

# ANNEX 2

# Technical details of the research design Limitations



# Content

Annex 2. Technical details of the research design .....	3
Limitations .....	3
Self-selection bias .....	3
Sample size .....	3
Statistical power .....	3
Variation in treatment implementation conditions .....	5
External validity.....	5

## **Annex 2. Technical details of the research design**

### **Limitations**

Some aspects to be taken into account when interpreting the results of this impact evaluation are detailed below. It is important to mention that most of the limitations of this experiment were linked to the emergency situation caused by COVID-19.

#### **Self-selection bias**

As described above, the lack of direct access to construction workers by the research team made it necessary to make a call through publications in social networks. In order to access their contact and socioeconomic data, participants were asked to fill out a registration form.

Therefore, the sample presents self-selection bias, given that the participants decided to participate in the study, since they self-selected at the time of completing the enrollment form. In other words, this bias may imply that the participants may have previously had some kind of interest in social security.

Although random assignment of the 281 participants to each type of treatment was then performed, it is important to keep this limitation in mind when interpreting the results.

#### **Sample size**

The study was conducted with a sample of 70 individuals. Despite the fact that the three assigned groups have optimal swing conditions, the sample size may interfere with the results obtained, especially those referring to the statistical power of the estimates. In other words, it should be taken into account that the effects detected may be greater, or that there may be some effects that have not been significantly quantified due to the low number of individuals.

#### **Statistical power**

As mentioned above, the sample selection strategy, as well as its size, must be taken into account in the interpretation of the results. However, despite the limitations, it is important to note that the 70 participants who finally participated in the experiment were part of randomly selected groups and the different descriptive statistics presented show the balance of the sample (Table 6). Although it was not possible to perform the power calculations related to the sample size, due to the limitations described above, the calculation of the minimum detectable effect (MDE) of the experiment is presented below.

The *minimum detectable effect* (MDE) is the smallest effect, which, if true, has a probability of P% of producing an impact that is statistically significant at a level of  $\alpha$ , where P is the statistical power of the experiment and  $\alpha$  is the level of statistical significance (Bloom, 1995). To perform this calculation, the *SEM approach* is used, which consists of setting the statistical power and sample size that the study can afford and determining the minimum detectable effect associated with these parameters.

Following Djimeu and Houndolo (2016), the following formula is used to calculate the MDE.

$$MDE = (t_1 + t_2) \sigma_y \sqrt{\frac{1}{p(1-p)n}}$$

Where:

- $t_1$  is the t-value of the desired  $\alpha$  significance level of the experiment in a one-tailed or two-tailed test. Generally, in social sciences,  $\alpha=0.05$  and given the nature of this experiment<sup>6</sup>, a two-tailed test is chosen.
- $t_2$  corresponds to the desired level of statistical power of the experimental design. It is a function of the desired statistical power (P) and the sample size (n). In the case of this intervention, it uses  $P=0.80=80\%$ , as suggested by the literature on experimental designs in social sciences and  $n=70$ .
- $\sigma_y$  is the total standard deviation of the estimated effect on the dependent variable. To obtain this parameter, we used the standard deviation of an index of knowledge about IPS, calculated with data from the Longitudinal Survey of Social Protection applied in 2015 to 15173 people aged 15 years or older within the entire Paraguayan territory. To construct this index, three questions were selected from this questionnaire: 1) Do you know what percentage of your taxable income is deducted (deducted or would be deducted) monthly for your contribution to the IPS or retirement fund? 2) Do you know at how many years of age a member of the IPS or retirement fund can retire or pension? 3) Do you know how the monthly pension or retirement is calculated? The PCA technique was used to assign a weight to each question. The Cronbach's Alpha test yielded a result of 0.64, which indicates an acceptable consistency for the index created. The results obtained in the calculation of this index are shown below:

Mean	Standard deviation ( $\sigma_y$ )	Minimum	Maximum
0.2234	0.4398	0	1.730

<sup>6</sup>A two-tailed test is recommended when researchers cannot predict a priori whether the intervention will have a positive or negative effect, but can hypothesize that there will be an impact (Djimeu and Houndolo, 2016).

Therefore,  $\sigma_y=0.4398$  was used as the parameter to perform the MDE calculation.

- $p$  is the proportion of the sample that is randomly assigned to treatment. In the case study, 46 individuals were assigned to treatment and 24 to control group. Hence,  $p=0.666$ .

Based on these parameters, the MDE of the experiment is 0.3. In other words, the minimum and statistically significant effect at 5% that has a probability of 80% of being detected by the study performed is 0.3. Given that the index taken as a reference to calculate the MDE has a maximum value of 1.73, it can be concluded that the minimum effect that can be detected by the experiment is 17.34% change over the base value ( $0.3/1.73*100=17.34\%$ ). That is, the experimental design used has the capacity to detect from a 17% change in the indices that may arise as a result of the intervention in the treatment groups, with respect to the control group. In this sense, it should be noted that this result is in line with the results obtained in the evaluation presented in this paper.

## Variation in treatment implementation conditions

One of the main consequences of the COVID-19 pandemic on the intervention proposed for this learning loop refers to the conditions of implementation, since, due to the restrictions imposed to reduce the circulation of the virus, it was not possible to carry out the different treatments in a single day.

Although the research team ensured compliance with all the requirements designed for each intervention, it is important to mention that the treatments could have suffered certain differences in terms of schedules, physical spaces, etc. It should also be noted that the two days of treatment 2 offered were given by different people, due to an outbreak of the virus within the facilitation staff.

Although the impact of these conditions on the intervention cannot be determined, they should be taken into account when interpreting the results.

## External validity

Finally, it should be noted that, due to all the limitations mentioned above, the results obtained in this analysis refer only to the 70 participants in the intervention and cannot be extrapolated to other individuals or populations. Furthermore, it cannot be inferred that an experiment of this type could have the same effects when applied by other experimenters and to another target population.