Pre-Analysis Plan for the study "Personal and social norms effects on pro-environmental behavior"

Our primary outcome variable measures weekly participation in waste sorting (i.e., a dummy variable taking value one every week that there is a disposal and zero otherwise).This variable captures regular participation in waste sorting, and it is the most reliable outcome that we can obtain with the available technology (which measures bin usage but not the content of each disposal).
Considering the panel data structure and the staggered recruitment, the treatment effect can be analyzed by aggregating periods or week-by-week. The first analysis summarizes the treatment effect for the whole period studied while the second breaks it down in different time periods and account for its evolution over time.

1 Aggregating Periods

Considering the panel data structure of our data and the staggered recruitment, the treatment effect can be analyzed by aggregating periods or week-by-week.
The aggregated specifications consists in averaging the outcome variable over the studied period (at least 15 weeks after the intervention).
By doing so, the outcome variable will capture the proportion of weeks recycling. By following this approach, the estimates for the effect of the treatment will be obtained through an OLS estimation of the following equation:

$$y_i = \beta_0 + \beta_1 SN_i + \theta X_i + \epsilon_i \tag{1}$$

where SN_i is an indicator taking value 1 if household i received the information about others' social norms and zero otherwise and ϵ_i is an error term. The term X_i is a set of household-specific controls that include: number of household members, personal norms and social norms obtained from the questionnaire.[1] We will also include as a control the index on preferences for altruism to compare the size ofs its associated coefficient to the one of social norms. Other controls that will be included if they available are household income and distance to the bin.[2] Controls will also include a set of dummies for postal code and recruitment week. Recruitment week fixed effects make sure that our estimates compare treated and control households that were recruited in the same week. For robustness we will control for particpating households' waste sorting before being recruited (i.e., the pre-recruitment value of the dependent variable). Since we have random treatment assignment, these controls should mainly improve the accuracy of our estimates. The main coefficient of interest is β_1, which captures the average treatment effect across the studied period. Considering the fractional or count nature of our outcomes, the residuals do not follow a normal distribution. Thus, the usual Huber-Eicker-White sandwich correction for standard errors will be applied.

2 Week-by-Week

By comparing the outcome the control and treatment group week-by-week, we can examine the dynamics of the treatment effect. To do so, we will divide the post-intervention period over 4 sub-periods (with 25% of the weeks located in each period). This leads to the following regression that will be estimated by a random effect model:

$$y_{it} = \beta_0 + \beta_1 SN_i + \beta_2 W_{it}^2 + ... + \beta_5 W_{it}^4 + \beta_6 SN_i * W_{it}^2 + ... + \beta_9 SN_i * W_{it}^4 + \theta X_i + \epsilon_{it} \tag{2}$$

[1] The primary variable used for personal norms will be the answer to the item "I consider myself ecologist" (5 possible answers from "totally disagree" to "totally agree"). The remaining controls for personal norms (items 2 to 4 in the questionnaire) will only be included as a control if their collinearity with the primary personal norm variable is low (correlation below 60%). The primary variable for social norms will the answer to the item "Out of every 10 people surveyed, how many do you think respond that they "agree" or "totally agree" with the statement "I consider that people in Palma should recycle their bio-waste"? (11 possible answers from 0 out of 10 to 10 out of 10).

[2] Household income and distance to the closest bin are not directly obtained from the questionnaire but could be respectively obtained with secondary data (census tract proxy) and by geolocation. These two controls will be available if having enough research funds for retrieving them.

Under this specification y_{it} is our primary outcome variable measuring weekly participation in waste sorting (i.e., a dummy variable taking value one every week that there is a disposal and zero otherwise).The variable W_i^j takes value 1 when the observation comes from period j and zero otherwise with $j = 1, 2, 3, 4$. X_i is the same set of household-specific controls as in the computation of the aggregate effect.

3 Other outcomes

To analyze the treatment effect on secondary outcomes, i.e. self-reports on the intention to act and beliefs about others, we will follow equation 1 (aggregate period).

4 Moderators

Providing information about others' social norms is expected to increase willingness to recycle of households that underestimate the beliefs of their co-citizens. Accordingly, we will interact a measure of discrepancy between beliefs and treatment to test such mechanism.
We expect that a majority of the participants underestimate the prescriptive social norms.

5 Analysis of the Survey on Social Norms

For analyzing the correlation between personal and social norms we will follow the aggregate approach. We will do so before and after answering the questionnaire.