of the RCT, is provided in **Section 5**. This report is supported by a baseline report which outlines the findings from exploratory research conducted to inform the design of interventions and RCT - **see Annex C**. This report is complemented by a separate Policy Note that discusses the policy implications of this trial.

2) Intervention design

The RCT tested three interventions. These interventions were split across three trial conditions, as outlined in the table below.

Trial condition	Interventions administered
Control	No interventions
Treatment 1	Household bins plus information (stickers)
Treatment 2	Household bins plus information (stickers) plus feedback messages

Each of these interventions are summarised below, along with the “theory of change” for why it was believed each intervention might work, and a summary of key evidence supporting each intervention.

Household bins

Description

The household bins were three indoor bins that were given to households to store their waste in. These included:

- A blue 40 litre bin for dry waste;
- A green 7 litre bin for wet waste; and,
- A red 7 litre bin for HH hazardous waste.

Photos of each of these bins are provided in Figure 1 below.

Figure 1: Blue, green and red household bins

Theory of Change

Providing households with bins may encourage waste segregation because it alters the "choice environment", that is, the available options.[2] In this context, the choice environment is the waste disposal options available when households make segregation decisions.

The exploratory research found that many households in Changzamtog do not have household bins, but instead make use of substitutes such as hessian bags or old paint buckets. These substitutes may not be of appropriate size or material and may therefore discourage correct segregation in favour of more practical solutions, such as putting all waste into the bin with the most space, regardless of waste type.

Providing bins to households that clearly indicate different kinds of waste will simplify the choice of how and where to dispose of waste, which should in turn lead to better segregation. Simply put, the bins should make it easier to manage waste properly, leading to an enhanced recovery rate.

Supporting evidence

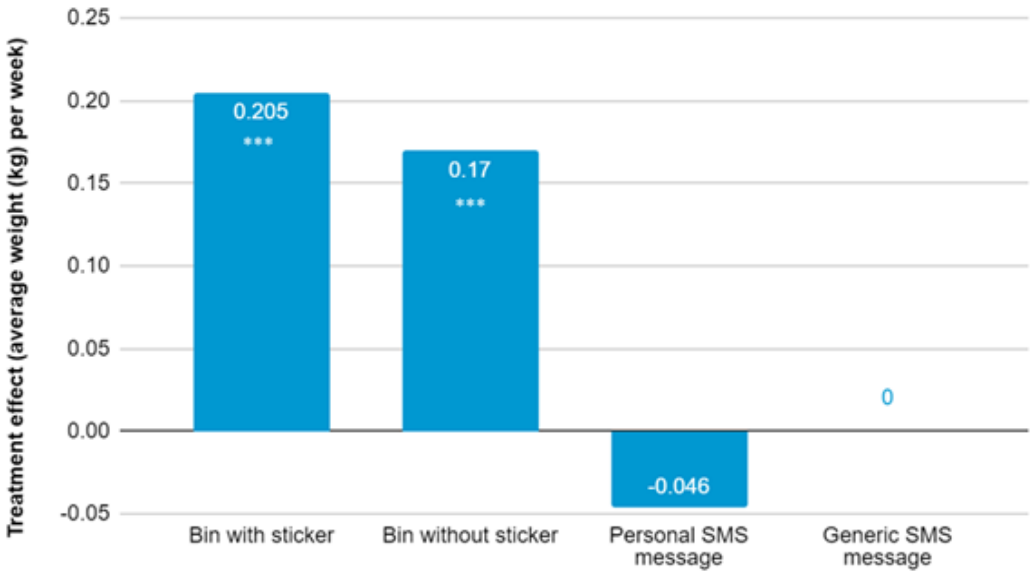
There is some existing evidence of the efficacy of providing households with bins to support segregation of waste at source. A World Bank study in Peru tested a number of interventions in a large scale RCT, where groups of randomly selected households received different messages and some groups of households were randomly selected to receive recycling bins (some with stickers on them encouraging recycling, some without).[3]

The researchers found that providing people with a recycling bin (with or without a sticker) had a statistically significant and positive effect on both the frequency and the amount of recycling. Other interventions, such as SMS reminders and environmental messages to improve participation, did not have a statistically significant effect on recycling behaviour. The treatment effects are shown in Figure 2 below.

Figure 2: Treatment effect of interventions in the World Bank study in Peru

[2] UK Cabinet Office (2010). MINDSPACE: Influence behaviour through public policy. Available at https://www.bi.team/publications/mindspace/

[3] Chong, A., Karlan, D., Shapiro, J., & Zinman, J. (2015). (Ineffective) messages to encourage recycling: evidence from a randomized evaluation in Peru. *The World Bank Economic Review*, *29*(1), 180-206.



Information (stickers)

Description

Information on how to correctly segregate household waste was given to each household in the form of three stickers, one for each of the three bins (stuck onto the side of the bin). The stickers had photos of items that should and should not go into each bin, with green ticks and red crosses, as shown in Figure 3 below.

Figure 3: Stickers for the blue, green and red household bins

Theory of Change

Information may encourage waste segregation because it addresses confusion households have about which type of waste items should go into which bins. Being displayed prominently on bins, the information will also act as a reminder to households at the point at which they are segregating their waste.

Supporting evidence

Several trials have tested the effects of providing information in different forms to encourage correct waste segregation. In relation to stickers, a large RCT in Surrey tested whether or not stickers on the lid of bins could prompt people to separate food waste from recycling (instead of placing everything in the refuse bin).[4] The stickers read, in large letters, “No Food Waste Please. Remember to use your food recycling caddy” (see Figure 4 below).

[4] Shearer, L., Gatersleben, B., Morse, S., Smyth, M., & Hunt, S. (2016). A problem unstuck? Evaluating the effectiveness of sticker prompts for encouraging household food waste recycling behaviour. *Waste Management* [In Press].

Figure 4: Sticker prompts used by Shearer et al.

While the control group saw no statistically significant change in the food waste collected on a weekly basis, the treatment group saw a 20.74 per cent increase in the weekly tonnage (from 1.23 tonnes to 1.49 tonnes).

Feedback messages

Description

Feedback was provided to households through SMS messages (in English). Each of these messages had three components. They:

1. Provided specific feedback on key items households were incorrectly segregating;
2. Made the effects of incorrect segregation more salient; and,
3. Highlighted neighbourhood behaviour.

The three messages that were sent to households over the course of the trial are provided below.

Feedback message 1

This week some households in Changzamtog put diapers and sanitary pads in their dry waste. Remember, these items should go into your red hazardous waste bin. This will help prevent your other dry waste being spoiled. We are seeing that more and more households just like yours are making the effort to segregate - keep up the good work! From NEC and UNDP

Feedback message 2

Did you know sauce spoils dry waste? Please put all food scraps (including sauce) into your green wet waste bin, and rinse your food containers before putting them in your blue dry waste bin. We appreciate your support in helping out the waste segregators at the Greenerway center. From NEC and UNDP

Feedback message 3

Some households in Changzamtog have been putting fruit and vegetable peels in their dry waste. Remember, all food scraps should go into your green wet waste bin. This will help prevent your other dry waste being spoiled. Lots of households have been segregating - let's all help keep Bhutan clean, green and beautiful! From NEC and UNDP

Theory of Change

The overarching theory for why feedback may encourage waste segregation is that people may not know they are segregating poorly. Targeted information can speak to specific errors in segregation. In this sense, the errors are made more *salient* and correcting the behaviour is *easier.*

To make this information provision as effective as possible, the feedback messages contained three key components:

1. *Provide specific feedback on key items households are incorrectly segregating*

This has three purposes. Firstly, it provides information that clarifies errors so that people can easily correct their behaviour. Secondly, it makes it clear that their waste segregation is being monitored - that somebody notices these errors. Lastly, it increases the credibility of the waste disposal system, by demonstrating that different types of waste are sorted and treated differently - suggesting that household segregation is worthwhile.

2. *Make the effects of incorrect segregation more salient*

This may correct fallacies that households may hold, for example that 'no one' is affected by how they segregate their waste. Highlighting consequences has proven effective in other domains, such as tax compliance, where inaction may seem costless.[5]

3. *Highlight neighbourhood behaviour*

Referring to neighbourhood behaviour can encourage greater effort because it communicates a peer "behavioural standard" of higher effort. People are motivated to match this standard, whether it is past behaviour, peer behaviour, or their own values or intentions.[6]

Supporting evidence

1. *Provide specific feedback on key items households are incorrectly segregating*

A study on energy consumption that sent people weekly postcards to simply tell them they were part of a study about electricity use and that no further action was required found a 2.7 percent reduction in monthly energy consumption.[7]

2. *Make the effects of incorrect segregation more salient*

An RCT in Guatemala found that framing a non-declaration as an active choice, rather than an unintentional error, increased the rate of payment (from 3.9 per cent to 5.4 per cent) as well as the average amount paid conditional on paying (from $6.67 to $24.62).[8]

3. *Highlight neighbourhood behaviour*

A small RCT in California tested the effect of placing personal feedback messages as well as group social norms feedback messages on door hangers in single-family dwellings.[9] The personal feedback messages stated the amount of each type of waste collected at the house

[5] The Behavioural Insights Team. Update Report 2013-2015. Available at https://www.bi.team/wp-content/uploads/2015/08/BIT_Update-Report-Final-2013-2015.pdf

[6] Kluger & DeNisi (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological bulletin, 119(2), 254.

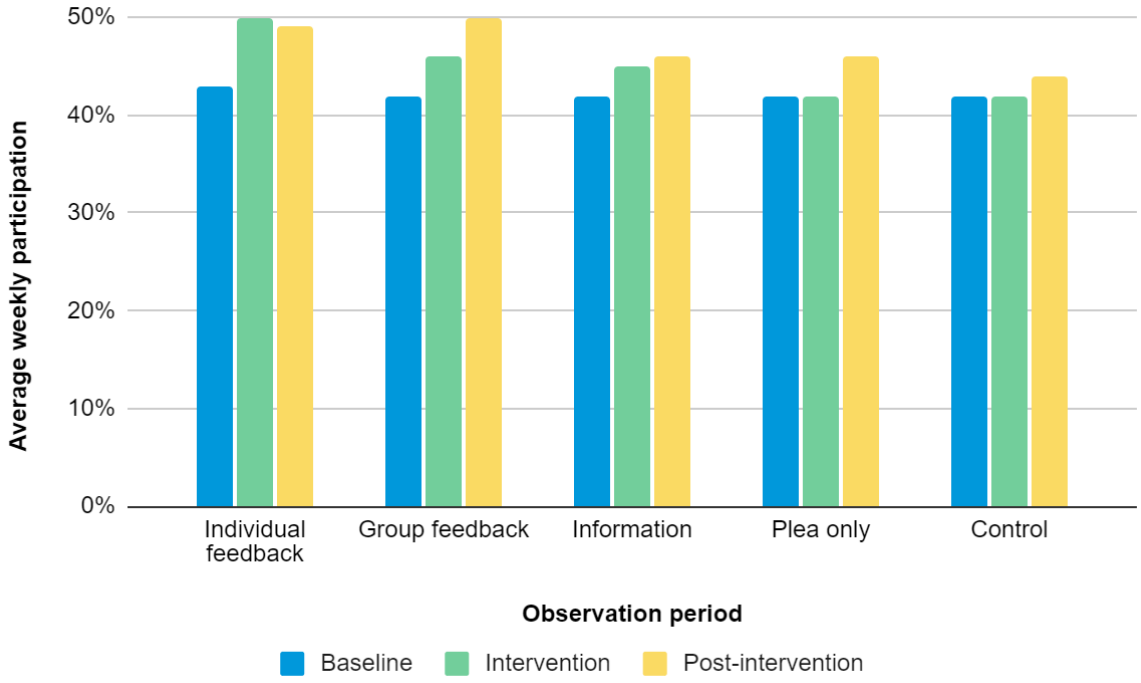
[7] Schwartz, D., Fischhoff, B., Krishnamurti, T., & Sowell, F. (2013). The Hawthorne effect and energy awareness. *Proceedings of the National Academy of Sciences*, *110*(38), 15242-15246.

[8] The Behavioural Insights Team. Update Report 2013-2015. Available at https://www.bi.team/wp-content/uploads/2015/08/BIT_Update-Report-Final-2013-2015.pdf

[9] Schultz, P. W. (1999). Changing behavior with normative feedback interventions: A field experiment on curbside recycling. Basic and applied social psychology, 21(1), 25-36.

the previous week, the current week, and the cumulative over the course of the study. The group social norms feedback messages stated the same information as the personal feedback condition, as well as the percentage of households that participated that week. The results (presented in Figure 5 below) showed a statistically significant increase in participation among those receiving the norms message (from 42 per cent to 50 per cent), as well an increase in the total amount recycled. The other intervention arm, providing individual feedback, saw similar results (from 43 per cent to 49 per cent).

Figure 5: Average recycling participation per condition



Another study in the United States that tested whether social norm messages ("the majority of guests reuse their towels") statistically significantly increased the rate at which guests reused their towels compared to the hotel's standard environmental message.[10] The social norm condition increased towel reuse rates from 35 per cent to 44 per cent, suggesting that social norms play a powerful role in influencing behaviour.

[10] Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of consumer Research*, *35*(3), 472-482.

3) Trial design

This trial was a three-arm randomised-controlled trial, randomised at the household level. There were approximately 200 households in each trial arm.

The procedure that was followed to conduct the trial, its outcome measures, and the statistical model used for analysing the trial results, are outlined below.

Trial procedure

To conduct the trial activities, 60 enumerators (DeSsups)[11] were engaged for six weeks. Enumerators were divided into eight groups and further split into pairs. Each group was supervised by environment inspectors from Thimphu Thromde and NECS.

Randomisation

Households were randomised prior to being recruited into the trial. Randomisation is key to an RCT, as it ensures that results between treatment conditions can be attributed to the treatment and not to other factors. The list of households in Changzamtog[12] was filtered to exclude any household with more than 12 residents. Three households from over 3,000 in the original list were excluded. All remaining households were randomised into the three trial conditions, as although it was aimed to recruit 600 households, it was not expected that all households would agree to participate.

Households were stratified before randomisation. The list included details on the approximate household size, and thus the households were grouped into three strata. The strata were: small (1-2 people), medium (3-4 people), and large (5-12 people).

Recruitment of households

The list of households in Changzamtog included a contact number for the head of each household. Enumerators called the head of household over the phone, following a script asking if the household would like to participate in the trial. Only those households that agreed to participate were included in the trial. Households that had moved out of Changzamtog, or were planning on moving outside of Changzamtog over the trial period, were also not included in the trial.

Enumerators stopped calling households after over 200 households from each condition agreed to participate. Enumerators recruited slightly over 200 for each condition, with the expectation that some households would drop out before the trial commenced.

Piloting

To confirm feasibility, the collection process was piloted with three households before initiating trial. No major amendments were required.

Distribution of bins

[11] De-suups were coordinated through the De-suung Office.
[12] The list of all households in the Changzamtog area of Thimphu was provided by the Ministry of Works and Human Settlement, Royal Government of Bhutan.

The household bins (with information stickers on them) were distributed to households in the Treatment 1 and Treatment 2 conditions by enumerators on Saturday 13 March 2021. When providing the bins to households, enumerators briefly described what waste was meant for each bin, and how to read the information sticker.

Collection of household waste

Once each week, for three weeks, enumerators collected waste from each household. For Treatment 1 and 2 households, waste were collected from the three households bins. For control households, the collection included waste that households would normally give to their waste collector and/or take to a drop-off centre. All households were asked not to give any of their waste to their waste collection provider or take any of their waste to a drop-off centre over the trial period.

Half of the households were visited on a Tuesday, and the other half were visited on a Saturday. Enumerators went to the front door of each household to collect waste. Enumerators pre-arranged a time they would visit each household during recruitment. Visiting time for the collection was confirmed with households, via text message, a day before the visit.

Initially, the plan to collect waste over three consecutive weeks. However, Bhutan entered a partial lockdown between 27 March and 3 April 2021, as the country rolled out COVID-19 vaccination. This prevented enumerators from being able to move around Thimphu to collect household waste. As a result, the second Saturday collection was brought forward to Friday, and the third collection for each household was postponed by a week.

Sorting of household waste

After collecting waste from each household, enumerators carried the waste to one of eight sorting sites set up across Changzamtog for the trial. Using scales[13] provided at each sorting site, enumerators first weighed the full bag of each type of waste.[14] Enumerators then sorted by waste type. For the dry waste, enumerators split the waste into three piles:

- Correct dry waste;
- HH hazardous waste (incorrectly put into the dry waste); and,
- Incorrect waste (any items that were not dry or HH hazardous waste).

Similarly, for the wet waste, enumerators split the waste into three piles:

- Correct wet waste;
- HH Hazardous waste (incorrectly put into the wet waste); and,
- Incorrect waste (any items that were not wet or HH hazardous waste).

For the HH hazardous waste, enumerators split the waste into two piles:

- Correct HH hazardous waste; and,
- Incorrect waste (any items that are not HH hazardous waste).

Enumerators were provided with a guide to help them sort the waste into different categories. If enumerators were unsure about how to sort any items, they were instructed to

[13] The scales were provided by the National Statistics Bureau.
[14] For the bag of wet waste, enumerators punctured the bag and let any liquid drain out before the bag was weighed.

include them in the 'incorrect waste' pile. Each of the above components were weighed separately. All weights were entered by enumerators into a Google form.

After weighing, enumerators left all dry waste together, and all wet and HH hazardous waste together. At the end of every collection day, a waste collection provider took the dry waste to a recovery centre and the wet and HH hazardous waste to landfill.

Provision of feedback to households

Feedback messages were sent to Treatment 2 households three times. Most feedback messages were provided via SMS, however some messages were provided via WhatsApp or WeChat, if preferred by households. Feedback messages were sent to households after the first two waste collections. Preference for the mode of feedback delivery was sought during the second visit to the households. This was so that the feedback messages reflected how households were segregating their waste (as recorded by enumerators during the data collection process). The third feedback message was sent to households during the week of lockdown. Households whose waste was collected on Tuesdays were sent messages on Fridays. Households whose waste was collected on Saturdays were sent messages on Mondays.

Endline survey

At the final collection, enumerators conducted an endline survey with households. The survey included questions on:

- Household demographics, including highest level of education in the household and number of people living there;
- Willingness to pay for the three household bins and bin liners;
- Households perception of the size of the bins, and what households liked/disliked about the bins;
- Willingness to pay for a timely waste collection service;
- Whether households would value, and be happy to help manage, a common residential bin; and,
- Suggestions for improving drop-off centres' services.

Outcome measures

Before the trial commenced, the outcome measures that would determine whether the trial was a success were pre-specified. The primary, secondary and exploratory outcomes measures are outlined in the tables below.

Primary outcome measures

Research question	Outcome measure
Do the interventions improve segregation of dry waste?	Proportion (by weight) of dry waste correctly placed in the dry waste bin (continuous variable between 0-100%).
Do the interventions improve segregation of wet waste?	Proportion (by weight) of wet waste correctly placed in the wet waste bin (continuous variable between 0-100%).

Secondary outcomes measures

Research question	Outcome measure
Do the interventions improve the correct disposal of HH hazardous waste?	Amount (by weight) of HH hazardous waste in dry and wet bins combined (continuous variable in grams).

Exploratory outcome measures

Research question	Outcome measure
Do the interventions increase the amount of HH hazardous waste households dispose?	Amount (by weight) of HH hazardous waste in all bins combined (continuous variable in grams).
Do the interventions increase the amount of total waste households dispose?	Amount (by weight) of waste in all bins combined (continuous variable in grams).
Do the interventions improve segregation of dry waste across time?	Primary analyses comparing the first week and last week (of the proportion (by weight) of dry waste correctly placed in the dry waste bin).

Analysis approach

Data cleaning

Before conducting the data analysis, the data was cleaned. Some weights were entered as zero, due to enumerators not being able to collect waste as no-one from the household was home, or a household did not provide a particular type of waste. These zeros were removed (and instead coded as ‘not applicable’).

There was some duplication in the entries. Where entries were essentially identical, the duplicate entries were removed. Where there were multiple entries with the same household ID and treatment condition, but different data, both entries were deleted as it was impossible to identify which entry was correct. Similarly, multiple entries with the same household ID on the same day were deleted as it was impossible to identify which entry was correct.

Some households who agreed to participate when recruited, dropped-out before the trial commenced. To compensate, enumerators asked other households to participate, assigning them to one of the three treatment conditions. As these households were not randomly allocated to a treatment condition they were dropped from the sample and not included in the trial analysis.

Error thresholds

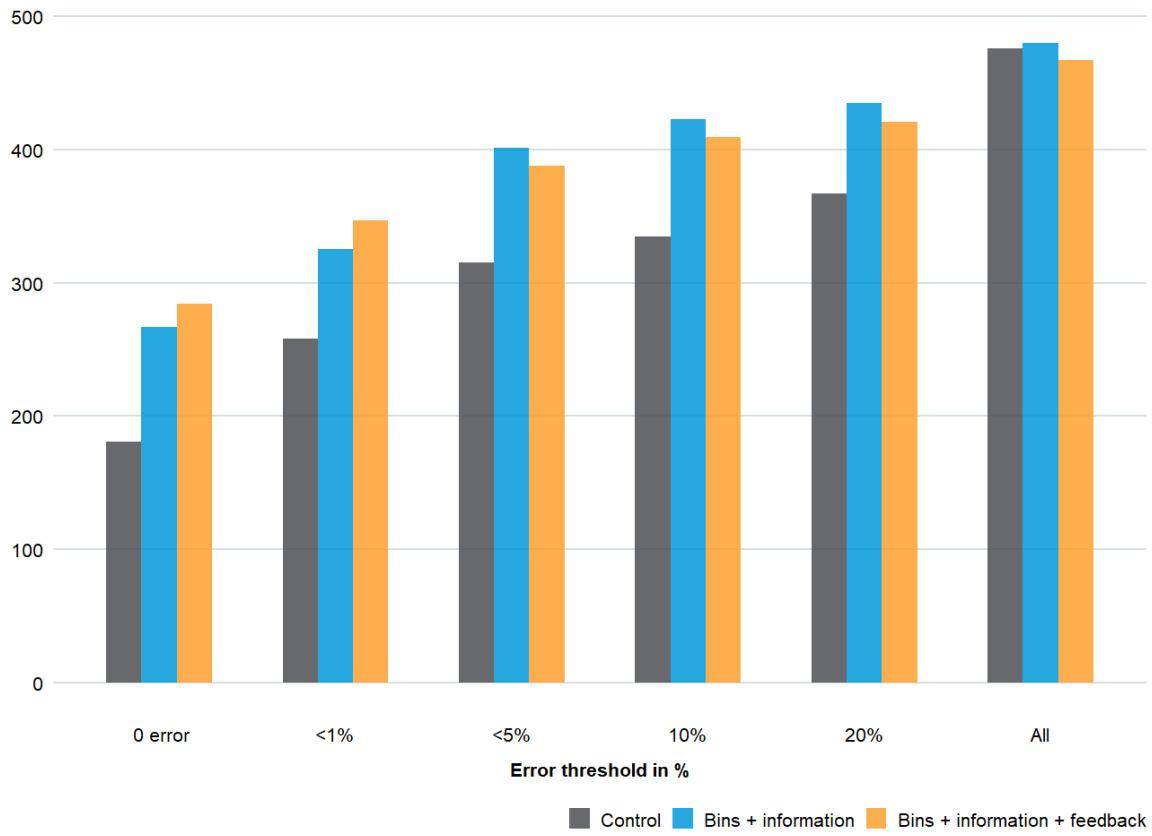
When weighing each type of waste, enumerators weighed the total bag of each type of waste. Enumerators then split the bag into piles, and weighed each pile. There were several discrepancies, where the weight of the total bag did not match the summed total weight of each pile of waste. There may be a range of reasons for this discrepancy, for example:

- Scales that are highly accurate can be sensitive to small shifts in the contents of a container;
- Environmental factors such as wind made it challenging to ensure segregated piles remained intact and uncontaminated; and,
- Manual data entry errors.

Notably, while data that merely represents a small discrepancy due to sensitive scales should not be excluded, it is prudent to exclude data that is incorrect (for example, due to a manual data error). Three different error thresholds were therefore looked at, where observations with discrepancies above these thresholds were removed from the analysis. Error thresholds were set at a 1%, 5% and 10% difference between the weight of the total bag of waste, and the summed total of the weight of each of the piles of waste.

The number of observations, by treatment, that would remain in the sample depending on where the error threshold is set for dry waste, as this appears to have the largest spread of values, is shown in Figure 6 below.

Figure 6: Number of observations remaining using 1%, 5% and 10% error thresholds



Notably, this suggests that while the arms were balanced on initial numbers (in fact, slightly more observations in the control condition), using any reasonable error threshold to exclude observations leads to an unbalanced sample, with the control condition particularly impacted.

Looking at the individual trial arms, it appears that this effect may be due to the impact of the treatment, i.e., the treatments lead to fewer observations with data errors.

Figures 7-9, below, show each observation, in ascending order of the proportion of dry waste correctly segregated. The shading indicates the relevant error threshold, with darker colours indicating a higher error threshold. Note, as the weight of correctly segregated dry waste is being measured, divided by the original total weight of all dry waste, it is possible for the total to be more than 100% if one of the entries was incorrectly entered. To ensure the charts are readable, the proportion has been capped at 120%.

Figure 7: Correctly segregated dry waste in the control condition across error threshold from 1-20%+

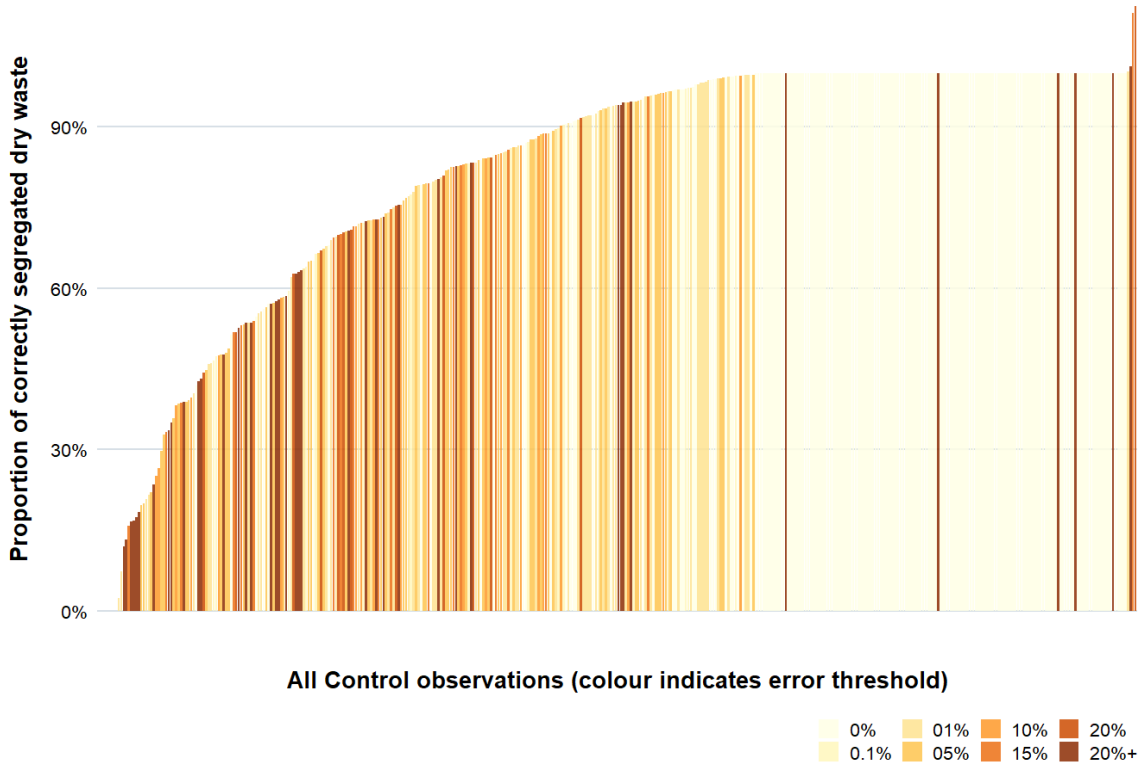


Figure 8: Correctly segregated dry waste in the Treatment 1 condition across error threshold from 1-20%+

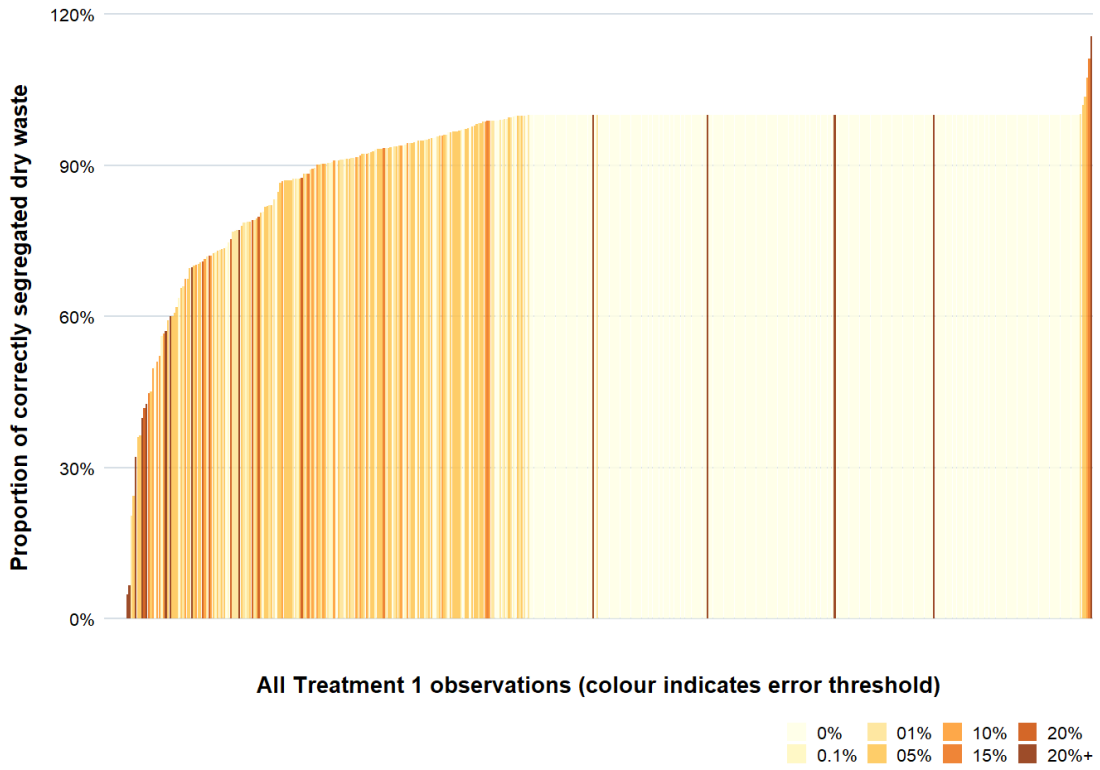


Figure 9: Correctly segregated dry waste in the Treatment 2 condition across error threshold from 1-20%+

The charts indicate two salient facts. First, as the proportion of correctly segregated waste moves away from 100%, the likelihood of an error increases substantially. It is likely that this

is because those entries required more manual handling and multiple weighing, allowing for errors to creep in. Conversely, if waste were perfectly segregated, enumerators could simply re-enter the total weight of the contents of the bin. Secondly, the treatments appear to increase the proportion of households that have correctly (100%) segregated waste.

This represented a challenge for the analysis. A tighter error threshold understates the impacts of the treatments, as it is more likely to exclude households with poor segregation. This disproportionately affects the control group, leaving mostly households with perfect segregation. A larger error threshold, however, may lead to results that are less reliable, as results are more likely to include data that has manual entry or other errors.

To address this, conducted regressions were applied using a number of thresholds, using reasonably conservative error thresholds (1%, 5% and 10%).

Statistical model

Before the trial commenced, the number of analyses was pre-specified to ensure that findings were robust for future scaling and replication, and to reduce the risk of spurious findings.

Three simple linear regressions were estimated for each outcome variable. The first was with just the treatment variable included, i.e.:

$$Outcome_i = \alpha + \beta \cdot Treatment_i + \epsilon_i$$

The second included the treatment variable and two covariates. These were:

- Number of people living in the household; and
- Enumerator pair assigned to collect waste.

$$Outcome_i = \alpha + \beta \cdot Treatment_i + \theta \cdot Covariates_i + \epsilon_i$$

In theory, enumerators should make the same decisions when determining whether waste is segregated correctly. However, it is possible that some might be systematically biased in some way, therefore this was captured through a control variable.

Finally, the third regression included random treatment allocation, prespecified covariates, and an additional covariate that was collected (highest education level in the household).

All regressions were run three times, using error thresholds of 1%, 5% and 10%. All regressions were conducted as linear regressions with robust errors. The Benjamini-Hochberg step-up procedure was used to correct for multiple comparisons. Results in the charts below reflect the results from the third regression including all covariates.

4) Trial results

The results of the analyses of the primary, secondary and exploratory outcome measures are presented below.

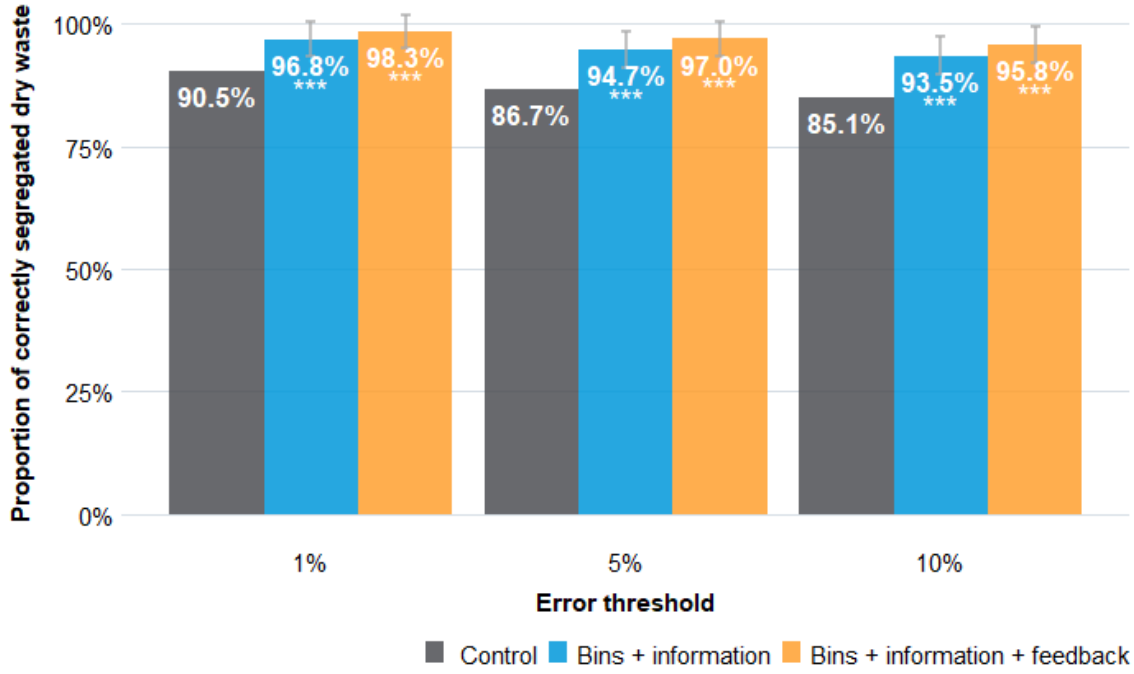
Segregation of dry waste

Primary outcome I: Proportion of correctly segregated dry waste

The first primary outcome measure was the proportion of correctly segregated dry waste in the dry bin (i.e., weight of correctly segregated dry waste, divided by total waste in the dry bin).

The results of the analyses, using a 1%, 5% and 10% error threshold are presented in Figure 10 below.

Figure 10: Segregation of dry waste using 1%, 5% and 10% error thresholds



The proportions shown in the figure above are the average proportions of the households correctly segregated waste collected in each condition. For each household, the observations from the three weeks are averaged to produce a single measure for each household of correctly segregated dry waste.

With a 1%, 5%, and 10% error threshold, both treatments had positive and strongly statistically significant results, i.e., the treatments appear to have improved the proportion of correctly segregated dry waste. These results persist even after multiple comparisons are corrected for using the Hochberg step-up procedure. In fact, these results would persist under

the most conservative corrections for multiple comparisons (Bonferroni correction), even if a thousand comparisons were to be conducted.

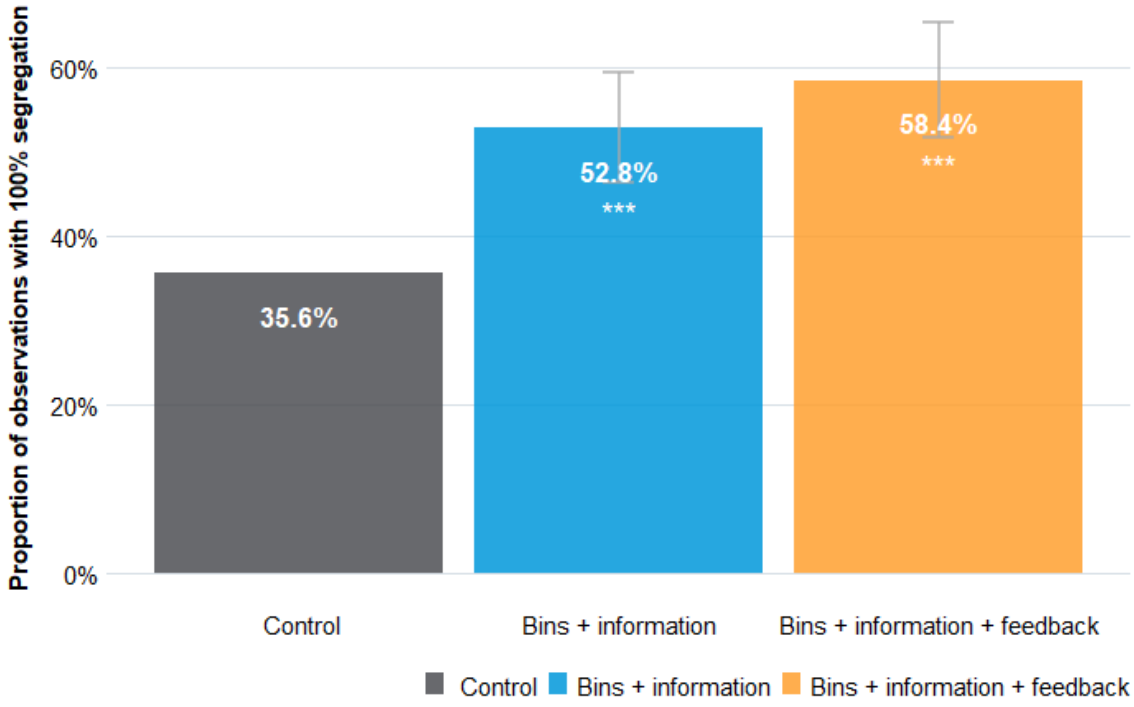
The estimates of the treatment effect are between 6.3 and 8.4 percentage points for the 'Bins + Information' treatment arm, and 7.2 and 10.7 percentage points for the 'Bins + Information + Feedback' treatment arm. These minimum and maximum treatment effects are the unadjusted OLS estimate for the 1% error threshold model and the OLS model with all covariates of the 10% error threshold model respectively.

If scaled up across Bhutan, this could lead to an additional 25,000 households correctly segregating their dry waste, and an additional 3,000 tonnes of HH hazardous waste being correctly segregated and collected each year.[15] However, it is important to note that our trial only ran for three weeks, so it is not entirely clear what the impacts will be over a longer period of time.

Proportion of households with perfect segregation of dry waste

The treatment effects for dry waste are confirmed by a different analysis that targets the same broad outcome as an exploratory analysis 'check'. The results of an analysis of the proportion of observations that have 'perfect' segregation (i.e., the amount of correct dry waste is exactly equal to the total weight) using all observations (regardless of error threshold) are shown in Figure 11 below.

Figure 11: Observations with perfect segregation of dry waste



The proportions shown in the figure above are of all observations from households across the three weeks in each condition.

[15] These calculations are based on an estimate of 150,000 households in Bhutan.

Regardless of the error threshold, both treatments had positive and strongly statistically significant results, i.e., the treatments appear to have increased the proportion of observations with 'perfect' dry waste segregation.

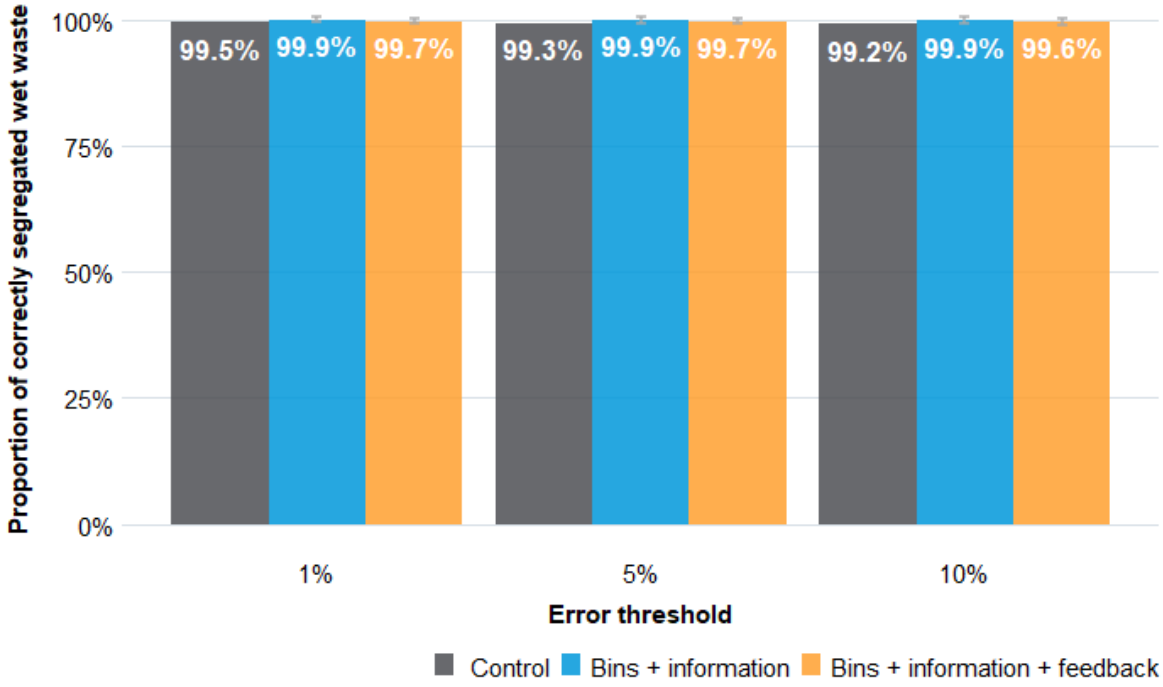
Segregation of wet waste

Primary outcome II: Proportion of correctly segregated wet waste

The second primary outcome measure was the proportion of correctly segregated wet waste in the wet bin (i.e., weight of correctly segregated wet waste, divided by total waste in the wet bin).

The results of the analyses, using a 1%, 5% and 10% error threshold are presented in Figure 12 below.

Figure 12: Segregation of wet waste using 1%, 5% and 10% error thresholds



The proportions shown in the figure above are the average proportions of the households in each condition. For each household, the observations from the three weeks are averaged to produce a single measure for each household.

With a 1% error threshold, both treatments had positive but not statistically significant effects. With a 5% error threshold, both treatments had positive effects, with the models including covariates indicating a statistically significant effect at the 5% level for the 'Bins + Information' treatment arm. However, this effect does not persist once multiple comparisons are corrected for. With a 10% error threshold, both treatments had positive effects, with all models indicating a statistically significant effect at the 5% level for the 'Bins + Information' treatment arm. However, this effect does not persist once multiple comparisons are corrected for.

Given the small effect sizes and the lack of significance once multiple comparisons are corrected for, it is reported that the treatments had no effect on this outcome measure.

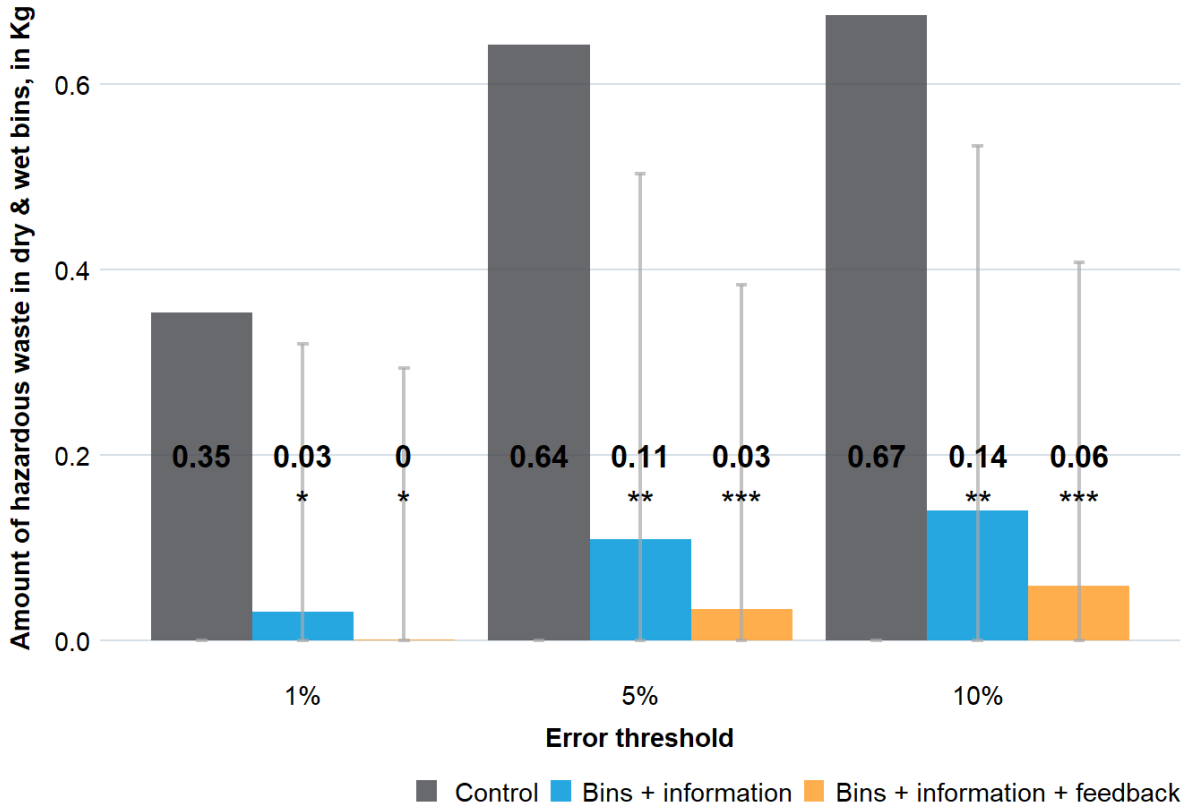
This may be because households are already mostly correctly segregating their wet waste. It may be that it is easier for households to distinguish what is 'wet' waste from other types of waste, than it is for households to distinguish 'dry' waste from other types of waste.

Segregation of HH hazardous waste

Secondary outcome I: Amount of HH hazardous waste in dry and wet bins combined

The secondary outcome measure was the amount of HH hazardous waste in the dry and wet bins combined. The results of the analyses, using a 1%, 5% and 10% error threshold are presented in Figure 13 below.

Figure 13: HH hazardous waste in dry and wet bins combined using 1%, 5% and 10% error thresholds



The amounts shown in the figure above are the average amounts of waste from households under each condition. For each household, the observations from the three weeks are averaged to produce a single measure for each household.

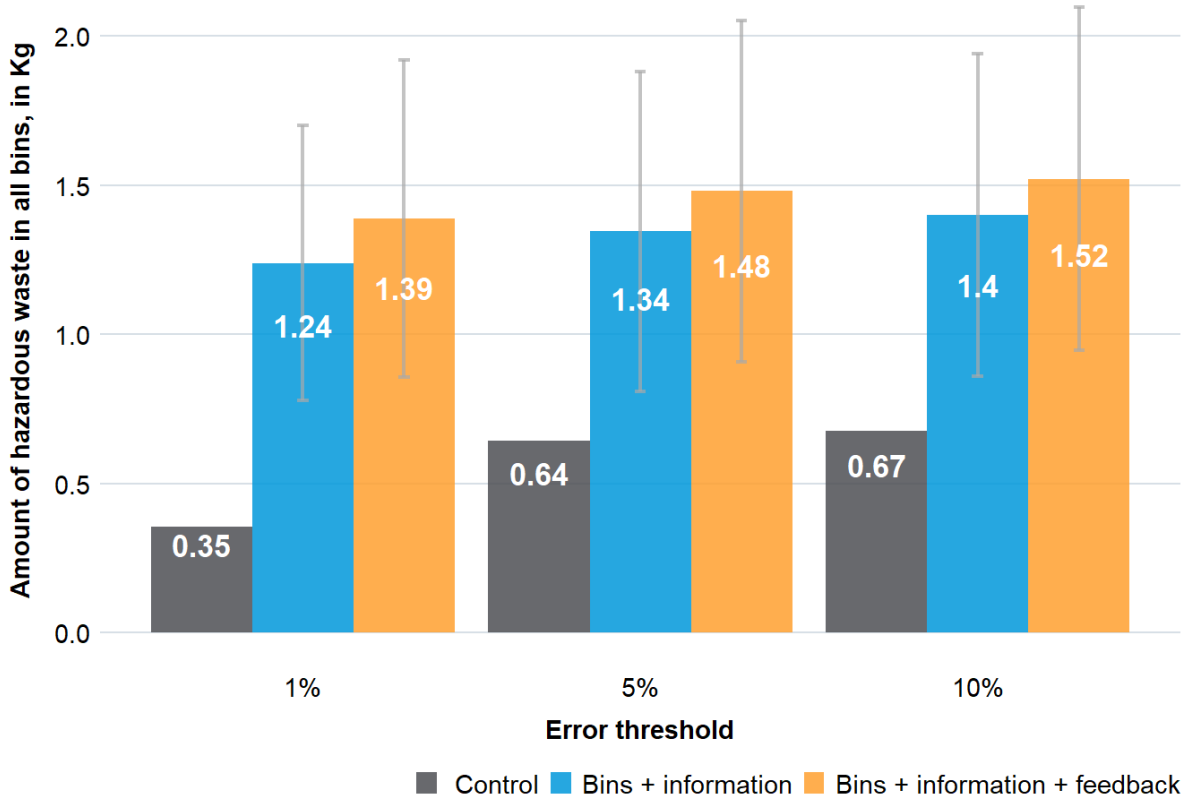
With a 1% error threshold, both treatments had negative and statistically significant effects at the 5% significance level, i.e., the treatments appear to have reduced the amount of HH hazardous waste incorrectly placed in dry and wet bins. With a 5%, and 10% error threshold,

both treatments had negative and statistically significant effects at the 1% significance level. All results continue to be statistically significant after applying the Hochberg procedure for multiple corrections.

Exploratory outcome I: Total amount of HH hazardous waste

The first exploratory outcome measure was the total amount of HH hazardous waste in all three bins combined. The results of the analysis, using a 1%, 5% and 10% error threshold are presented in Figure 14 below.

Figure 14: HH hazardous waste in all bins using 1%, 5% and 10% error thresholds



The amounts shown in the figure above are the average amounts of waste from households under each condition. For each household, the observations from the three week trial period are averaged to produce a single measure for each household.

Regardless of the error threshold used, the treatments appear to have increased the total amount of HH hazardous waste disposed of by households. This is interesting considering the reduced amount of HH hazardous waste in dry and wet bins in the treatment conditions. It appears that when provided with a HH hazardous bin, not only do households transfer HH hazardous waste from dry and wet bins into it, but they also begin disposing of more HH hazardous waste.

It is unclear where this waste was being deposited in control households. There are a range of possibilities including:

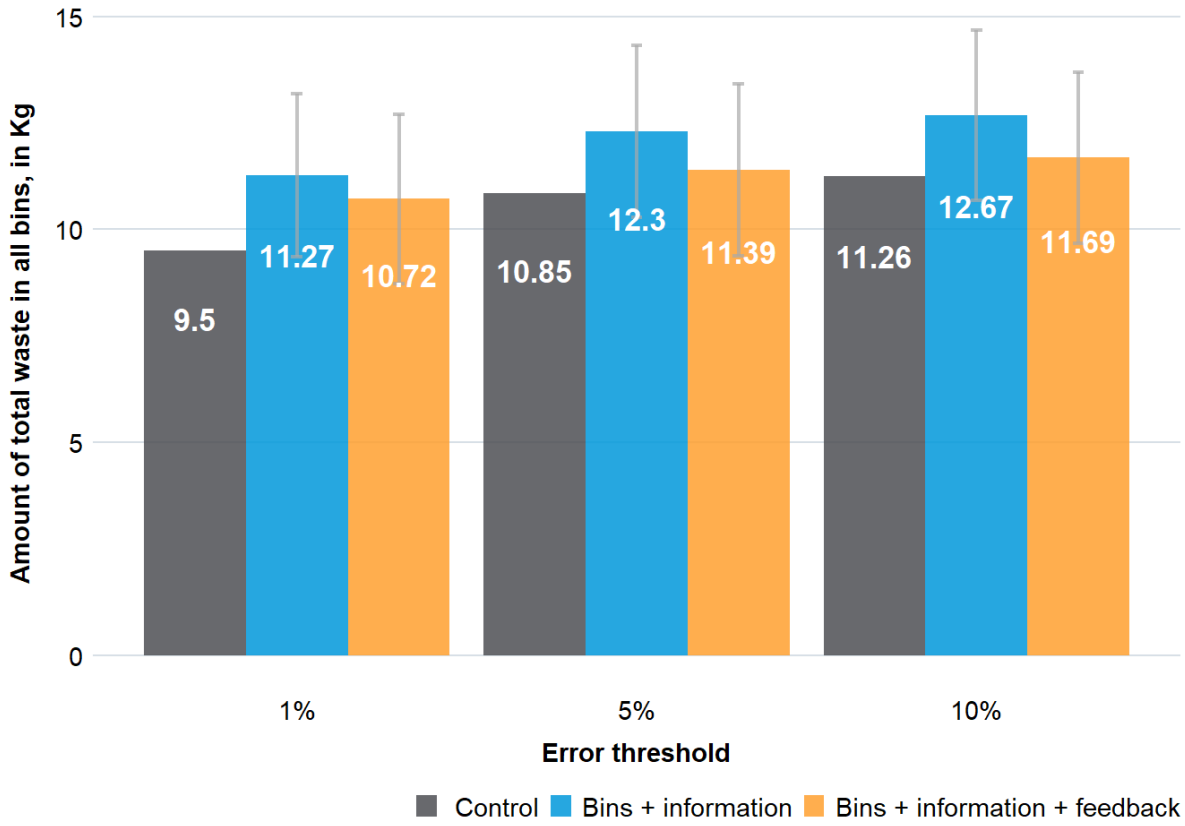
- The gap reflects waste that control households disposed of correctly in other ways - for example, at drop-off centres.
- The gap reflects waste that control households disposed incorrectly, for example, by littering in the street.
- The gap reflects Hawthorne effects from the trial. That is, households in the control condition would normally dispose certain types of HH hazardous waste in dry/wet bins, but did not beacuse enumerators were collecting waste directly. This waste may have been held on to, for disposal via the normal dry/wet collection.

Total waste

Exploratory outcome II: Total amount of waste

The second exploratory outcome measure was the total amount of waste in all three bins combined. The results of the analyses, using a 1%, 5% and 10% error threshold are presented in Figure 15 below.

Figure 15: Total waste in all bins using 1%, 5% and 10% error thresholds



The amounts shown in the figure above are the average amounts of waste from households under each condition. For each household, the observations from the three weeks are averaged to produce a single measure for each household.

It appears that the treatments resulted in slightly higher amounts of waste being disposed of than the control condition. However, noting the increase in the amount of HH hazardous waste disposed of in the treatment conditions, a significant portion of this may have been due to

households in the control condition disposing of HH hazardous waste in other ways (i.e., not providing it to enumerators at collection).

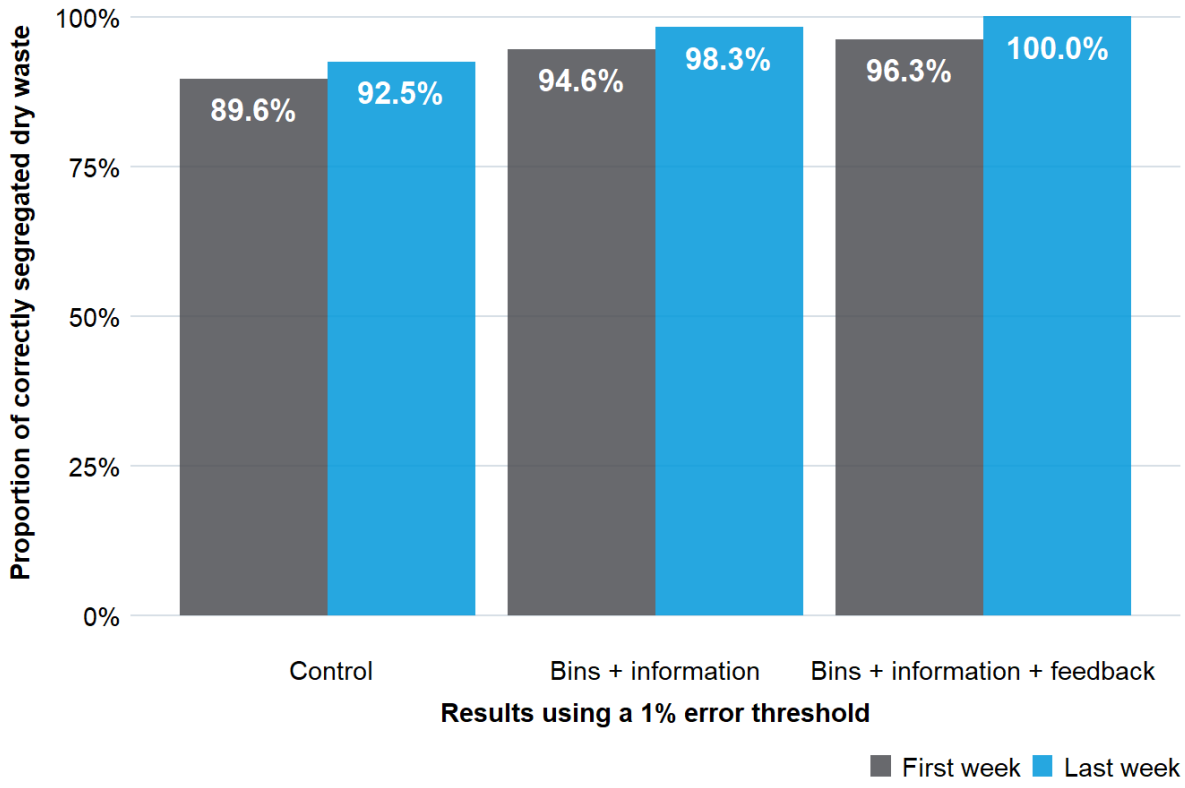
The end-line survey found that most households indicated the household bins were the right size.[16] About a third of households indicated the green wet waste bin was too small.

Segregation across time

Exploratory outcome III: Segregation of dry waste across time

The third exploratory outcome measure was the proportion of correctly segregated dry waste in the first week versus the last week. The results of the analysis, using a 1% error threshold are presented in Figure 16 below.

Figure 16: Segregation of dry waste in the first week vs the last week using 1% error threshold



The proportions shown in the figure above are the average proportions of the households in each condition, in the respective week.

Using a 5% and 10% error threshold yields similar results. Overall, regardless of the error threshold used, it appears that segregation improved by approximately four percentage points uniformly across all three trial arms. It is unclear why this occurred. This may reflect

[16] Between 80% and 90% of households indicated the red HH hazardous waste bin and the blue dry waste bin were the right size, and 66% of households indicated the green wet waste bin was the right size.

monitoring effects from the trial, as households felt the need to segregate better when they realised their waste was being collected directly by an enumerator.

5) Recommendation

The trial proved that providing households with three bins (with information stickers) for free improves household segregation behaviour, particularly in terms of not putting HH hazardous waste into their dry waste bins.

The household perspectives survey showed that the majority of households found the following bin sizes were the right size:

- 7 litre red bin for HH hazardous waste;
- 7 litre green bin for wet waste; and
- 40 litre blue bin for dry waste.

It is therefore recommended that each household be provided with three bins (with stickers) for free.

It is understood the Government may have limited funds to spend on supplying households with materials. In this case, it is recommended that the Government:

1. Provide each household with a HH hazardous bin (with an information sticker) for free, when the HH hazardous waste collection service commences.
2. Provide each household with dry and wet information stickers for free.
3. Make it extremely easy for households to buy matching dry and wet bins for a small fee.

For more details refer the Policy Note.

Annex A: Definitions

Some key technical terms used in this report are defined in the table below.

Term	Definition
Statistically significant	A result is statistically significant when it is very unlikely to have occurred by chance. For this trial the threshold for statistical significance was set at 5%. That is, a result is considered statistically significant if there is a less than 5% probability that this result would be seen under normal circumstances.
Percentage point	A percentage point is the unit for the difference between two percentages. For example, moving up from 40% to 44% is a 4 percentage point increase.

Annex B: Sample size per condition

The table below shows the number of households that were included in each condition, by the error threshold that was used.

Error threshold	Control	Treatment 1	Treatment 2
1%	148	157	155
5%	157	155	160
10%	160	168	163

Annex C: Baseline report