Low taxes and informative nudges in the control of negative externalities: substitutes or complements?

Carolina López

Abstract: *We present the results of a series of public-bad laboratory experiments in which we test whether an informative message of suggested behavior with an implicit moral appeal, and a tax that is insufficient to induce the optimal level of the externality, can complement each other when implemented jointly, or not, and by how much. Our results confirm that the average subject, (a) behaves consistently with having moral preferences, (b) is "nudgeable" by such a such a message, (c) exhibits preferences that are non-separable from the choice of a tax as instrument (the tax crowds-out part of its moral preferences). In what is our main contribution, we find that the tax and the nudge exhibit a negative synergy. Additionally, when comparing the relative effects of both instruments, we find that the tax has a greater impact on reducing the externality than the nudge, even though the tax was only half of what would be needed to reach the socially optimal level. The tax appears to complement the nudge, but not the reverse, challenging the policy suggestion that nudges can effectively complement low taxes while awaiting the political will to raise them.*

Keywords: Economic experiment, nudge, public bad, tax.

Acknowledgements: this study was made possible by the funding of Agencia Nacional de Investigación e Innovación (ANII), grants FCE_1_2019_1_156336 and POS_FCE_2020_1_1009209.

1 Introduction

Most economists have traditionally favored taxes or tradable permits (prices) over other policy instruments to reduce negative environmental externalities (Fuller and Geide-Stevenson, 2003; Alston et al., 1992; Ricketts and Shoesmith, 1992; Frey and Eichenberger, 1992). This support continues to be strong (Climate Leadership Council, 2004; Ilzetzki and Jain, 2024; van den Bergh, and Botzen, 2024). Possibly as a result of policy recommendations from economists, these instruments have been widely adopted in the past (Harrington et al., 2004; Sterner and Coria, 2013) and continue to do so in the present (World Bank, 2024).

Although the number of price-based emission control programs is increasing, due mainly to political-economy reasons, prices for pollution are frequently set at low levels. Carbon pricing is a recent and important example of this. Around the world, carbon prices range between less than one US dollar in Poland to 167 USD in Uruguay (World Bank, 2024). Most of these values are lower than the recent estimates of the social cost of carbon. According to Tol (2023), estimates of the social cost of carbon have increased from USD 122 per tCO2 to USD 525 in the past 10 years, for a relatively low discount rate. More recently, Bilal and Känzig (2024) put the social cost of carbon above USD 1,000. Moreover, values of the implemented taxes are also lower than the levels needed to induce an abatement of greenhouse gases (GHG) sufficient to avoid exceeding the 2ºC target of the Paris Agreement, according to recent estimates of their rather low impact on GHG emissions (van den Bergh and Savin, 2021; Lilliestam et al., 2021).

This situation has led several economists to argue, more recently, that carbon prices should be part of a broader set of instruments (Carattini et al, 2018; Blanchard, et al., 2023; Stiglitz, 2024; Sterner et al, 2024).

Nudges seem to be a good candidate for complementing prices (Carlsson and Johansson-Stenman, 2019). First, since the publication of the popular book by Thaler and Sunstein (2008), we have witnessed an impressive expansion of both academic studies and actual implementation of nudges by governments, to deal with "internalities" and externalities. Even though their effectiveness can vary to a great degree (DellaVigna and Linos, 2022), and that we know little about what type of nudges work under what conditions (Szaszi et al., 2018; 2022), on average, nudges interventions do work (Mertens et al., 2022). This is true also for green nudges (Gravert and Olsson Collentine, 2021), which have made their way into the environmental and energy literatures and policies (Carlsson et al., 2021).

In addition to their effectiveness, there is evidence that nudges are cost-effective means to change behaviors (Hahn et al., 2024).

The observation that nudges could complement prices in environmental policy is not new. Thaler and Sunstein (2008) state that: "… the most important step in dealing with environmental problems is getting the prices … right". Nevertheless, "…such an approach is politically difficult" (pg. 190). "(A)long with getting the prices right (or while we are waiting for the political courage to set the prices right), we should take other

nudgelike steps that can help to reduce the problem in politically more palatable ways." (pg. 191). A similar point was made by Carlsson et al., (2021): "...if the existing tax is too low, … a nudge could play a bigger role for policy. … a pure green nudge could be used to complement the tax so that the combination mimics the outcome that would result from an optimal tax…. With a non-optimal tax, there is also more room for moral nudges." Concluding that "(w)hen taxes are not set optimally, both moral and pure nudges can be efficiency improving complements to taxes".

A challenge that this recommendation faces is that comprehensive theoretical models that may guide how policy makers should think about the interactions of a nudge and a price in the control of a negative externality have yet to be developed. In an early effort, Stern (1999) provides a conceptual framework, listing the factors that affect the pro-environmental behavior of consumers and its policy implications. He concludes that, due to their synergy, providing consumers with information as well as material incentives in combination may have a greater effect than the sum of their own effects. With a caveat: it is only "once incentives are large enough for consumer to take it seriously" (p. 474) that it may be more effective to invest in information than to increase the incentive.

The empirical evidence regarding the complementarity of nudges and prices on negative externalities is also rather limited. Moreover, it is inconclusive about the size and the direction of the synergy (Drews et al., 2020). Some of the studies show that a nudge could add to the effect of a price increase (Hernández et al., 2024; Panzone et al., 2021; Hilton et al. 2014; Mizobuchi and Takeuchi, 2013; Spraggon and Oxoby, 2010; Ambec et al., 2024). Others show that prices totally crowd out the effect of the nudge (Sudarshan, 2017; Mackay et al., 2019), or vice versa (Dolan and Metcalfe, 2015). Finally, another works perfect complementarity (zero synergy; Fanghella, et al. 2021; Schall et al. 2016). (See section 2 for a more detailed summary of the literature). Furthermore, most of the existent evidence corresponds to the complementarity between a tariff increase and a nudge informing households what the level of consumption of water and electricity of similar households is.

The lack of evidence on the complementarity of prices and nudges in the control of a negative externality is particularly binding in the domain of the interactions of nudges and environmental taxes. We know little about whether a nudge could complement a low tax on an externality, and if so, by how much. Contributing to answering these two questions is the main motivation of this work. To the best of our knowledge, ours is the first paper to do so. Moreover, it does so in the light of a model in which economic incentives may crowd out social preferences and the social planner is "sophisticated" enough to recognize this (Bowles, 2016). There is ample evidence of motivational crowding by economic incentives (see Frey and Stutzer, 2008, and Bowles and Polania-Reyes, 2012). If economic incentives may crowd out the same social preferences that the nudge wants to trigger, allowing for non-separability between economic incentives and social preferences is crucial to inform a regulator on the synergy between a nudge and a tax on controlling a negative externality.

To assess the degree of complementarity and possible synergies between a nudge and a price, we conduct a series of public bad laboratory experiments and test whether the joint effect of (a) a message informing subjects what the optimal level of a negative externality is, and (b) a tax that is insufficient to induce the optimal level of this externality, is higher or lower than the effect of the tax and the nudge standing alone. This set of experiments seeks to mimic a local public bad situation, such as garbage disposal or air pollution in a town or neighborhood, in which socially disconnected individuals contribute to a pollution problem.

Providing information is one of the most important means of nudging people (Sunstein, 2014). Information provision can take several forms. A commonly used type of message is an injunctive one: to inform what other people are doing and communicate approval or disapproval to the person´s relative behavior. The majority of the messages in the most relevant literature has these features. Notwithstanding, information that consumers and citizens receive about the environment regularly comes in the form of information about the state of the environment and the environmental impacts of consumption choices. In these messages, explicitly or implicitly, citizens receive tips or prescriptions on how to behave in order to avoid these negative impacts on the environment or on other citizens, in a significant and welfare decreasing way. Our message is a simple, informative nudge, mimicking this situation. The message does not have an explicit moral appeal. Nevertheless, it does implicitly call for a subject to behave consistently with the group´s welfare. In this sense, it may classify also as a suggestion. An example could be the information provided by local air quality offices under poor air quality conditions in urban, central-southern Chile in winter. During the mildest episodes of a three-tyer system of environmental warnings, local air quality offices issue recommendations and advocate for the efficient use of heating stoves. Messages of "suggested play" have proved to be effective in the presence of heterogenous preferences over the public good in question (Marks et al., 1999; Croson and Marks, 2001) and particularly, when it is combined with moral suasion (Dal Bó and Dal Bó, 2014). Ours could be considered an informative message of suggested behavior with an implicit moral appeal based on utilitarianism.

Our results show that our average subject (a) behaves consistently with having moral preferences, (b) is "nudgeable" by this type of message, (c) exhibits preferences that are non-separable from the choice of a tax as instrument (the tax crowds-out part of its moral preferences). When applied separately, the average and median effect of the insufficient tax is higher than that of the message. In what is our main contribution, we find that the tax and the nudge exhibit a negative synergy: the effect is lower than the sum of the two effects. This effect is observed even among individuals with "stronger" baseline social preferences who were more influenced by the nudge, implying that, to some extent, the tax "crowds-out" the impact of the nudge.

The paper proceeds as follows. In the second section, we present the most relevant literature. The third section presents and explains the theoretical model used to carry out

this study and the hypotheses that guide our research. Section 4 presents the experimental design, treatments and expected results, and the procedures we used to implement our experiments. In Section 5 we present the results of the experiment. Finally, in Section 6 we conclude.

2 Most relevant literature

In this section, we review the most relevant literature for our work. We divide this in two: (1) a literature that compares the effect of nudges versus prices when implemented independently, and (2) a literature that tests the effect of nudges and prices implemented simultaneously.

2.1 Nudges versus prices

The empirical literature that compares the effect of a nudge and a price, but not the complementarity of them, basically applies a nudge and a price to different sets of subjects, and compare the effects (Nakagawa, et al., 2022; Buckley and Lerena, 2022; Bucholz et al., 2021; Antinyan et al., 2020; My and Ouvrard, 2019; Xu et al, 2018, Ito et al., 2018; Delaney and Jacobson, 2016; Romaniuc, 2016, López et al., 2012). These studies vary in several important features, such as the context of the test (agricultural production, household consumption, waste separation, fishing), the externality being targeted (pesticide use, waste generation, consumption of positional goods, overuse of a CPR), the type of nudges (traffic light labelling, communication, demonstration, information, graphics, normative messages, private or public disclosure of individual behavior), the type of prices (taxes, subsidies, increase in tariffs, redeemable points to be used in shopping, automatic penalties, uncertain penalties), and the level and design of the prices being tested. Beyond these differences, general results do emerge; both prices and nudges seem to be effective, but the effect of prices is generally higher and more persistent in time. An exception to this is López et al. (2012), who found that, in framed field experiments in fishing communities in Colombia, expected penalties (the prices) high enough to induce efficient levels of contributions to the public good produced lower contributions than the threat of public disclosure of individual behavior (the nudge).

2.2 Nudges plus prices

The empirical literature on the degree of complementarity between a price and a nudge to reduce an externality generating activity is rather limited.

The studies vary in several features. One of these is the externality the intervention aims to control. A number of them studied the effect the interventions on GHG emissions, either through choices in groceries shopping (Panzone et al., 2021; Ambec et al., 2024), transport means (Hilton et al., 2014), fuel consumption in driving (Schall et al., 2016), and electricity consumption (Mizobuchi and Takeuchi (2013), Sudarshan (2017), Dolan and Metcalfe (2015), Fanghella et al (2021)). Other focus on residential water consumption (Hernández et al., 2023), recreational fishing (Mackay et al. (2019)),

volunteering for nature restoration (Maris et al. (2024)). Finally, a particular case is Spraggon and Oxoby (2010), whose framework is a public good in a lab.

Another feature in which these studies differ is in the type of nudge tested. In five, the nudge tested is a social comparison (Hernández et al. (2023), Mizobuchi and Takeuchi (2013), Sudarshan (2017), Dolan and Metcalfe (2015), Fanghella et al., 2021). In the rest, nudges include detailed instructions (Spraggon and Oxoby, 2010), an injunctive message (Hilton et al., 2014), information and tips (Schall et al., 2016), an environmental recall (Panzone et al., 2021), goal setting and feedback (Fanghella et al., 2021), traffic light labelling (Ambec et al., 2024), and moral suasion (Maris et al., 2024).

Papers also differ in the amount and the type of economic incentive. Most of the studies test the effects of a change in an underlying price. Only three of them (Maris et al., 2024; Schall et al., 2016; and Spraggon and Oxoby, 2010) test the introduction of an economic incentive on a previously unpriced activity. An increase in expected fines is the price studied by Mackay et. al. (2019). Among the studies that introduce a change in the underlying price, the mechanism for introducing it also differs. Panzone et al. (2021), Ambec et al., 2024 and Spraggon and Oxoby (2010) introduce taxes, Sudarshan (2017) a reward/penalty scheme, Mizobuchi and Takeuchi (2013) and Dolan and Metcalfe (2015) a reward, Hilton et al. (2014) a bonus/malus price scheme, Hernández et al. (2024) a tariff increase.

Finally, some of the studies in this literature have an incomplete design, testing the effect of the two instruments when implemented jointly against that of a stand-alone nudge, but not against a stand-alone price, or vice versa. The studies with an incomplete design provide no clear general results on the effect of combining the two instruments. Sudarshan (2017), the only example in this set of papers whose design lacks a stand-alone price variation treatment, finds that communicating to households in India what the average consumption of electricity by similar households is, plus providing them with energy saving tips, produced a reduction in electricity consumption of 7%. Nevertheless, households did not change the level of electricity consumption when this communication was accompanied by a reward/penalty scheme on top of the underlying price. Among the papers lacking a treatment with a stand-alone nudge in their design, Mizobuchi and Takeuchi (2013) did not observe a statistically significant difference between the effect of the joint implementation of a social comparison and a reward on top of the underlying price and the effect of the reward standing alone. Contrary to this apparent lack of complementarity between a nudge and a price instrument, Spraggon and Oxoby (2010) found that providing subjects in a public good game in the lab with enhanced instructions increased the efficiency induced by the price instruments. In another laboratory experiment, Mackay et. al. (2019) found that the communication of the average catch by others increased compliance to a catch limit by 10% when the deterrence level was low, but it did not decrease it in a statistically significant way in a higher deterrence situation.

It may be argued that the lack of general results on the synergies between nudges and prices is the result of incomplete designs of these papers (Drews et al., 2020).

Nevertheless, it is not. If we focus only on the studies that have a complete design (a nudge treatment, a price treatment, and a treatment with both) and are therefore able to identify and test the synergies of the nudge and the price, no consistent evidence emerge either.

Dolan and Metcalfe (2015) found that social comparison decreased electricity consumption by 6%, a £100 reward decreased it by 8%, but the two instruments had no effect when implemented together, suggesting a negative synergy or substitutability between the two types of instruments.

Moving up in the ladder of complementarity, in a hypothetical transport choice experiment in France, Hilton et al. (2014) found that a price increase for the plane ticket coupled with a price decrease of equal value for the train ticket had more effect on choices favoring the train when presented as a bonus/malus system due to the lower CO2 emissions of the train with an injunctive message. The result, nevertheless, did not hold for a higher bonus / malus price difference.

One step up the ladder, Hernández et al. (2024) found that a tariff increase of 12% for the first consumption block (1-20 m^3) and 38% for the second one (>20 m^3)) decreased household water consumption by 11%, with respect a control untreated group. On the other hand, a report containing a social comparison component and an informative component on the environmental effects of water use decreased water use by 7%. Finally, water consumption of the group of households that received the report and, six months after this nudge had ended saw their tariff increase, decreased 14% with respect to the control group of households. Because this decrease in water use is lower than the sum of the effects of the two instruments applied alone, the authors observed that there is a degree of complementarity between the two instruments, but this is not complete.

Panzone et al. (2021) used an experimental online grocery store to assess (a) whether being required to recall past environment-friendly behavior before shopping led consumers to purchase baskets of food products with lower carbon footprint, (b) whether the effect of this nudge ("environmental recall") is comparable to that of an income - neutral tax on the carbon content of foods (equivalent to £70/ton of CO2eq) and (c) how they interact when imposed together. They find that the environmental recall had a similar effect to that of the tax (both decreased the carbon footprint of a food basket by around 3 kg CO2eq) and when implemented together, the joint effect was the sum of the two separate effects, suggesting perfect complementarity with zero synergy. Ambec et al. (2024) conducted a similar experiment using a traffic light labeling nudge for carbon contents of baskets, with two levels of carbon tax (80€/ton CO2eq and 250€/ton CO2eq) with rebate. They found that only the combination of the nudge with the high tax significantly reduced the carbon footprint of the average basket. For the other treatments (low tax plus traffic light or taxes alone), they found no effect. Because they found more evidence of a reduction in the carbon footprint per euro spent, they attributed the reduction in carbon footprint to an income effect. Schall et al. (2016) report the same type of result,

although less promising. These authors found that a training course with fuel saving tips had no effect on fuel consumption by a company's drivers, contrary to non-monetary rewards, which had an immediate effect that attenuated over time. The joint implementation of both instruments showed no additional effect. In an experimental simulated environment, Fanghella et al (2021) also found a perfect complementarity with zero synergy between a monetary reward coupled with a nudge for goal setting. When implemented jointly, as when standing alone, these instruments had no effect on the consumption of electricity.

Finally, Maris et al. (2024) using a field experiment, found a positive synergy between and economic incentive of 50 NZD and a message highlighting its personal and environmental benefits on volunteering for nature restoration. They found that the monetary incentive significantly increased volunteering, the message alone was ineffective, but the monetary incentive became more than twice as effective when combined with the message.

In sum, among those studies with a complete design, we have an array of heterogeneous results. A set of results is consistent with a partial complementarity, negative synergy (Dolan and Metcalfe, 2015; Hernández et al. 2024). Another one is consistent with perfect complementarity and zero synergy (Panzone et al, 2021; Schall et al., 2016; Fanghella et al., 2021). A final set is consistent with positive synergy (Maris et al., 2024). Among those studies that report a negative synergy, we find no study reporting a backfire, but we do have results of total crowding out (Dolan and Metcalfe, 2015).

3 Theoretical Framework

In this section, we present the theoretical model from which we derive the hypotheses that we test with our experiments. The setting of our model is that of a local public bad, in which the aggregate level of a negative externality generated by a set of sources linearly affects every one of these sources in the same amount. Apart from experiencing the negative externality, we allow these sources to experience disutility for causing a public bad on third parties. The sources are polluters with moral preferences.

3.1 A moral polluter's behavior

Suppose that a subject generates a quantity e of emissions of a given pollutant. The subject could be the owner of a firm, a citizen, or a household that produces an unspecified good and generates emissions as a byproduct, or that incorrectly disposes of its waste. In either case, generating e is beneficial for the decision maker, because it saves her the higher costs associated with the use of more efficient, less polluting technology and inputs. In the case of waste, it saves her the additional effort needed to classify, temporarily store, and correctly dispose of it. Let $g(e)$ be the function that measures net economic benefits associated with generating e units of this externality. We assume that $g'(e_i) > 0$ and $g''(e_i) < 0$. Suppose that there are n sources of this pollutant, whether

firms, households, or citizens. The aggregate level of emissions of the n sources is $E = \sum_{i=1}^{n} e_i$. E is a public bad. It produces a negative externality (cost) of γE $(\gamma > 0)$ to *each* of the n subjects.

Following Levitt and List (2007) and Bowles and Polania – Reyes (2012), we assume that the utility that a subject derives from emitting is additively separable in profits and a moral term $M(e_i)$ that captures her moral benefit or cost associated with the action: $U(e_i) = g(e_i) - \gamma E + M(e_i)$. According to Levitt and List (2007), $M(e_i)$ captures the desire to "do the right thing". Different motivations may trigger this desire. In the words of Levitt and List, "*decisions that an individual views as immoral, antisocial, or at odds with his or her own identity… may impose important costs on the decision-maker*" (p. 156). Levitt and List (2007) focus on three factors that may influence the level of the "moral utility", each of which implies a different motivation. One is the size of the negative externality imposed on others. The higher the level of the externality that she generates, the higher the moral cost of emitting. Such moral utility (or disutility) may be the result of pure altruism, inequality aversion, or other types of "other regarding" preferences. It may also be the result of a purely selfish individual that experiences guilt or pride as a result of deviating or not from what she considers the correct level of action; her private moral threshold. Andreoni (1989) called this type of "selfish" moral preferences "impure altruism", commonly known as "warm glow" (p. 1448-1449). A second factor affecting M according to Levitt and List could be the existence of *social norms or legal rules* that govern the level of emissions in a society. A subject deviating from a social norm or not complying with a legal rule may also experience guilt, which increases moral disutility (Loewenstein and O'Donoghue, 2006). A final factor that may affect $M(e_i)$, is the extent to which the action causing the externality is or can be *scrutinized by others*. This motive is absent in our experiment, as the subject´s decision is completely anonymous. Because motivations behind the moral term $M(e_i)$ may differ, Bowles and Polania-Reyes (2012) call this term "social preferences", referring to motives that make subjects behave differently from what a purely profit-maximizing person would. This moral term $M(e_i)$ is consistent with the "personal domain" of Stern (1999), consisting of the individual´s basic values, social norms, emotions and beliefs about the functioning of ecosystems and the environmental consequences of its actions.

Building upon Allcott and Kessler (2019), who follow Glaeser (2006, 2014) and Lowenstein y O´Donoghue (2006), we model the moral term $M(e_i) = \mu_i[\varphi_i(m_i - e_i) + (1 - \varphi_i)(s - e_i)] = \mu_i[\varphi_i m_i + (1 - \varphi_i)s - e_i]$, where $0 \leq \varphi_i \leq 1$, m_i is the person´s individual moral threshold level of emissions (its intrinsic values, according to Frey and Stutzer (2008)), and s_i is the person´s perception of what the social norm about emissions is. The expression $\varphi m_i + (1 - \varphi_i)s_i$ is a weighted average of the individual moral threshold and the perceived social norm that represents the personal threshold level of emissions. Finally, the parameter $\mu_i > 0$ is a moral or psychological tax/subsidy for emitting above/below the personal threshold level of emissions. Modelled in this way, the individual moral price μ_i captures feelings of pride/guilt arising from deviating from the

level of emissions consistent with one´s personal values. Including this specification of the moral term, the utility function for each of our subjects is:

$$U(e_i) = g(e_i) - \gamma E + \mu_i[\varphi_i m_i + (1-\varphi_i)s - e_i] \quad (1)$$

Assuming that a subject chooses e_i to maximize (1); the first order condition characterizing the choice of *e* by this moral polluter is:

$$g'(e_i) - \gamma - \mu_i = 0 \quad (2)$$

Given our assumption that $g''(e_i) < 0$, this condition is sufficient to characterize an interior optimal choice. We call this optimal choice by a moral polluter in an unregulated setting $e_i^{um}(\gamma, \mu_i)$, where the "*u*" in the superscript refers to "unregulated" and the "*m*" refers to "moral". Note when setting $\mu_i = 0$, equation (2) characterizes the utility maximizing choice of emission by an amoral polluter, which we called $e_i^u(\gamma)$.

3.2 Social optimum with moral subjects

We now characterize the socially optimal distribution of emissions among the group of emitters. This is given by the set $(e_1, \dots, e_n)$ that solves the following social planner problem

$$\max_{\{e_1,\dots,e_n\}} \sum U_i = \sum (g_i(e_i) - \gamma E + \mu_i[\varphi_i m_i + (1-\varphi_i)s - e_i])$$

The set of first order conditions

$$g_i'(e_i) - \gamma n - \mu_i = 0, i = 1, \dots, n \quad (3)$$

These conditions implicitly define the socially optimum level of emissions with moral subjects, which we call $e_i^{mso}(\gamma, n, \mu_i)$.

Assuming decreasing marginal benefits $g'(e_i)$ as we do, two results are easy to show. First, $e_i^{mso}(\gamma, n, \mu_i)$ is lower that the socially optimum level of emissions with amoral subjects ($\mu_i = 0$), $e_i^{aso}(\gamma, n)$. Second, first order conditions (2) and (3) imply $e_i^{mso}(\gamma, n, \mu_i) < e_i^{um}(\gamma, \mu_i)$. Note that this is true even when the moral subject fully internalizes its marginal externality $(\mu_i = \gamma(n-1))$. The reason is that this form of morality gives rise to another social benefit or cost, additional to the public bad, as first noticed by Andreoni (1990). Alternatively put, the "impure altruistic" affects her behavior to take care of her private "warm glow" effect. This new private benefit decreases the privately chosen level of emissions with respect to the amoral subject ($e^{um} < e^u$), but it also decreases the socially optimum level of emissions ($e_i^{mso}(\gamma, n, \mu) < e_i^{aso}(\gamma, n)$), by the same amount. In this special case when $\mu = \gamma(n-1)$, the social planner's first order condition becomes $g_i'(e_i) - \gamma - 2\gamma(n-1) = 0$, which says that the social planner, in the margin, needs to account not only for the externality $\gamma(n-1)$ itself, but also for the emitter´s moral costs of causing the externality, also $\gamma(n-1)$.

3.3 Regulation

We now examine the response of moral polluters to a tax on emissions and nudges, designed to reduce the aggregate level of emissions from the unregulated level.

3.3.1 A nudge

Nudges may affect the individual´s moral utility M through three different mechanisms. First, they can affect the social norm s_i, such that $s_i = s_i(z)$, where z is the nudge. A social norm is a convention; this is what everybody expects others to believe (a normative expectation) or do (an empirical expectation) (Bicchieri and Dimant, 2022). A social norm is therefore a belief or an expectancy, and as such it can be modified by a nudge that takes the form, for example, of communication of what others believe or are doing. This type of nudge has been extensively studied in the literature (see for example, the literature on the Home Energy Reports for energy conservation).

Second, nudges may affect the person´s moral threshold, such that $m_i = m_i(z)$.[1] For example, a message highlighting the benefits of a healthy environment along with its current threats, may affect the person´s individual moral threshold level of emissions. A similar effect could have an informative nudge, providing citizens with tips or prescriptions on how to behave in order to avoid negatively impacting the environment, or other citizens.

Notwithstanding, it is clear from equation (2) that nudges that affect the perceived social norm or the individual moral threshold do not affect the individual´s choice of its level of the externality. For nudges to affect the choice of emissions in a Levitt-List moral utility function, they have to affect the moral price, μ_i. We model this effect a shift in the moral price, $\mu_i = \mu_i^0(1 + \mathbf{1}\{z > 0\}\mu_i^z)$, where μ_i^0 is the baseline moral price of the subject and the indicator $\mathbf{1}\{z > 0\} = 1$ if $z > 0$ (a nudge is implemented). The shift parameter μ_i^z measures the effect of the nudge on the moral price. μ_i^z could be of either sign, depending on the nudge, the subject and the situation. In the case of a public bad, and a nudge that is intended to decrease the negative externality, $\mu_i^z > 0$. This could be the case of nudges of the moral suasion type, that convey information that makes more salient the level of individual wrongdoing, and this increases a subject´s guilt (or shame if the action is being observed by others). It may also be the case of message of suggested play with an implicit moral suasion, as the example mentioned above an informative nudge that provides citizens with tips or prescriptions on how to behave (what is the socially optimum behavior). Finally, the effect of a nudge may differ between individuals. The interaction of personal traits and features of the message, such as its quality, the

[1] A nudge may affect only the social norm, while a different nudge may affect only the individual moral threshold. For ease, we only use a general parameter z to indicate the presence of a nudge.

technology used to deliver it and its frequency, determine the "nudgeability" of the subject. Hence the subscript *i* in the shift parameter μ_i^z.

When nudges may alter the moral price, the social norm, and the individual moral threshold, the Levitt-List utility function of a moral subject may be written as:

$$U_i(e_i) = g_i(e_i) - \gamma E + \mu_i^0(1 + \mu_i^z)[\varphi_i m_i(\mathrm{z}) + (1 - \varphi_i)s_i(\mathrm{z}) - e_i\,]$$

In this case, the optimal choice of emissions satisfies

$$g_i'(e_i) - \gamma - \mu_i^0(1 + \mu_i^z) = 0 \tag{4}$$

3.3.2 A tax on emissions

Assume the regulator sets a uniform tax *t* per unit of emissions. The utility function of the representative moral subject in this case is given by

$$U_i = g(e_i) - \gamma E - e_i t + \mu_i^0[\varphi_i m_i + (1 - \varphi_i)s - e_i\,]$$

The first-order condition that implicitly defines the level of emissions $e_i^{tm}(\gamma, \mu_i^0, t)$ that maximizes utility is

$$g_i'(e_i) - \gamma - t - \mu_i^0 = 0 \tag{5}$$

Comparing the first order condition defining the social optimum level of emissions (in (3)) with equation (5), we can conclude that the optimal tax for moral subjects should be set as:

$$t^m = \gamma(n - 1) \tag{6}$$

Note that this tax is equal to the classical Pigouvian tax in the case of amoral subjects. The morality of the subjects does not affect the level of the optimal tax, a result obtained by Johansson (1997). This is, again, because, whether motivated by "impure" or pure altruism, morality creates a new private utility/disutility, which the subjects consider when deciding how much to emit. For this reason, the marginal externality remains uninternalized.

3.3.2.1 Non-separability

The model above, by Levitt and List, assumes separability (no interaction) between the emissions tax and the social preferences of those regulated by the tax. However, this is contestable. There is substantial evidence indicating that the use and the size of economic incentives may affect people´s social preferences (see Bowles and Polania, 2012). This crowding of intrinsic motivation (Frey, 1992; Frey and Oberholzer-Gee, 1997)) enters our model through the moral price. The use of an economic incentive may affect the moral price of the citizens through different mechanisms. In the cases deterring negative

externalities, one of such possible mechanisms is that economic incentives trigger "moral disengagement" (Bandura, 1991). According to this interpretation, an economic incentive signals that the situation is not an ethical one but a market-like one, decreasing the moral price that a subject imposes on itself for taking a socially undesirable action. Another possibility is that a tax (as well as a regulation or appeal) may deactivate the moral norm guiding a behavior by depriving the subject of the personal satisfaction of acting according to one´s values (Schwartz, 1977). Also, an economic incentive can decrease intrinsic motivations because it can undermine people´s sense of self determination (Deci and Ryan (2013)).

If, as a result of non-separability, an economic incentive crowds out moral motives to behave pro-socially, it may therefore have less of an effect than expected under separability. It may even have the opposite effect, if the incentive more than crowds out the moral motives completely.

To allow for non-separability, we follow Bowles and Polania Reyes (2012) and postulate that $\mu_i = \mu_i^0(1 + \mathbf{1}\{t > 0\}\mu_i^t)$, where the indicator $\mathbf{1}\{t > 0\} = 1$ if $t > 0$ (a tax is implemented). The shift parameter μ_i^t measures the effect of the tax on the moral price. Because we are modelling a subject that participates in a public bad game, we are going to assume that. We do not distinguish between the categorical effect of the tax (due to the presence of the tax, whatever the value) and the marginal effect (due to the level of the tax) in the specification of separability above, as Bowles and Polania - Reyes did. The reason is that in our experiments we treat subjects with only one level for the tax, and therefore, we are not able to disentangle the categorical effect from the marginal effect. Our μ_i^t captures both effects.

Note that with an emissions tax and non-separability equation (5) becomes $g_i'(e_i) - \gamma - t - \mu_i^0(1 + \mu_i^t) = 0$. The Pigouvian tax in (6) is still optimal, in the sense that it internalizes the externality in the margin. Nevertheless, given that under non-separability subjects have a lower moral price than under separability ($\mu_i^0(1 + \mu_i^t) < \mu_i^0$), the level of emissions with which subjects respond with the tax under non-separability is higher than the level of emissions with the tax under separability, $e_i^{tm}(\gamma, \mu_i^0, t)$.

Our regulator in the experiments may be "sophisticated" (Bowles, 2016), in the sense that it recognizes non-separability. Nevertheless, it is politically constrained. Consequently, it sets a tax $t < t^m = \gamma(n - 1)$.

3.3.3 A tax and a nudge

We now consider the possibility that a regulatory agency uses both a tax and a nudge jointly. Given the lack of theory and the lack of conclusive empirical evidence, our model is silent with respect to the possible complementarity or substitutability of these two instruments. Following the previous discussion, the individual utility in the case in which it faces a nudge and a tax at the same time is given by:

$$U_i(e_i) = g(e_i) - \gamma E - te_i + \mu_i^0(1 + \mu_i^z + \mu_i^t)[\varphi_i m_i(z) + (1 - \varphi_i)s_i(\mathrm{z}) - e_i\,]$$

An individual's choice of emission in this case, $e^{tz}((\gamma, \mu_i^z, \mu_i^t, t)$ satisfies

$$g'(e_i) - \gamma - t - \mu_i^0(1 + \mu_i^z + \mu_i^t) = 0\,, \tag{7}$$

3.4 Hypotheses

In this the final subsection of our theoretical framework, we present the hypotheses that we evaluate with our laboratory experiment. We follow the enunciation of each of the hypotheses with the corresponding proof.

Hypothesis 1 (Morality): *Subjects behave consistently with having "moral preferences".*

Null hypothesis (H_0): The baseline (no treatment) average level of emission is equal to the amoral individual profit maximizing level of emissions ($\boldsymbol{H_0}$: $\bar{e}_{baseline} = e^u$).

Alternative hypothesis (H_1): The baseline (no treatment) average level of emissions is lower than the amoral individual profit maximizing level of emissions. ($\boldsymbol{H_1}$: $\bar{e}_{baseline} < e^u$)

Proof: Recall that equation (2) implicitly defines the optimal choice of emissions of an unregulated moral subject as a function of the marginal externality γ and the moral price μ_i, $e_i^{um}(\gamma, \mu_i)$. Given our assumption that $g''(e_i) < 0$, $e_i^{um}(\gamma, \mu_i)$ is decreasing in μ_i (a simple exercise of comparative statics with Equation (2) shows that $\frac{de_i^{um}}{d\mu_i} = \frac{1}{g''(e_i)} < 0$). Recall also that when setting $\mu_i = 0$, equation (2) characterizes the utility maximizing choice of emission by an amoral polluter, which we called $e_i^u(\gamma)$. It is then easy to see that $e^{um} < e^u$; an unregulated moral subject emits less than an unregulated amoral subject. Therefore, if we reject H_0 in favor of H_1, we can conclude that subjects behave consistently with having "moral preferences". QED.

Hypothesis 2 (Nudgeability): *A nudge informing subjects of the social optimum level of emissions reduces the average level of emissions with respect to the baseline level, consistently with moral preferences of some subjects being affected by the nudge ($\mu_z > 0$).*

Null hypothesis (H_0): The average level of emissions in the nudge treatment is the same as or higher than the emissions in the baseline treatment. ($\boldsymbol{H_0}$: $\bar{e}_{nudge} \geq \bar{e}_{baseline}$)

Alternative hypothesis (H_1): The average level of emissions in the nudge treatment is lower than the emissions in the baseline treatment. ($\boldsymbol{H_1}$: $\bar{e}_{nudge} < \bar{e}_{baseline}$)

Proof: Equation (4) implicitly defines the utility-maximizing choice of emissions by a nudgeable ($\mu_i^z > 0$) moral subject facing a nudge, $e_i^{zm}(\gamma, \mu_i^0, \mu_i^z)$. Given our assumption that $g''(e_i) < 0$, $e_i^{zm}(\gamma, \mu_i^0, \mu_i^z)$ is decreasing in μ_i^z (a simple exercise of comparative

statics using Equation (4) shows that $\frac{de_i^{zm}}{d\mu_i^z} = \frac{\mu_i^0}{g''(e_i)} < 0$). This implies that nudgeable subjects emit less when nudged (in which case its moral price become $\mu_i^0(1+\mu_i^z)$, $\mu_i^z > 0$) than when they are not (in which case its moral price equals μ_i^0). QED.

Hypothesis 3 (Non-separability): *A tax does not reduce the average level of emissions as much as it would under separability ($\mu_t < 0$).*

To test this hypothesis, we conduct a test in which the null hypothesis is that the average level of emissions in the tax treatment is the same as or higher than the expected level of emissions for a subject exhibiting separability ($\boldsymbol{H_0}$: $\bar{e}_{tax} \leq e_{tax}^{separability}$). The alternative hypothesis (H_1) is that the average level of emissions in the tax treatment is higher than the average level in the hypothetical scenario in which morality is not affected by the tax (separability). ($\boldsymbol{H_1}$: $e_{tax}^{separability} < \bar{e}_{tax}$)).

Proof: as noted in section 3.3.2.1, with an emissions tax and non-separability, equation (5) becomes $g_i'(e_i) - \gamma - t - \mu_i^0(1+\mu_i^t) = 0$. Assuming decreasing marginal benefits of emissions ($g_i''(e_i) < 0$) as we do, it is easy to see that the level of emissions solving this equation is increasing in $\mu_i^0(1+\mu_i^t)$ (doing comparative statics, $\frac{de_i^{tns}}{d\left(\mu_i^0(1+\mu_i^t)\right)} = \frac{1}{g''(e_i)} < 0$, where e_i^{tns} is the solution to $g_i'(e_i) - \gamma - t - \mu_i^0(1+\mu_i^t) = 0$. Given that we assume that $\mu_i^t \leq 0$, $\mu_i^0(1+\mu_i^t) \leq \mu_i^0$. Therefore, the level of emissions with which citizens respond to the tax when there is no separability is higher than that when there is. QED

Hypothesis 4 (Complementarity): *A tax set at a level that is lower than that inducing the social optimal level of emissions implemented together with a nudge informing subjects what the social optimum level of emissions is, reduces the average level of emissions with respect to that of the tax or nudge levels ($\mu_z + \mu_t > 0$).*

Null hypothesis (H_0): The average level of emissions in the treatment in which a tax and the nudge are implemented together is the same as or higher than the minimum between the average level of emissions in the tax treatment and that in the nudge treatment ($\boldsymbol{H_0}$: $\bar{e}_{tax+nudge} \geq \min(\bar{e}_{tax}, \bar{e}_{nudge})$).

Alternative hypothesis (H_1): The average level of emissions in the treatment in which the tax and the nudge are implemented together is lower than the minimum between the average level of emissions in the tax treatment and that of the nudge treatment ($\boldsymbol{H_1}$: $\bar{e}_{tax+nudge} < \min(\bar{e}_{tax}, \bar{e}_{nudge})$).

Proof: Comparing equation (7) with equation (4), and assuming $g_i''(e_i) < 0$, as we do, it is easy to see that $e^{(t+z)}(\gamma, n, \mu_i^0, \mu_i^t, \mu_i^z, t) < \min\left(e_i^t(\gamma, n, \mu_i^0, \mu_i^t, t), e_i^z(\gamma, n, \mu_i^0, \mu_i^z, t)\right)$. QED

As we stated above, we are going to test these four hypotheses for the case of a local public bad, affecting every one of the decision makers. Although we formally test the four hypotheses enunciated, our hypothesis of interest is Hypothesis 4.

4 Experimental Design and Procedures

In this section, we present the experimental design, treatments, and the procedures we used to implement our experiments.

We conducted a series of lab experiments with university students in Montevideo, Uruguay. We framed the experiment as a neutral production decision of an unspecified good q, the production of which generates economic benefits for its producer. Every subject had a production capacity of up to 10 units (whole numbers). The schedule of marginal benefits (benefit per unit produced) is presented in Table 1 and is the same for each producer and throughout the experiments. Each individual decides how many units of the unspecified good to produce (from 1 to 10). Starting from a baseline situation without regulation, we study the effectiveness of three policy interventions: a uniform unit tax on production, an informative message (nudge), and a combination of the tax and the nudge.

Table 1. Marginal benefits per unit of production (Ur $)

Unit of production	Marginal benefits
1	$30
2	$22
3	$18
4	$14
5	$11
6	$9
7	$7
8	$6
9	$5
10	$4

Apart from generating economic benefits to its producer, each unit of production in these experiments generates a public bad, affecting the source (the producer) and the rest of the producers in its group. Each group consist of five subjects. We set these damages, the value of the public bad, as a linear function of the aggregate level of production, γQ, where $\gamma > 0$ is a constant parameter capturing the marginal (and average) value of the damage that each unit of production generates to each of the 5 members of the group, and $Q = \sum_{i=1}^{5} q_i$ is the group level of production (q_i is the production level of individual *i*). Consequently, the level of economic benefits that a producer *i* obtains from producing q_i units of this good, are given by the profits obtained when producing q_i units of the good (according to the marginal benefits schedule of Table 1), minus the value of

damages caused by the aggregate level of production in the group, γQ. In our experiments, we set $\gamma = \$U\ 2$ (two Uruguayan pesos).

4.1 Treatments and theoretical benchmarks

We implemented the following treatments:

Baseline: In this treatment, subjects decide freely and uncoordinatedly the number of units that each one wants to produce in each round. With the chosen parameterization, producing 10 units is a dominant strategy for those interested in maximizing profits (amoral subjects). Of course, given the public bad, if all end up producing 10 units, the individual profit is $26. Instead, if the 5 subjects in the group produce 5 units each, every subject earns $45, the maximum possible. Therefore, while 10 units is the Nash equilibrium, 5 units is the social optimum.

Low tax: The second treatment is a uniform tax per unit of production. We set the level of the tax to UY$ 5 per unit. At this level, an amoral profit maximizer individual faced with this tax would choose to produce 6 or 7 units (the marginal benefit of the seventh unit of production, net of the self-imposed $2 from the public bad, is $5). This level of production is higher than the level of production that maximizes welfare of the group (5 units per individual). The reason is that the UY$ 5 tax is lower than the UY$ 8 tax that is needed for amoral producers to fully internalize the externality. (With a UY$ 8-tax, an amoral subject would produce 5 units (the marginal benefit of the sixth unit of production, net of the self-imposed $2 from the public bad, is UY$ 7 (UY$ 9 – UY$ 2)). Our choice of a tax of UY$ 5, lower than the optimal tax, is consistent with our motivation to study the complementarity of nudges and taxes on negative externalities, when taxes are low due to political-economy reasons.

Nudge: Our nudge consists of a message informing the level of individual production that maximizes group economic benefits to the subjects (5 units). The message that appeared on the decision screen was the following:

"The individual production level that maximizes the group´s profits is 5 units.

To choose an individual production level higher than 5 means that the aggregated profits of the group would be lower than when choosing a production level of 5."

Instructions informed subjects that "(t)he income per unit produced is the same for all participants". Therefore, our message cannot provide information to rational, attentive, and capable subjects. However, a nudge that provides information to inattentive or cognitively limited subjects may not affect their behavior either, if this information only helps these subjects to form their belief of the social norm (s_i), their private moral threshold (m_i). Therefore, in order to alter behavior, our nudge must also affect the moral

price. Given that the individual behavior of subjects was not revealed to the other players in the group, or even the experimenter, shame is not a mechanism at play. If our message alters behaviors, it must affect feelings such as guilt (pride).

Of course, we are unable to provide a theoretical benchmark of the level of individual production with which the subjects will respond when facing such a nudge, since we do not observe the individuals' moral term of their utility function.

Low tax + nudge: Lastly, we include a treatment in which we implement the low tax and the nudge jointly. In this case, the message reads

"The individual production level that maximizes the group's profits plus the tax revenues is 5 units.

To choose an individual production level higher than 5 means that the aggregated profits of the group would be lower than when choosing a production level of 5."

The message is slightly different and adds a reference to the tax revenues. This modification is necessary, given that a tax decreases private production profits. A full rebate of the tax revenues among the five subjects, according to some rule, would have made the modification of the message unnecessary. Nevertheless, actual rebates are not full rebates. At least, revenues have to finance the implementation of the pollution control program, including tax collection and administration cost. In addition, revenues frequently finance environmental education campaigns, restoration of habitats, defense measures (adaptation in the case of climate change) and/or technology adoption. Implementing any other rebate different than a full one in our experiments would have needed a similar message or a message with a reference to profits after taxes.

4.2 Procedures

We conducted computer-based experiments in the Experimental Economics Laboratory of the University of Montevideo (UM). We recruited the participants via email invitations sent to university students in Montevideo. Invitations to programmed sessions to those registered for the experiments were administered through ORSEE.

The day of the session, the subjects showing up were received by the experimenter and an assistant. Participants were randomly assigned to groups of five. A maximum of six groups of five participated in a given session. Each session began with the instructions of the game (A full transcript of the whole set of instructions are included in Appendix A). Instructions were played from a previously recorded audio, accompanied by a Power Point presentation highlighting the main points and illustrating the tasks and screens (see Appendix B). After playing the instructions, the experimenter answered the remaining questions. After these questions, the subjects had two practice rounds.

Each session started with the baseline treatment, followed by a second treatment consisting in a policy intervention to control the externality. The intervention was one of the three treatments discussed above: the "low" tax, the nudge (message), or both. The total number of rounds per session was 10, equally divided between the two treatments (baseline + intervention).

After the 10 rounds, the subjects answered a questionnaire. Questions seek to gather information about the participants' sociodemographic characteristics, about their pro-environmental attitudes, religious beliefs, political orientation, and beliefs about other people's attitudes and beliefs. In addition, questions about motives behind their decisions in the experiment were added at the end of the questionnaire. The complete version of the questionnaire is in Appendix A, at the end of the instructions.

An important feature in the procedures of our experiments was the payment procedure. This preserved the confidentiality of the participants' decisions. It did this in the following way. When the activity finished, the experimenter left the room, and the assistant, who stayed outside the room during the entire session, entered the room and extracted the information on the participants´ final earnings from the server, but not the information about how they played. With the earnings information, the assistant prepared the payments and a receipt for each participant. Once this task was completed, the assistant proceeded with the payments. Each participant received their earnings in private. The instructions explained this payment procedure and underscored to participants that with this procedure, nobody (neither the experimenter, nor the assistant) could know what decisions they made in the experiment. In addition to the earnings from the exercises, participants were paid UY$ 150 for showing up on time for the experiment.

4.3 Participant´s characteristics

In total, we conducted 18 experimental sessions, recruiting 200 subjects, in 40 groups of 5 participants each. (See Table 2).

Table 2. Subjects per treatment

Treatment	Groups	Subjects
Baseline	40	200
Low Tax	14	70
Nudge	13	65
Low Tax + Nudge	13	65

Most of the recruited subjects were between 18 and 21 years old. Forty-six percent (46%) were females. Almost all the subjects were undergraduate students from the University of the Republic (74%), or from the University of Montevideo (21%). Most of the subjects majored in economics (57%). Concerning university majors, 77% of the participants were pursuing a degree within the STEM/ECON fields. Subjects reported

household income levels that are well distributed across the different income ranges, with nearly equal numbers in each range. Appendix D presents more detailed descriptive statistics on the characteristics of the participants, based on their responses to questionnaire.

5 Results

In this section, we present the results of our work. First, we present the descriptive statistics of the outcome variable in each treatment, as well as their theoretical benchmarks. We then present the results of the parametric and non-parametric tests of our hypotheses. Finally, we present our regression analysis, as additional tests of our hypotheses.

5.1 Descriptive statistics

Table 3 presents the descriptive statistics of our outcome variable, the individual production level per round, by treatment. Table 3 also presents the corresponding theoretical benchmarks (expected values) of this variable, depending on the assumed morality of the producer.

Table 3. Descriptive statistics and theoretical benchmarks for the per- round individual level of production (*q*), by treatment

	Baseline	Nudge	Low Tax	Low Tax + Message
Statistics				
Mean	7.45	7.13	6.31	6.43
Median	8.00	7.00	6.00	6.00
Std. Deviation	2.32	2.19	2.11	2.04
Observations	1000	325	350	325
Theoretical benchmarks				
Amoral producer	10	10	6 - 7	6 - 7
Moral producer				
with separability	<10	< 10	<6	<6
without separability	<10	<10	? (*)	? (*)
Group profit maximizer	5	5	5	5

(*): the subject with non-separability may respond to a tax with a level of emissions that could be lower or higher than 6-7, depending on the degree in which the tax affects social preferences (μ^t) and, in the case of the T+M treatment, the synergy between the two instruments.

5.2 Parametric and non-parametric tests

Hypothesis 1 (Morality): To evaluate our first hypothesis, concerning the morality of subjects, we test whether the average level of production in the baseline treatment is equal to the amoral individual profit maximizing level, against the alternative hypothesis that is lower. According to their marginal benefits of production presented in Table 1, an amoral, profit – maximizing producer should produce 10 units of the good. Formally stated then, our null hypothesis is $\boldsymbol{H_0}$: $\bar{q}_{baseline} = 10$ and our alternative hypothesis is $\boldsymbol{H_1}$: $\bar{q}_{baseline} < 10$. As seen in Table 3, we have : $\bar{q}_{baseline} = 7.45$. The result of a t-test is presented in the first line of Table 4. According to this test, we can reject the null in favor of the alternative (p-value: 0.0000). Therefore, the behavior of the subjects is consistent with them having moral preferences in the form of equation (1).

To illustrate, we estimate the magnitude of the baseline moral price of the average subject in our experiments, $\bar{\mu}^0$. To obtain it, we use Equation (2) ($g'(e_i) - \gamma - \mu_i^0 = 0$) to solve for the value of μ_i^0 that is consistent with the average level of production in the baseline treatment. In order to do it, we fit a continuous function to the discrete values of the subjects' marginal benefits of production, and we substitute q for $q_{baseline} = 7.45$. The fitted function is $g'(q) = 0.3295q^2 - 6.3159q + 34.65$. Doing this, we may write $0.3295q^2 - 6.3159q + 34.65 = \gamma + \mu_i^0 = 2 + \mu_i^0$, which solves for $\bar{\mu}^0 \cong 3.9$. This is the average moral price estimated for the 200 subjects that played 5 rounds of the baseline treatment. Knowing that it produces a public bad (a negative externality of \$2 to each of the other 4 member of its group), due to their social preferences in the form of a moral price, each additional unit of production costs an average of \$3.9 to its producer.

Table 4: Results of parametric and non-parametric tests of the hypotheses

Hypothesis		Test	H_0	H_1	Statistic	p-value
1	Morality	t-test	$\bar{q}_{baseline} = 10$	$\bar{q}_{baseline} < 10$	-34.69	0.0000
2	Nudgeability	t-test	$\bar{q}_{nudge} \geq \bar{q}_{baseline}$	$\bar{q}_{nudge} < \bar{q}_{baseline}$	2.12	0.0171
		Rank-sum	$\bar{q}_{baseline} - \bar{q}_{nudge} = 0$	$\bar{q}_{baseline} - \bar{q}_{nudge} \neq 0$	2.78	0.0055
		Median	$q_{baseline}^{median} = q_{nudge}^{median}$	$q_{baseline}^{median} \neq q_{nudge}^{median}$	9.69	0.002
3	Tax	t-test	$\bar{q}_{tax} \geq \bar{q}_{baseline}$	$\bar{q}_{tax} < \bar{q}_{baseline}$	8.05	0.0000
		Rank-sum	$\bar{q}_{baseline} - \bar{q}_{tax} = 0$	$\bar{q}_{baseline} - \bar{q}_{tax} \neq 0$	8.29	0.000
		Median	$q_{baseline}^{median} = q_{tax}^{median}$	$q_{baseline}^{median} \neq q_{tax}^{median}$	66.61	0.000
	Separability	t-test	$\bar{q}_{tax} \leq q_{tax}^{separability} = 5$	$\bar{q}_{tax} > q_{tax}^{separability} = 5$	11.55	0.0000
4	Complementarity	t-test	$\bar{q}_{tax+nudge} \geq \bar{q}_{tax}$	$\bar{q}_{tax+nudge} < \bar{q}_{tax}$	0.76	0.7765
		Rank-sum	$\bar{q}_{tax+nudge} - \bar{q}_{tax} = 0$	$\bar{q}_{tax+nudge} - \bar{q}_{tax} \neq 0$	-0.070	0.9445
		median	$q_{tax+nudge}^{median} = q_{tax}^{median}$	$q_{tax+nudge}^{median} \neq q_{tax}^{median}$	1.9507	0.163

Hypothesis 2 (Nudgeability): To evaluate our second hypothesis, we test whether, consistently with moral preferences being affected by the nudge ($\mu_i^z > 0$), a nudge informing subjects what the social optimum level of emissions is reduces the subject´s chosen level of emissions with respect to the baseline level. To test Hypothesis 2, we performed three different types of tests: two non-parametric tests (the Wilcoxon rank-sum test (also known as the Mann-Whitney U test) and the median test) and a t-test. The results of these tests are presented in lines 2 to 4 of Table 4.

In the t-test (line 2), we test whether the average level of production in the nudge treatment ($\bar{q}_{nudge} = 7.13$) is the same as or higher than the average level of emissions in the baseline treatment ($\bar{q}_{baseline} = 7.45$), against the alternative hypothesis that is lower. More formally, H_0: $\bar{q}_{nudge} \geq \bar{q}_{baseline}$ and H_1: $\bar{q}_{nudge} < \bar{q}_{baseline}$. According to the t-test, we should reject H_0 in favor of H_1 (p-value: 0.0171). In other words, the message is effective in reducing the average level of production (from 7.45 units to 7.13 units). The observed difference in the levels of average production (0.32 units), a 4.3% decrease, has a standard deviation of 0.14 (95% confidence interval for the difference in the average levels of production: [0.023, 0.598]).

The results of the non-parametric tests are consistent with the result of the t-test. According to the Wilcoxon rank-sum test (line 3, Table 4), we should reject the null that the two samples of production levels come from the same distribution (p-value= 0.0055). Likewise, the value of the Pearson chi-squared statistic of the median test (9.69) indicates that we should reject the hypothesis that the median production level in the baseline treatment is equal to the median production level in the message treatment (p-value = 0.002).

In sum, we conclude that a nudge informing subjects of the social optimum level of production reduces the average level of production with respect to the baseline level. This is consistent with at least some of the subjects exhibiting moral preferences and these being affected by the nudge ($\mu_i^z > 0$). This result is consistent with Antinyan et al. (2020), who find that, similarly to us, informing participants of the joint welfare maximizing consumption bundle changes their behavior by increasing the moral price and psychological cost of generating the externality.

To illustrate, similarly to what we did with μ_i^0, we can estimate the value of $\mu_i^z > 0$ that is consistent with the average level of production of subjects in the nudge treatment. Recall that the optimal choice of production for a subject with a Levitt-List utility function is characterized by the following condition (Equation (4)): $g_i'(e_i) - \gamma - \mu_i^0(1 + \mu_i^z) = 0$. Substituting $g_i'(q_i)$ by the fitted continuous function used above, $\gamma = 2$ and μ_i^0 for $\bar{\mu}^0 = 3.9$, we obtain $\bar{\mu}^z = 0.121$. In other words, the nudge increases the subject´s average moral price by $\approx 3\%$, with respect to its baseline level $\bar{\mu}_i^0$. This is the mechanism by which as a response to the message, subjects decrease production by 4.3%, on average.

Hypothesis 3 (Separability): To test our third hypothesis, concerning the separability of economic incentives and moral preferences, we first test whether the low tax is

effective in reducing the level of production, with respect to the baseline. That is, we test whether the average level of production in the tax treatment ($\bar{q}_{tax} = 6.31$) is the same as or higher than that of the baseline treatment ($\bar{q}_{baseline} = 7.45$), against the alternative hypothesis that is lower. More formally, H_0: $\bar{q}_{tax} \geq \bar{q}_{baseline}$ and H_1: $\bar{q}_{tax} < \bar{q}_{baseline}$.

According to the t-test presented in line five of Table 4, we should reject H_0 in favor of H_1 (p-value: 0.0000). In other words, the tax effectively reduces the average level of production of subjects (from 7.45 units to 6.31 units). The observed difference in the levels of average production (1.14 units), a 15.2% decrease, has a standard deviation of 0.14 units (95% confidence interval for the difference in the levels of average: [0.86, 1.414]).

This result is confirmed by the non-parametric tests. According to the rank-sum test (line 6 of Table 4), we must reject the null hypothesis that the sample observation of the levels of production in the baseline treatment come from the same distribution than those coming from the tax treatment. According to the median test, we must reject also the null hypothesis that the median level of production in the baseline treatment (8 units) is equal to the median level in the tax treatment (6 units).

Having concluded that the average decrease in production levels from 7.45 units to 6.31 units caused by the $5 tax is statistically different from zero, we now test whether the decrease is consistent with the separability assumption. More formally, we test whether $\overline{q}_{tax} \leq q_{tax}^{separability}$, against the alternative that is higher, where $q_{separability}$ is the predicted level of production under the separability assumption, the solution to equation (5). To calculate $q_{tax}^{separability}$ using equation (5) we need the average baseline moral price. Using the value estimated for μ_i^0 above, $\gamma = 2$, $t = 5$, and the fitted $g'(e)$ function, the solution to equation (5) is $q_{tax}^{separability} = 5$.

Now we are able to test the separability hypothesis $\boldsymbol{H_0}$: $\bar{q}_{tax} \leq q_{tax}^{separability}$ against the non-separability alternative hypothesis $\boldsymbol{H_1}$: $q_{tax}^{separability} < \bar{q}_{tax}$. The result of the t-test for this hypothesis (in Table 4), we should reject the null hypothesis in favor of the alterative that the subjects reacted to the tax with a production level that is higher than the one predicted for subjects with separability. The difference ($6.31 - 5 = 1.31$ units) is statistically significant (p-value = 0.0000).

We can therefore conclude that the tax does reduce the average level of emissions, but not as much as it would under separability. This result is consistent with the subjects having non-separability between their degree of social preferences and the use of a tax to control the externalities.

Hypothesis 4 (Complementarity): Finally, having tested the morality, nudgeability and separability hypotheses, we now have a tested model upon which to stand in order to test the degree of complementarity (or substitutability) of a low tax with an informative message as policy instruments to reduce negative externalities. We test the null hypothesis of no complementarity (negative synergy) between the two instruments, against the alternative of some degree of complementarity (positive synergy). We call "no

complementarity" or "negative synergy" the situation in which the effect of the two instruments implemented jointly is lower than the maximum of the two effects of the instruments standing alone. More formally, we test $\boldsymbol{H_0}$: $\bar{q}_{tax+nudge} \geq \min(\bar{q}_{tax}, \bar{q}_{nudge})$, against $\boldsymbol{H_1}$: $\bar{q}_{tax+nudge} < \min(\bar{q}_{tax}, \bar{q}_{nudge})$. The reason to call H_1 "strong negative synergy" is H_1 is a specific case of negative synergy, which includes the cases in which the instruments have some degree of complementarity. Recalling that in our experiments $\min(\bar{q}_{tax}, \bar{q}_{nudge}) = \bar{q}_{tax}$, some degree of complementarity involves $\bar{q}_{tax} < \bar{q}_{tax+nudge} \leq \bar{q}_{tax} - (\bar{q}_{baseline} - \bar{q}_{nudge})$, i.e.: the implementation of the two instruments jointly reduces production more than what the tax does, but not more than what the sum of the effects of the two instruments standing alone would.

As it can be seen in the last three lines of Table 4, the results of the t-test and the non-parametric test indicate that we cannot reject that the sample mean level of production of the "tax + message" treatment (6.43) is higher than or equal to that of the tax alone ($\min(\bar{q}_{tax}, \bar{q}_{nudge}) = \bar{q}_{tax} = 6.31$ units). Therefore, our experiments show results consistent with the hypothesis that there is no complementarity or some degree of negative synergy between the two instruments. More specifically, we cannot reject that the average effect of the two instruments jointly implemented is equal to the effect of the low tax implemented alone. In other words, we cannot reject that the message adds no effect to that of the tax.

To sum up, the results of the parametric and non-parametric tests indicate that (a) the average and median subjects in our sample exhibit baseline social preferences in the form of a moral price, consistent with our theoretical model, (b) a message suggesting a behavior to internalize a negative externality with an implicit moral appeal or suasion is effective in reducing the level of the negative externality of the average and median subjects, consistent with the hypothesis derived from our model that this type of message increases a subject´s moral price, (c) a tax that is not high enough to induce the proper level of the externality is effective in reducing the level of the negative externality of the average and median subjects, consistent with the hypothesis derived from our model that a tax decreases a subject´s moral price, and (d) the message adds no effect to that of the tax when both are implemented jointly. Having found evidence of non-separability (that is, the tax crowding out morality), this last result is also consistent with the hypothesis that it is the crowding out of social preferences of the tax that crowds out the effect of the nudge.

5.3 Regressions

To complement the results of the parametric and nonparametric tests previously presented, we carried out an econometric analysis. Table 5 presents the results of two random-effects linear panel regression estimation of this analysis in which our outcome variable is the level of production of subject *i* in round *t* (q_{it}). In Model 1, the covariates are a set of treatment indicator variables. In Model 2, the covariates include the former

plus a vector of controls for socio-economic characteristics. More formally, the specification for the regression in model 2 is $q_{i,t} = \beta_0 + \beta_1 Message_{i,t} + \beta_2 Tax_{i,t} + \beta_3 (Message + Tax)_{i,t} + X_i + \varepsilon_{i,t}$, where X_i is the vector of socio-economic characteristics. These include the university major of the subjects, their declared household income and the university they assist.

The results in Table 5 show that the three interventions are effective in decreasing the average individual production level, compared to the Baseline. They also confirm the results obtained by our parametric and non-parametric tests, regarding our hypotheses 1 and 2. If we look at column 2 (Model 1), we can see that $\bar{q}_{baseline} = 7.45 < 10$ (p-value = 0.0000). This result is consistent with subjects having moral preferences of the Levitt and List form, our ($\mu^0 > 0$), our Hypothesis 1. If we turn to Hypothesis 2 (Nudgeability), the $\hat{\beta}_{Message} = -0.47$ is statistically different form zero, which indicates that t $\bar{q}_{nudge} < \bar{q}_{baseline}$. This is consistent with subjects being nudged by the message through a change in their moral preferences ($\mu^z > 0$).

The econometric analysis does not allow us to test for the non-separability hypothesis directly. Nevertheless, we observe in Table 5 that the estimate of the constant is identical to the $\bar{q}_{baseline}$ presented in Table 3. Therefore, we could use our estimate $\bar{\mu}_i^0 = 3.9$ and obtain, as before, $q_{tax}^{separability} = 5$. At the same time, according to our regression results, $\bar{q}_{tax} = \widehat{Constant} + \hat{\beta}_{tax} = 7.45 - 1.02 = 6.43$. Since we reject that this value is equal to 5 in favor of the alternative that is higher (p value = 0.0000) the econometric analysis confirms the evidence favoring the non-separability hypothesis that we obtained with the tests.

The estimations we report in Table 5 also confirm the results obtained for our main hypothesis, Hypothesis 4 (complementarity). To see it, observe that while both, the tax and the nudge prove to be effective, the impact of the tax (~1 unit) doubles that of the message (~0.5 units). As observed in the lower part of Table 5, we reject the hypothesis that $\hat{\beta}_{Tax} = \hat{\beta}_{Message}$ are equal in favor of the alternative that $|\hat{\beta}_{Tax}| > |\hat{\beta}_{Message}|$ (p-value = 0.002). We also reject the hypothesis that $|\hat{\beta}_{Tax+Message}| = |\hat{\beta}_{Message}|$ against the alternative that $|\hat{\beta}_{Tax+Message}| > |\hat{\beta}_{Message}|$ (p-value is 0.004). Nevertheless, we cannot reject the hypothesis that $\hat{\beta}_{Tax+Message} = \hat{\beta}_{Tax}$, (the associated p-value is 0.84). We, therefore, have that $|\hat{\beta}_{Message}| < |\hat{\beta}_{Tax}| \approx |\hat{\beta}_{Tax+Message}|$. This is the same result obtained with the tests. reported in Table 4. Namely, the impact of jointly implementing the two instruments is not different to that of implementing the tax alone. The message does not add any effect to that of the tax. There exists a negative synergy between the tax and the message in which the tax crowds out the effect of the nudge. When introduced together, the effect of the tax on social preferences crowds out the effect of the nudge.

The results presented in Table 5, column (2), show that the absolute and relative effects of the different treatments are robust to adding controls based on observable characteristics of the subjects in our sample. Moreover, the coefficients associated with

these different characteristics we controlled for are not statistically significant at the 5% significance level (not shown). Thus, we can conclude that despite having an unbalanced sample, this does not affect the conclusions drawn from our previously presented results.

Table 5. Linear Random Effect regression results

Dependent variable: level of production	Model 1	Model 2
Message (M)	-0.47***	-0.50***
	(0.13)	(0.13)
Tax (T)	-1.02***	-1.00***
	(0.13)	(0.13)
Message and Tax (M+T)	-0.98***	-0.97***
	(0.13)	(0.13)
Constant	7.45***	7.92***
	(0.11)	(0.33)
Controls	No	Yes
Chi-squared	126.51	150.16
N	2000.00	2000.00
Hypothesis Tests		
M vs T	0.56***	0.50***
	(.177)	(.178)
T vs (M+T)	-0.04	-0.02
	(.177)	(.178)
M vs (M+T)	0.52***	0.47***
	(.181)	(.18)

Notes: This table presents estimates from random-effect GLS regressions, estimated on a panel data structure grouped in two levels: a concatenation of session and subject as the group identifier, and round as the time identifier. In each column, the dependent variable is the production level chosen by subject i of group g, in round t. There is one independent variable for each treatment, recalling that a subject only participated in one of them. Each treatment is represented by a dummy variable, which is 1 if subject i has faced that treatment, and 0 in other case. The constant on each estimation reflects the average production level in baseline. Column 1 reflects the regression results without controlling for participants characteristics and Column 2 the results but controlling for several covariates where we identified imbalances between treatments, which are income, field of studies and University. The last 3 rows exhibit the t-tests performed for the difference of each treatment effect. Significance levels are indicated as * 0.10 ** 0.05 *** 0.01

6 Discussion

Our experimental evidence is consistent with the hypothesis that an informative nudge changes the behavior of our average subject by altering her baseline moral price. It is also consistent with the non-separability hypothesis, based on a tax triggering the same mechanism (in the opposite direction). Finally, the negative synergy between the message and the tax that we find could also be modeled as a similar mechanism, involving a negative interaction between the former effects. The objective of this section is to unravel these mechanisms. To do it, we use the subjects´ responses to the questionnaire to classify

individuals into different groups, based on some of their characteristics, and we analyze how individual reactions to the instruments are associated with these characteristics.

6.1 STEM/ECON vs other university majors

A first characteristic that we use to classify subjects is the participants' university major. We divide participants into two categories. The first one, which we call STEM/ECON, include all students majoring in the following areas: Science, Technology, Engineering, Mathematics, Economics (including related fields such as accountancy, business administration, finance), and Data Science. The second category includes students majoring in Communications, Architecture, Psychology or Social Work. The rationale behind this classification is the thesis that individuals in the first group, for different reasons, might be more inclined to follow rational thinking when making decisions, which in this setting would conduce to the maximization of the private benefits, consistent with what a "homo-economicus" would do. According to our model, this would translate into having different moral prices than those of subjects in the other group. In particular, STEM/ECON students could be less affected by the nudge, compared to individuals from majoring in other areas. After classifying subjects into these two groups, we explore the potential differential effects of treatments in these two groups by running random effects regressions. Results from this exercise are presented on Table 6. Column 1 shows the results for STEM/ECON group of 156 subjects, Column 2 for the non-STEM/ECON group of 56 subjects and Column 3 shows the results for the pooled sample, with the corresponding indicator variables. The dependent variable in all regressions is the level of production of subject *i* in round *t* (q_{it}).

A first observation is that subjects majoring in Non-STEM/ECON areas produced, on average 0.62 units less than student majoring in STEM/ECON areas in the baseline. Thes results suggest that there are indeed differences in their baseline moral preferences and are consistent with the hypothesis that those who are studying outside the STEM/ECON area have "stronger" moral preferences (a higher moral price μ^0). In addition to having more baseline moral preferences, those subjects majoring in Non-STEM/ECON area are more "nudgeable". We can see in column (2) that they reduce their production by 1 unit, on average, when nudged, compared to the 0.3 units that those in the STEM/ECON group reduce. On the other hand, for both types of individuals, the stand-alone tax leads both groups to reduce their average production level by the same amount (1 unit). Noticeably, when both instruments are combined, their reactions differ, but not in the expected manner. For those majoring in Non-STEM/ECON areas, the joint implementation of both instruments has a clear negative synergy. The average effect of the combined instruments (a decrease in 0.45 units) is half the effect of the effect any of the two instruments standing-alone, and statistically significant only at the 10% level. This result is novel. We find that a price does not only may crowd out baseline social preferences, consistently with the non-separability hypothesis ($\mu^t < 0$), but it may crowd out the additional social preferences that a nudge may trigger. When jointly implemented, a price diminishes the nudge's ability to alter these preferences (μ^z). Of course, the negative synergy between

the price and the nudge could be explained the other way around; the nudge may crowd out the effect of the price.

Table 6. Linear Random Effect regressions results depending on university major

Dependent variable: Level of production	(1) STEM/ECON	(2) Non-STEM/ECON	(3) All subjects
Message (M)	-0.29**	-1.00***	-0.29*
	(0.14)	(0.30)	(0.15)
Tax (T)	-1.04***	-1.03***	-1.05***
	(0.13)	(0.40)	(0.13)
Message and Tax (M+T)	-1.22***	-0.45*	-1.22***
	(0.15)	(0.25)	(0.16)
Non-STEM/ECON (NSE)			-0.62**
			(0.27)
NSE*M			-0.76**
			(0.31)
NSE*T			0.07
			(0.39)
NSE*(M+T)			0.78***
			(0.28)
Constant	7.58***	6.96***	7.58***
	(0.13)	(0.24)	(0.13)
Chi-squared	125.59	19.48	146.98
N	1,560	440	2,000
M vs T	0.75***	0.03	
	(.19)	(.488)	
T vs (M+T)	0.18	-0.58	
	(.197)	(.463)	
M vs (M+T)	0.93***	-0.55	
	(.206)	(.377)	

Notes: This table presents estimates from random-effect GLS regressions, estimated on a panel data structure grouped in two levels: a concatenation of session and subject as the group identifier, and round as the time identifier. In each column, the dependent variable is the production level chosen by subject i of group g, in round t. There is one independent variable for each treatment, recalling that a subject only participated in one of them. Each treatment is represented by a dummy variable, which is 1 if subject i has faced that treatment, and 0 in other case. We also include a dummy variable to measure the impact of university major over production and treatment effects. This variable is called non-STEM/ECON, and takes values 1 if subject is studying a major outside the STEM/ECON area, and 0 otherwise. We also include the interaction of this variable with the treatment ones, to check if there are differential effects of each instrument over production depending on majors. The constant on each estimation reflects the average production level in baseline. Column 1 reflects the regression results for the group compound only by the STEM/ECON subjects, Column 2 the same but for the students of other areas and Column 3 has the effects for the pooled case where all subjects are considered. The last 3 rows exhibit the t-tests performed for the difference of each treatment effect. Significance levels are indicated as * 0.10 ** 0.05 *** 0.01

Notably, the synergy between a tax and an informative nudge is totally different for STEM/ECON students. Since we cannot reject the hypothesis that $\hat{\beta}_{Message} + \hat{\beta}_{Tax} = \hat{\beta}_{Message+Tax}$, we conclude that for these students the instruments are perfect complements.

Assuming that the classification in STEM/ECON students and other is representative of classification between "more rational" and "less rational" subjects, we can conclude that the effect jointly implementing a price and a nudge on an externality will depend on the distribution of these traits in the population. While such a policy may harness the power of both instruments in rational individuals (albeit the effect of a nudge for this type of subjects is relatively lower), this complementarity can turn into a negative synergy if the share of less rational subjects is sufficiently large. If not, as in the case of our population, the result can be something in between, as we report in table 5, where, on the average subject, the negative synergy between the two instruments tax completely crowds out the effect of the nudge.

6.2 Pro-sociality

To further unravel the mechanisms behind the interaction between a low price and an informative nudge, we explore the effect of implementing both instruments jointly on different sets on subjects, when subjects in the different sets differ on social preferences. To classify subjects in this manner, we use the answers to the section in the questionnaire designed to assess prosocial behaviors (see questions 18 to 23 in Appendix C). Given that the responses were closed-ended, we assigned a score to each response, ranging from 0 (if they answered 'never') to 4 (if they answered 'always'). After computing the scores for each question, we summed the total points for each subject and divided this final score by the maximum possible total points (24) to obtain a "pro-sociality index". In this index, a score closer to 1 indicates that the subject has 'stronger' social preferences. After calculating the index for each subject, we categorize individuals into three groups, according to this index. Those whose index lies in the lowest 33% of values, were categorized as "Less Prosocial". Those whose index lies between the lowest and the highest 33% of the values, were categorized as "Moderately prosocial". Finally, those whose score lies at the top 33% of the values of the index were classified as "Very Prosocial". Of a total of 200 hundred subjects, 25 classified as "Very Prosocial", 35 as "Less Prosocial" and 140 "Moderately Prosocial". Following this classification, we conducted regression analyses for each group separately, and for the full sample of subjects, with the corresponding indicator variables, to examine the differences in the effects of the treatments between these groups. Table 7 presents the results of linear random effects estimations of these estimations.

Table 7. Linear Random Effect regressions by social preferences

Dependent variable: Level of production	(1) Less Prosocial	(2) Moderately prosocial	(3) Very prosocial	(4) All Subjects
Message (M)	-0.55*	-0.29*	-1.77***	-0.56*
	(0.32)	(0.15)	(0.45)	(0.33)
Tax (T)	-1.05***	-1.05***	-0.89***	-1.05***
	(0.27)	(0.16)	(0.31)	(0.28)
Message and Tax (M+T)	-1.06***	-0.93***	-1.27***	-1.04***
	(0.30)	(0.15)	(0.45)	(0.32)
Moderately prosocial (MP)				-0.27
				(0.30)
Very prosocial (VP)				-0.61
				(0.42)
MP*M				0.27
				(0.36)
VP*M				-1.22**
				(0.54)
MP*T				-0.00
				(0.32)
VP*T				0.18
				(0.41)
MP*(M+T)				0.11
				(0.35)
VP*(M+T)				-0.27
				(0.53)
Constant	7.71***	7.44***	7.10***	7.71***
	(0.25)	(0.14)	(0.31)	(0.27)
Chi-squared	27.80	80.26	29.07	142.78
N	350	1,400	250	2,000
M vs T	0.50	0.76	-0.88	
	(.408)	(.215)	(.531)	
T vs (M+T)	0.01	-0.12	0.39	
	(.397)	(.216)	(.531)	
M vs (M+T)	0.50	0.65	-0.49	
	(.431)	(.209)	(.621)	

Notes: This table presents estimates from random-effect GLS regressions, estimated on a panel data structure grouped in two levels: a concatenation of session and subject as the group identifier, and round as the time identifier. In each column, the dependent variable is the production level chosen by subject i of group g, in round t. There is one independent variable for each treatment, recalling that a subject only participated in one of them. Each treatment is represented by a dummy variable, which is 1 if subject i has faced that treatment, and 0 in other case. We also include a categorical variable, prosocial, to measure the impact of social preferences over production and treatment effects. If subject i has a pro-sociality index (which goes from 0 to 1) between 0 and 0.33, the variable takes value 0 and is tagged as less prosocial; if it has an index between 0.33 and 0.66 it takes value 1 and tagged as moderately prosocial and if the index score of subject *i* is higher than 0.66 it takes value 2 and tagged as very prosocial. We also include the interaction of this variable with the treatment ones, to check if there are differential effects of each instrument over production depending on the preferences. The constant on each estimation reflects the average production level in baseline. Column 1 reflects the regression results for the group compound only by the less prosocial subjects, Column 2 the same for the moderately ones, Column 3 for the most prosocial and Column 4 presents the pooled case. The last 3 rows exhibit the t-tests performed for the difference of each treatment effect. Significance levels are indicated as * 0.10 ** 0.05 *** 0.01

As in table 6, the dependent variable in the regressions results shown in Table 7 is the level of production of subject *i* in round *t* (q_{it}). Column 1 contains the results for the "Less Prosocial" subjects, column 2 for the "Moderately Prosocial" subjects and column 3 for the "Very Prosocial" subjects. Finally, column 4 replicates the analysis for the pooled sample with the group dummies.

A first result is that the value of their baseline level of production (the Constant) is consistent with what one would expect. Less prosocial subjects produce more than moderately prosocial subjects and this, in turn, more than the very prosocial ones. Nevertheless, we cannot reject the hypothesis that these levels are equal.

A second observation is that not all the average subjects of each group react to the informative message in the same way. The nudge is very effective for the "Very prosocial" average subject (decreasing its production level by 1.77 units), but significantly less so, both in units of production and statistically, for the other types of subjects. A result that is robust to this exercise also is that all three types of subjects react very similarly to the tax, decreasing their average level of production by around 1 unit. This result is statistically very significant. With respect to the synergy between both instruments when implemented jointly, results show heterogeneity across groups.[2] For the very prosocial group, the effect of both instruments implemented jointly (-1,27 units) is ~48% of the sum of the effects of the two instruments standing alone. This difference is statistically significant (p-value = 0.047, CI: [-2.7412, -0.0181]), suggesting that the two instruments exhibit negative synergy when implemented jointly. For the less prosocial group, the effect of both instruments implemented jointly is a decrease of 1.06 units of production. This number is ~66% of the sum of the effects of the two instruments standing alone. Nevertheless, this difference is not statistically significant (p-value=0.285, CI: [-1.5427,0.4531]). In other words, we cannot reject the hypothesis that for the less prosocial group, the effect of the two instruments implemented jointly is the same as the sum of the instruments implemented alone. Finally, for the moderately prosocial, the majority of subjects, the effect of implementing both instruments jointly is a decrease of 0.93 units with respect to the baseline. This decrease is approximately ~69% of the sum of the decreases caused by the two instruments separately (1.34 = 0.29+1.05). The difference, 0.41 units, is not statistically significant for this group either (p-value=0.122, CI: [-0.9216,0.1091]). Therefore, we cannot reject the hypothesis that the effect of the two instruments implemented jointly is the same as the sum of the instruments implemented alone for the moderately prosocial. In sum, for all subjects except the more prosocial, we cannot reject that the instruments could show perfect complementarity (zero synergy).

[2] These results do not change if we split the sample into two groups, one with a value of the social preferences index below the median and one above.

If we focus on the differences between groups, our results suggest that a low tax is evenly effective in decreasing production across subjects with different prosocial preferences. (The coefficients of the corresponding interactions are not statistically significant in column 4 of table 7). An informative nudge, on the other hand, is not. It is 1.22 units more effective with very prosocial subjects, than with less prosocial ones (p-value=0.024). On the other, there is no statistically significant difference between the effect of the message on less prosocial subjects and the effect on moderately prosocial ones. Finally, implementing both instruments jointly does not add any effect to that of the tax for all the considered groups, and we cannot reject the hypothesis that the effect is the same for all subjects despite their level of prosociality.

In sum, two results emerge from this exercise. One is that the instruments exhibit a degree of complementarity, but this degree differs between subjects according to their social preferences. For those with lower social preferences, the price and nudge incentive act as perfect complements, while for those with "stronger" social preferences, they show partial complementarity with negative synergy. Second, we cannot reject the hypothesis that the instruments, separately or jointly, affect subjects with different levels of pro social preferences differently; except for the tax, which affects subjects with different social preferences evenly.

6.3 Welfare

To finish, we perform a welfare/profit analysis of the effects of the different interventions. To do it, we calculate the earnings of the average subject in each treatment, including the net profits from production, the externality generated by the other members in the group, the total tax payment, and the moral utility. To calculate the moral utility term, we use our estimations of the different moral prices $\mu_i^0, \mu_i^z \mu_i^t$, and we impute the participants' response to the question regarding the level of production they believed every participant should achieve in each round, as the subjects' personal moral threshold $(\varphi_i m_i + (1-\varphi_i)s_i)$. Additionally, we calculate the profits/utility of a theoretical amoral subject in the tax and no tax treatments, as comparison. Table 8 presents the result of this analysis.

A first observation is that moral subjects, on average, achieve a higher level of utility than what a hypothetical amoral subject would achieve, when comparing each treatment to its corresponding theoretical counterpart. Without a tax, an amoral subject in a group of amoral subjects would obtain \$26 in profits, while the average moral subject in our experiments obtains \$31.9 in the baseline treatment and \$32.9 in the nudge-alone treatment. The reason is that the benefits they gain from reducing their levels of production (and thereby externalities) more than offset their additional moral cost. Although profits decrease due to the tax bill, the same happens in the tax treatment. Individual profits are between \$6 and \$14 in a group of amoral subjects facing a tax and who produce 7 units or 6 units, respectively, while there are \$10.8 with the tax alone and \$10.1 with the tax + nudge in a group of average moral subjects.

A second observation is that both types of subjects achieve lower levels of utility when individuals are exposed to a tax, either standing alone or in combination with a nudge, than under the nudge alone or in the baseline scenario. This is due to the tax bill. Therefore, although the tax is more effective in reducing the externalities, if subjects have a Levitt-list utility function, there is the possibility that they could be better off with a less effective nudge or even with no intervention at all than with a non-revenue-neutral tax, even if this is low.

Table 8. Estimated utility levels of average subject, by type of preferences and treatment

	Type of subject	Treatment	q_i	Moral price	$\varphi_i m_i + (1-\varphi_i)s$	$U_i(q_i)$
Experiments	**Moral**	Baseline	7.45	$\mu^0 = 3.9$	5.6	31.9
		Nudge	7.13	$\mu^0 + \mu^z = 3.9 + 0.12$	5.4	32.9
		Tax	6.31	$\mu^0 + \mu^t = 3.9 - 0.77$	5.6	10.8
		Tax + Nudge	6.43	$\mu^0 + \mu^t + \mu^z = 3.9 - 0.83$	5.7	10.1
Theory	**Amoral**	No tax	10	-	-	26
		Tax	6 - 7			6 - 14

Notes: This table presents the result for the welfare analysis of the effects of the different interventions. To do so, we use the utility function described in section 3, $U_i(q_i) = g(q_i) - \gamma Q - tq_i + \mu_i^0(1 + \mu_i^z + \mu_i^t)[\varphi_i m_i(z) + (1-\varphi_i)s_i(\mathrm{z}) - q_i\ ]$, where: q_i is the average production level observed at each treatment; $g_i(q_i)$ is the profit function derived from the marginal benefits table presented to the subjects and described in section 4 (not the quadratic profit function we use to estimate the parameters in section 5, which is a fitted function and does not show real profit earned by subjects); with $\gamma = 2$ as we mentioned in section 4; $\mathrm{Q} = \sum_{i=1}^{5} q_i$. We use the marginal benefits earned by the next unit produced to estimate the potential benefit that could be obtained by producing a fraction of that unit (for example, marginal benefit of unit 8 is \$6, so producing 0.45 units above 7 will generate an extra benefit of 0.45*6=\$2.7); $M_i(e_i) = \mu_i^0(1 + \mu_i^z + \mu_i^t)[\varphi_i m_i(z) + (1-\varphi_i)s_i(\mathrm{z}) - e_i\ ]$ is the moral term, where we substitute the moral price and the effects of each treatment on it using the estimations presented in section 5 and $\varphi_i m_i + (1-\varphi_i)s$ is the average answer to the question about how much each participant believes should be the level each participant should produce in each round. The levels of $U_i(q_i)$ in the last column are calculated assuming every one of the 5 subjects in the group behaves as the average subject.

7 Conclusions

We assess the degree of complementarity between (a) a message informing subjects what the optimal level of a negative externality is, and (b) a tax that is insufficient to induce the optimal level of this externality (a nudge) in a set of experiments that seek to mimic a local public bad situation in which socially disconnected individuals contribute to a pollution problem. Using both parametric and non-parametric tests, along with econometric analysis, we find evidence that supports the following hypotheses. First, subjects exhibit baseline moral preferences. Second, subjects react to the message, decreasing their baseline level of production in a manner consistent with the hypothesis that such a message increase their moral price. Third, a tax partially crowds out these baseline moral preferences (non-separability). Finally, we find that the average level of the negative externality under the joint implementation of the Tax and Nudge exhibit a degree of complementarity, but this is not perfect. The two instruments exhibit a negative synergy when jointly implemented, showing that the effect of both instruments implemented together is lower than the sum of both instruments standing alone.

When comparing the relative effect of both instruments, our experiments provide evidence consistent with the effect of a tax on a negative externality being higher than those of a nudge in the form of a message informing players what the optimal level of the externality is, implicitly suggesting what to play to a utilitarian player. We obtain this result even though the implemented tax was only half of the tax needed to induce the socially optimum level of the externality. Moreover, such a low tax may add to the effect obtained by the implementation of the nudge alone. The reverse is not true, nevertheless, suggesting that taxes may complement nudges but not the other way around. This evidence is inconsistent with the policy recommendation of implementing nudges to complement low taxes, while we wait for the political will to increase taxes to develop.

Regarding the evidence in favor of baseline moral preferences, although subjects share very similar baseline social preferences (according to their behavior in the baseline treatment), the message do trigger greater morality (is more effective in affecting this moral price), in subjects with higher baseline social preferences and those not majoring in STEM/ECON majors. The effects of the nudge standing alone and both instruments jointly implemented differ between "Very prosocial" and non-STEM/ECON individuals, as well as between "moderately" and "less prosocial" and STEM/ECON, individuals.

8 References

Allcott, H. and Kessler, J. B. (2019). "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." *American Economic Journal: Applied Economics*, 11 (1): 236-76.

Alston, R. M., Kearl, J. R., & Vaughan, M. B. (1992). Is There a Consensus among Economists in the 1990's? *The American Economic Review*, 82(2), 203-209.

Ambec, S., Andersson, H., Cezera, S., Kanay, A., Ouvrard, B., Panzone, L. and Simon, S. (2024). "Taxing and nudging to reduce carbon emissions: Results from an online shopping experiment." *Working Paper*

Andreoni, J. (1989). Giving with impure altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy*, *97*(6), 1447–1458.

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, *100*(401), 464–477.

Antinyan, A., Horváth, G., & Jia, M. (2020). Curbing the consumption of positional goods: Behavioral interventions versus taxation. *Journal of Economic Behavior & Organization*, 179, 1-21.

Bandura, A. (1991). Social cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development* (Vol. 1, pp. 45-103). Hillsdale, NJ: Erlbaum.

Barmettler, F., E. Fehr, and C. Zehnder (2012). "Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory." *Games and Economic Behavior* 75(1): 17-34.

Bicchieri, C., & Dimant, E. (2022). Nudging with care: The risks and benefits of social information. *Public choice*, 191(3-4), 443-464.

Bicchieri, C., Dimant, E., Gächter, S., and Nosenso, D. (2019). "Social Proximity and the Evolution of Norm Compliance." *Games and Economic Behavior* 132: 59-71.

Blanchard, O., Gollier, C., & Tirole, J. (2023). The portfolio of economic policies needed to fight climate change. *Annual Review of Economics*, 15(1), 689-722.

Bowles, S., and S. Polania-Reyes (2012). "Economic incentives and social preferences: substitutes or complements?" *Journal of Economic Literature*, 50(2), 368-425.

Bowles, S. (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.

Bucholz, M., & Musshoff, O. (2021). Tax or green nudge? An experimental analysis of pesticide policies in Germany. *European Review of Agricultural Economics*, 48(4), 940-982.

Buckley, P., & Llerena, D. (2022). Nudges and peak pricing: A common pool resource energy conservation experiment. *Journal of Behavioral and Experimental Economics*, 101, 101928.

Carattini, S., Carvalho, M., & Fankhauser, S. (2018). Overcoming public resistance to carbon taxes. *Wiley Interdisciplinary Reviews: Climate Change*, *9*(5), e531.

Carlsson, F. and O. Johansson-Stenman (2019). "Optimal Prosocial Nudging." *Working Paper in Economics No 757, Department of Economics, University of Gothenburg.*

Carlsson, F., Gravert, C., Johansson-Stenman, O., & Kurz, V. (2021). The use of green nudges as an environmental policy instrument. *Review of Environmental Economics and Policy*, *15*(2), 216-237

Climate Leadership Council (2004). *Economists' Statement on Carbon Dividends.* Economists' Statement on Carbon Dividends | Climate Leadership Council (clcouncil.org)

Croson, R, and Marks, M. (2001). The Effect of Recommended Contributions in the Voluntary Provision of Public Goods. *Economic Inquiry, 39, issue 2, p. 238-49*

Dal Bó, E. and P. Dal Bó (2014). ""Do the right thing:" The effects of moral suasion on cooperation." *Journal of Public Economics* 117: 28-38.

Deci, E. L., & Ryan, R. M. (2013). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.

Delaney, J., & Jacobson, S. (2016). Payments or persuasion: common pool resource management with price and non-price measures. *Environmental and Resource Economics*, *65*, 747-772.

DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90(1), 81-116.

Dolan, P., & Metcalfe, R. (2015). Neighbors, knowledge, and nuggets: two natural field experiments on the role of incentives on energy conservation. *Becker Friedman Institute for Research in Economics Working Paper*, (2589269).

Drews S., Exadaktylos F., van den Bergh J C. (2020). Assessing synergy of incentives and nudges in the energy policy mix. *Energy Policy, Volume 144, 111605, ISSN 0301-4215.*

Fanghella, V., Ploner, M., & Tavoni, M. (2021). Energy saving in a simulated environment: An online experiment of the interplay between nudges and financial incentives. *Journal of Behavioral and Experimental Economics*, 93, 101709.

Frey, B.S. (1992). Tertium Datur: Pricing, Regulating and Intrinsic Motivation. *Kyklos*, 45(2), 161-184.

Frey, B. S., and R. Eichenberger (1992). Economics and economists: A European perspective. *The American Economic Review*, 82(2), 216-220.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American economic review*, *87*(4), 746-755.

Frey, B. S., and A. Stutzer (2008). "Environmental Morale and Motivation." In *The Cambridge Handbook of Psychology and Economic Behaviour*, edited by Alan Lewis, 406-28. Cambridge and New York: Cambridge University Press.

Fuller, D., & Geide-Stevenson, D. (2003). Consensus among economists: revisited. *The journal of economic education*, 34(4), 369-387.

Glaeser, E. L., 2006. "Paternalism and Psychology," *Working Paper Series rwp06-006, Harvard University, John F. Kennedy School of Government.*

Glaeser, E. L, 2014. "The Supply of Environmentalism: Psychological Interventions and Economics," *Review of Environmental Economics and Policy, Association of Environmental and Resource Economists,* vol. 8(2), pages 208-229.

Gravert, C., & Olsson Collentine, L. (2021). When nudges aren't enough: Norms, incentives and habit formation in public transport usage. *Journal of Economic Behavior & Organization*, 190, 1–14. https://doi.org/10.1016/j.jebo.2021.07.012

Hahn, R. W. and Hendren, N. and Metcalfe, R. and Sprung-Keyser, B., 2024. "A Welfare Analysis of Policies Impacting Climate Change". NBER Working Paper No. w32728, Available at SSRN: https://ssrn.com/abstract=4908545

Harrington, W., Morgenstern, R. D., & Sterner, T. (2004). Choosing environmental policy. *Resources for the Future Press, Washington*.

Hernández, F., Jaime, M., and F. Vásquez (2024). Nudges versus prices: Lessons and challenges from a water-savings program. *Energy Economics* 134: 107546.

Hilton, D., L. Charalambides, C. Demarque, L. Waroquier and C. Raux (2014), "A Tax Can Nudge: The Impact of an Environmentally Motivated Bonus/Malus Fiscal System on Transport Preferences," *Journal of Economic Psychology*, 42, 17-27.

Ilzetzki, E. and S. Jain (2024, July 15). *From pledge to practice: Assessing the UK's net zero commitment*. VoxEU - CEPR. https://cepr.org/voxeu/columns/pledge-practice-assessing-uks-net-zero-commitment

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy*, 10 (1): 240-6

Johansson, O. (1997). Optimal Pigovian taxes under altruism. *Land Economics*, 297-308.

Levitt, S. and J. List (2007). "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?" *Journal of Economic Perspectives* 21(2): 153-174.

Lilliestam, J., Patt, A., & Bersalli, G. (2021). The effect of carbon pricing on technological change for full energy decarbonization: A review of empirical ex-post evidence. *Wiley Interdisciplinary Reviews: Climate Change*, *12*(1), e681.

Loewenstein, G. and T. O'Donoghue (2006) "We Can Do This the Easy Way or the Hard Way": Negative Emotions, Self-Regulation, and the Law," *University of Chicago Law Review*: Vol. 73: Iss. 1, Article 10.

Lopez, M. C., J. Murphy, J. Spraggon and J. Stranlund (2012). "Comparing the Effectiveness of Regulation and Pro-Social Emotions to Enhance Cooperation: Experimental Evidence from Fishing Communities in Colombia." *Economic Inquiry* 50(1): 131-142.

Mackay, M., Yamazaki, S., Jennings, S., Sibly, H., van Putten, I. E., & Emery, T. J. (2019). The influence of nudges on compliance behaviour in recreational fisheries: a laboratory experiment. *ICES Journal of Marine Science*, *77*(6), 2319-2332.

Maris, R., Dorner, Z., & Carlsson, F. (2024). Nudges and Monetary Incentives: A Green Partnership? *Working Papers in Economics 842, University of Gothenburg, Department of Economics.*

Marks, M.B., Schansberg, D.E. and Croson, R.T.A. (1999) 'Using Suggested Contributions in Fundraising for Public Good', *Nonprofit Management & Leadership, 9(4), p. 369. doi:10.1002/nml.9403.*

Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, 119(1), e2107346118.

Miao, Haoran, Trandafir, Simona and Uchida, Emi. (2017). "The effect of informational nudges to promote voluntary behavior to reduce nonpoint source pollution: A randomized controlled trial in the field." *Unpublished working paper.*

Mizobuchi, K., & Takeuchi, K. (2013). The influences of financial and non-financial factors on energy-saving behaviour: A field experiment in Japan. *Energy Policy*, 63, 775-787.

My B. and B. Ouvrard (2019). "Nudge and Tax in an Environmental Public Goods Experiment: Does Environmental Sensitivity Matter?" *Resource and Energy Economics*, 55, 24-48.

Nakagawa, M., Lefebvre, M., & Stenger, A. (2022). Long-lasting effects of incentives and social preference: A public goods experiment. *Plos one*, *17*(8), e0273014.

Panzone, L., A. Ulph, D. Zizzo, D. Hilton and A. Clear (2021). "The impact of environmental recall and carbon taxation on the carbon footprint of supermarket shopping." *Journal of Environmental Economics and Management* 102137.

Pellerano, J. A., Price, M. K., Puller, S. L., & Sánchez, G. E. (2017). Do extrinsic incentives undermine social norms? Evidence from a field experiment in energy conservation. *Environmental and resource Economics*, *67*, 413-428.

Ricketts, M., & Shoesmith, E. (1992). British economic opinion: positive science or normative judgment? *The American Economic Review*, 82(2), 210-215.

Romaniuc, R. (2016). What makes law to change behavior? An experimental study. *Review of Law & Economics*, 12(2), 447-475.

Schall, D. L., Wolf, M., & Mohnen, A. (2016). Do effects of theoretical training and rewards for energy-efficient behavior persist over time and interact? A natural field experiment on eco-driving in a company fleet. *Energy Policy*, *97*, 291-300.

Schwartz, S. H. (1977). Normative influences on altruism. In *Advances in experimental social psychology* (Vol. 10, pp. 221-279). Academic Press.

Spraggon J, J. Oxoby R (2010). Ambient-Based Policy Instruments: The Role of Recommendations and Presentation. *Agricultural and Resource Economics Review*. 2010;39(2):262-274.

Stern, P.C. (1999). Information, Incentives, and Proenvironmental Consumer Behavior. *Journal of Consumer Policy* 22, 461–478.

Sterner, T., & Coria, J. (2013). *Policy instruments for environmental and natural resource management*. Routledge.

Sterner, T., Ewald, J., & Sterner, E. (2024). Economists and the climate. *Journal of Behavioral and Experimental Economics*, *109*, 102158.

Stiglitz, J., Barrett, S., & Kaufman, N. (2024). How Economics Can Tackle the 'Wicked Problem'of Climate Change.

Sudarshan A. (2017). Nudges in the marketplace: The response of household electricity consumption to information and monetary incentives. *Journal of Economic Behavior & Organization, Volume 134, Pages 320-335, ISSN 0167-2681.*

Sunstein, Cass and Thaler, Richard. (2009). *NUDGE: Improving Decisions About Health, Wealth, and Happiness.* Penguin Books.

Sunstein, Cass. (2017). "Nudges that fail." *Behavioral Public Policy* 1(1): 4-25.

Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2018). A systematic scoping review of the choice architecture movement: Toward understanding when and why nudges work. *Journal of Behavioral Decision Making*, 31(3), 355-366.

Szaszi, B., Higney, A., Charlton, A., Gelman, A., Ziano, I., Aczel, B., ... & Tipton, E. (2022). No reason to expect large and consistent effects of nudge interventions. *Proceedings of the National Academy of Sciences*, 119(31), e2200732119.

van den Bergh, J., and Savin, I. (2021). Impact of carbon pricing on low-carbon innovation and deep decarbonisation: controversies and path forward. *Environmental and Resource Economics*, *80*(4), 705-715.

van den Bergh, J. C., & Botzen, W. W. (2024). Assessing Criticisms of Carbon Pricing. *International Review of Environmental and Resource Economics*, 18(3), 315-384.

Vélez, M. A., J. Murphy, and J. Stranlund (2010). "Centralized and Decentralized Management of Local Common Pool Resources in the Developing World: Experimental Evidence from Fishing Communities in Colombia." *Economic Inquiry* 48(2): 254-265.

World Bank Group (2023). State and Trends of Carbon Pricing 2019. Washington, DC: World Bank. © World Bank.

Xu, L., Ling, M., & Wu, Y. (2018). Economic incentive and social influence to overcome household waste separation dilemma: A field intervention study. Waste management, 77, 522-531.

9 Appendix A: Instructions

9.1 Baseline + Tax

Introduction

Welcome and thank you for agreeing to participate in this activity!

You are about to participate in two exercises to inform social researchers about persons´ decisions. These exercises will be carried out through a computer. In total, today's activity will not take more than 75 minutes to complete.

You have already earned $U 150 for agreeing to participate and show up on time for this activity. The decision-making exercises you are about to participate in will give you a chance to earn more money. The amount of additional money you earn will depend on the decisions you make in the exercises.

Below, we will describe the general characteristics of today's activity. Next, we will describe the first exercise and explain how to enter your decisions into the computer. After the first exercise, we will describe the second.

Before reading the instructions, we want to announce some simple rules:

- Please silence your phone and use it only as a calculator, if needed.
- Please do not talk to any other participants during the exercise.
- If you have a question, please raise your hand.
- If you do not follow the instructions, we will ask you to conclude your participation.
- We hope you can stay until the exercise is over. However, if you must leave before the exercise ends, you will be able to keep the $U 150 you earned by showing up today, but you will lose any additional money you have earned.

Let us move on to describe the general characteristics of today's activity and read the instructions of the exercises. Please pay careful attention to the instructions. If you have any questions while we explain in detail the exercises, please raise your hand.

Activity Overview

Today's activity consists of two decision-making exercises. Each of the exercises will last 5 rounds of 2 minutes each. In each round we will ask you to make the same type of decision. You may think of these rounds as years or months.

The two exercises are completely independent. When Exercise 1 ends, we will resume the activity by starting Exercise 2 with a new round 1.

You may have heard about these exercises from someone. Although some of these exercises may be similar to some that other people have participated in, there may be some differences as well.

Once today's two exercises have been completed, we will ask you to answer a short questionnaire.

Group

In today's activity, you are part of a group of 5 subjects. You will only interact with the others 4 members of your group.

Payment procedure to participants

As mentioned above, you will be paid $U 150 for agreeing to participate in this activity, plus your earnings from the exercises. It is important for us to let you know that the payment procedure preserves the confidentiality of your decisions. We do this in the following way. First, we store your personal data collected for payment in a different file from the one where we store your decisions in the activity and your answers in the questionnaire. The experimenter will not see your personal information at any time.

The payment procedure is as follows. When the activity finishes, the experimenter will leave and the assistant will extract the information of your final earnings from the server. With this information, the assistant will prepare payments and a receipt for each participant. The receipt contains the name of each participant and also indicates the amount of money they won. Once this task is completed, the assistant will proceed with the payments. Each participant must show their identity card to the assistant to ensure that the participant receives the correct amount. When the participant receives the money, he has to sign the receipt.

Each participant receives their earnings in private.

With this procedure, nobody (including the experimenter and assistant) could know what decisions you made in the experiment.

Description of the first exercise

In the first decision-making exercise, you are a producer of a non-specific good and must choose the level of production of this good. In each round, you can produce any quantity between 0 and 10 units of the good. For each unit you produce, you will earn an economic income, expressed in Uruguayan pesos ($U). The more units of good you decide to produce, the greater your income in $U. It is your decision how much to produce.

In a moment, we will tell you how much money you will earn for producing units.

The rest of the 4 members of your group are also producers like you. They must also choose, in each round, an individual quantity of production between 0 and 10. The income per unit produced is the same for all participants.

In addition, each unit you produce reduces the earnings of each of the 5 members of the group, including yourself, by a certain amount of money. Similarly, each unit produced by each of the other members of your group reduces the earnings of you and the rest by the same amount.

Your earnings for each round will then be that resulting from the level of production that you decide to carry out and the level of production that the other members of your group choose.

The earnings of each period are added to the earnings of previous periods to make up the accumulated earnings up to that period.

The screens

In today's activity, you will need to enter your decisions into the computer.

During the first exercise, you will see two screens. A decision screen, in which you decide your production level for the period, and an information screen, in which you will be informed of your earnings for the period and accumulated earnings.

You can see at a purely HYPOTHETICAL example of the decision screen that you will see during this experiment in Figure 1. The screen is divided into several parts. We will explain each of those parts below.

Period

In the upper left of Figure 2, where the blue arrow points, you will be able to see the current period number of the experiment and the total number of periods. In this HYPOTHETICAL example, it informs that you are in period 1 out of a total of 3 periods. This is an example of 3 rounds but remember that the activity consists of 5 rounds.

Income per Unit of Production

As we explained previously, in each round, you can produce any amount between 0 and 10 units of the good. For each unit you produce, you will automatically obtain a certain economic income, expressed in Uruguayan pesos ($U). The more units of the good you decide to produce, the higher your income in $ U will be. We will refer to these benefits as "Ingresos por Producción". The income per unit produced is the same for all participants.

During the exercises, a table located on the left of the screen, with the title "Ingresos por Producción", will inform you the amount of economic income that you will be able to obtain for each unit that you produce, as well as the total income. The values that you will see in this table will not change during all periods or rounds.

A HYPOTHETICAL example of this table is illustrated in Figure 3, where the blue arrow points. In this HYPOTHETICAL example, the table tells you that you get an income of $U 45 for the first unit you produce. If you decide to produce a second unit, you get an additional $U 33. If you decide to produce the third unit, your profits increase by $U 27. And so on. Note that you may receive a different amount of money for each unit produced. Suppose you decide to produce 3 units in total. The total income from the production of the 3 units will then be $U 105 (= 45+33+27).

If you decide to produce a fourth unit, you will receive an additional $U 22. Your total income in this case will be $U 127 for the four units (= 45+33+27+22).

Remember that these values are HYPOTHETICAL and may look nothing like the values you will see during the actual exercise you are about to participate in.

Loss experienced by all members of the group for each unit that is produced in the group

Where the blue arrow points in Figure 4, you are informed by how much the earnings of the rest of the members of your group are reduced per each of the units you decide to produce and by how your earnings are reduced per unit that the rest of the members of your group decide to produce. In the HYPOTHETICAL example in Figure 5, the earnings of each of the members of the group are reduced by $U 3 for each unit that you and the rest of the members of your group decide to produce.

For example, suppose that you produce 4 and the other group members produce 11 combined, for a total of 15. The reduction of the group's earnings will be $U 3 x 15=$U 45.

Your decision: the number of units produced

Now let us see how to enter your decision on how much to produce into the computer. This is done as follows. At the bottom of the screen, you will see a cell next to the title "Cantidad de unidades que desea producir ". In that cell (which is pointed to with a blue arrow in Figure 5), you must enter the number of units you choose to produce.

In each round, you must decide the production level only for that round.

Once you have entered the number of units you want to produce, press the red button with the legend "Ingresar" shown in Figure 5.

Remaining time

At the top right of the screen, where the blue arrow points out in Figure 6, you will be able to see the time remaining, in seconds, for the end of the period. In this HYPOTHETICAL example illustrated in Figure 6, you have 112 seconds left to decide how many units to produce.

In the experiment, you will have a maximum of 2 minutes (120 seconds) to make your decision. Once the 120 seconds have passed, the timer will stop at 0 and below the timer a red flashing warning will appear, telling you to enter a production level.

You must enter a production level to move on to the next stage of the experiment.

The Earnings of the period

After the round is completed, you will receive a summary of your earnings of the period. This summary is presented on a new screen that opens automatically.

A HYPOTHETICAL example of this screen is shown in Figure 7.

In the center of the screen, you are informed:

(a) the number of units you chose to produce (to the right of the legend "Su Producción"),

(b) the number of units the group decided to produce (to the right of the legend “Producción total del grupo”) and

(c) your earnings of this period in Uruguayan pesos.

In the HYPOTHETICAL example illustrated in Figure 7, you produced 4 units, the group jointly produced 14 units, and your profit in this period was $U 85.

Below, you can see a table. This table informs you, for the last and previous periods, the number of the period, your level of production, your earnings for the period, and the earnings accumulated in all periods.

In this HYPOTHETICAL example You produced 4 units in period 3, obtained a profit of $U 85 in that period, which added to the accumulated until period 2 ($U 85), yields a cumulative of $U 170.

This screen will be active for 20 seconds, after which it automatically closes and opens again the main screen where you must choose the production level of the next period.

This process is repeated at the end of each of the periods that the game lasts.

Example

Let's look at a HYPOTHETICAL example of the first exercise.

Round 1

Suppose that, in this HYPOTHETICAL example, you decide to produce 4 units in the first round. Your total income from that production will then be $U 127 (= 45 + 33 + 27 + 22).

Additionally, suppose the rest of the members of your groups decide to produce 0 units in the first round, so that the total number of units produced by the group in the first round is 4 (the 4 units that you decided to produce). Your losses from the group's production are then $U 12 (= 4*$3); the same ones that will have each of the remaining 4 members of the group. Your final net earnings in this round will then be $ 115.

Your production level (4), the group production level (4), your earnings in this period ($U 115) will appear on the information screen.

Round 2

Continuing with the hypothetical example, suppose in the second round you decide to produce 0 units. Therefore, you will not get income from that production.

Suppose the total produced by the group in this round are 10 units. Therefore, you (like the other members of your group) will have a loss of $U 30 (=10*$3) in this round.

Your production level (0), the group's production level (10), your earnings in this period ($ U -30) and your accumulated earnings in the two rounds ($ U85), are reported to you on the information screen.

Round 3

To finish with the example, suppose that in the third round you decide to produce 4 units. Your total income from that production will then be $U 127, the same as in round 1, in which you also produced 4.

Also suppose that the total produced by the group is 14 units (the 4 units you chose to produce + 10 units produced in total by the remaining 4 members of the group). Your losses from group production will then be $U 42 (=14*$3).

Therefore, you will have earning of $ 85 in this round.

Your decision to produce 4, the total production of the group (14), your earnings in the third round ($ 85) and your accumulated earnings in the three rounds ($ 170), are reported to you on the information screen.

Trial periods

The first exercise includes two trial periods, in addition to the actual periods. The earnings that you obtain in these trial periods will not count as part of the actual earnings that you will obtain at the end of the experiment.

We have reached the end of the explanation for exercise 1. If you have a question, please ask it right now.

If there are no further questions, we will start with the first exercise.

Description of the Second Exercise

We shall now proceed to explain the second exercise.

The second exercise is the same as the first exercise, except that You and the rest of the group must pay a tax per unit produced.

The decision screen for the second exercise

The amount of tax per unit produced will be indicated in the upper right corner of your decision screen, where the blue arrow in the Figure 11 points. In this HYPOTHETICAL example, you must pay a tax of $U 7 per unit produced.

Example

Let's look at an example to illustrate the exercise.

Round 1

As in the first exercise, suppose you decide to produce a total of 4 units in the first round. As we have seen, for the production of these 4 units you will receive a total income of $ U 127.

However, you now have to pay taxes on this amount of production. The tax is $ 7 for each unit produced. If you decide to produce 4 units, your total tax payments will be = $ 7 * 4 units = $ 28.

As in the first exercise, it is still the case that the income of each member of the group are reduced by an amount equal to the number of total units produced by the group, multiplied by $U 3. As in the first exercise example, suppose that the rest of the members of your group decide not to produce. In this hypothetical case, the group produced a total of 4 units (yours), so your losses due to the group production are $U 12 (= $U 3*4).

Subtracting these losses and your tax payments from your production income is $U 87, your earnings for the period.

Your production level (4), the group production level (4) and your earnings in this period ($U 87) will appear on the information screen.

Round 2

Continuing with the example, suppose that in the second round you decide not to produce. Therefore, you will not earn income from production, and you will not pay taxes. Assuming that the rest of the members of your group decide to produce 10 units together, the losses from group production will be $U 30. Therefore, your net earnings from this round will be $U -30.

Your decision to produce 0, your second-round earnings, which, in this hypothetical example, are $U – 30 (i.e. a loss of $U 30) and your accumulated earnings in the two rounds ($U 57), are reported to you on the information screen.

Round 3

To finish the example, we assume that you decide to produce 5 units in the third round.

Your income from that production will be $U 144 (= 45 + 33 + 27 + 22 + 17).

The tax payment, on the other hand, is $U 35 (=$7*5).

Assume that in round 3 the rest of the members of your group decide to produce 10 units together. The Total of Units Produced by the group is 15. The losses due to group

production in this round will be $U 45 (15*$3) and your net earnings will be $U 64 (144-35-45).

Your decision to produce 5, your third round earnings ($U 64), and your accumulated earnings in all three rounds ($U 121), are reported to you on the information screen.

Remember that the gains that you obtain in the second exercise are added to those obtained in the first exercise of the activity. So, in this hypothetical example, your total earnings for participating in the activity are the sum of your accumulated earnings in the three rounds of the first exercise ($ 170), plus the accumulated earnings in the three rounds of the second exercise ($U 121), which equals $ U 291.

Remember that these values are HYPOTHETICAL and may look nothing like the values you will see during exercise.

Finally, the earnings obtained in this second exercise will be accumulated to those obtained in the previous exercise of the activity.

We have reached the end of the explanation for exercise 2. If you have a question, please ask it right now.

If there are no further questions, we will start with exercise 2.

Questionnaire-completion of the activity

The last part of today's activity is to answer a questionnaire.

We will now ask you to answer the questionnaire that will appear on your screens. This will take about 10 minutes. You must complete the questionnaire before leaving the room.

If there are no more questions we will start with the questionnaire.

End of the experiment

9.2 Baseline + Message

The introduction, the general description of the activity, the description of the first exercise and the questionnaire-completion of the activity are the same that in the Baseline + Tax instructions, what differs from the instructions is the description of the second exercise, which in this case deals with the message and is presented below.

Description of the Second Exercise

We shall now proceed to explain the second exercise.

The second exercise is the same as the first one, except that in this exercise you and the rest of the members of your group will receive an informative message about what is the level of individual production that maximizes the earnings of the group. You may consider this information for your production decisions or not.

The decision screen for the second exercise

This message will appear on the right of the screen, where the blue arrow in the Figure 11 points.

In this second exercise, production decisions are entered into the computer in the same way as in the previous exercise.

Finally, the earnings obtained in this second exercise will be accumulated to those obtained in the previous exercise of the activity.

We have reached the end of the explanation for exercise 2. If you have a question, please ask it right now.

If there are no further questions, we will start with exercise 2.

9.3 Baseline + Message+Tax

The introduction, the general description of the activity, the description of the first exercise and the questionnaire-completion of the activity are the same that in the Baseline + Tax instructions, what differs from the instructions is the description of the second exercise, which in this case deals with the message and the tax as a whole and is presented below.

Description of the Second Exercise

We shall now proceed to explain the second exercise.

The second exercise is the same as the first exercise, except for two issues.

1. In this exercise You and the rest of the members of your group will receive an informative message about the level of individual production that maximizes the

earnings of the group, which you may or may not consider for your production decisions.

2. At the same time, you and the rest of the group must pay a tax per unit produced.

The decision screen for the second exercise

In this second exercise there are some modifications to what you will see on the decision screen.

First, in the upper right corner of your screen, where the blue arrow in the Figure 11 points, it will indicate the amount of tax. In this HYPOTHETICAL example, you must pay a tax of $U 7 per unit produced.

Second, the informational message about the level of production that maximizes your group's profits will appear on the right of the screen, where the blue arrow in the Figure 12 points.

Example

Let's look at an example to illustrate the exercise.

Round 1

As in the first exercise, suppose you decide to produce a total of 4 units in the first round. As we have seen, for the production of these 4 units you will receive a total income of $ U 127.

However, you now have to pay taxes on this amount of production. The tax is $ 7 for each unit produced. If you decide to produce 4 units, your total tax payments will be = $ 7 * 4 units = $ 28.

As in the first exercise, it is still the case that the income of each member of the group are reduced by an amount equal to the number of total units produced by the group, multiplied by $U 3. As in the first exercise example, suppose that the rest of the members of your group decide not to produce. In this hypothetical case, the group produced a total of 4 units (yours), so your losses due to the group production are $U 12 (= $U 3*4).

Subtracting these losses and your tax payments from your production income is $U 87, your earnings for the period.

Your production level (4), the group production level (4) and your earnings in this period ($U 87) will appear on the information screen.

Round 2

Continuing with the example, suppose that in the second round you decide not to produce. Therefore, you will not earn income from production and you will not pay taxes. Assuming that the rest of the members of your group decide to produce 10 units together, the losses from group production will be $U 30. Therefore, your net earnings from this round will be $U -30.

Your decision to produce 0, your second round earnings, which, in this hypothetical example, are $U – 30 (i.e. a loss of $U 30) and your accumulated earnings in the two rounds ($U 57), are reported to you on the information screen.

Round 3

To finish the example, we assume that you decide to produce 5 units in the third round.

Your income from that production will be $U 144 (= 45 + 33 + 27 + 22 + 17).

The tax payment, on the other hand, is $U 35 (=$7*5).

Assume that in round 3 the rest of the members of your group decide to produce 10 units together. The Total of Units Produced by the group is 15. The losses due to group production in this round will be $U 45 (15*$3) and your net earnings will be $U 64 (144-35-45).

Your decision to produce 5, your third-round earnings ($U 64), and your accumulated earnings in all three rounds ($U 121), are reported to you on the information screen.

Remember that the gains that you obtain in the second exercise are added to those obtained in the first exercise of the activity. So, in this hypothetical example, your total earnings for participating in the activity are the sum of your accumulated earnings in the three rounds of the first exercise ($ 170), plus the accumulated earnings in the three rounds of the second exercise ($U 121), which equals $ U 291.

Remember that these values are HYPOTHETICAL and may look nothing like the values you will see during exercise.

Finally, the earnings obtained in this second exercise will be accumulated to those obtained in the previous exercise of the activity.

We have reached the end of the explanation for exercise 2. If you have a question, please ask it right now.

If there are no further questions, we will start with exercise 2.

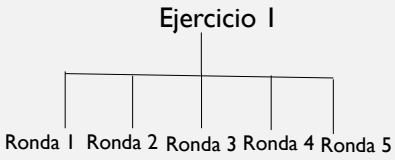
10 Appendix B: Power Point presentation

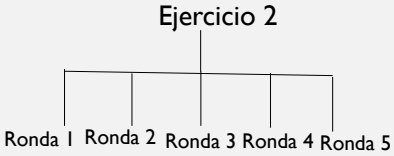
10.1 Baseline + Tax

REGLAS

INSTRUCCIONES

- Actividad de hoy: **2 ejercicios de toma de decisiones**.
- Cada uno consta de 5 **rondas** de 2 minutos.





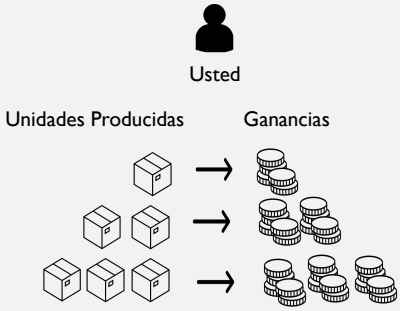
GRUPO

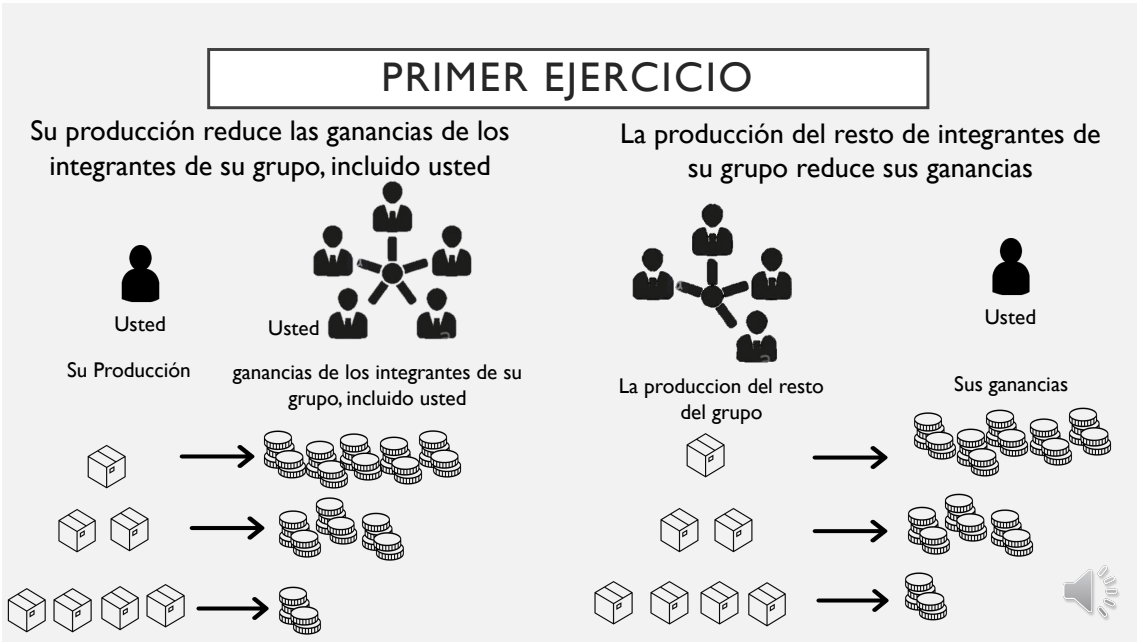
En la actividad de hoy Usted forma parte de un grupo de 5 sujetos. Usted va a interactuar solamente con los 4 miembros de su grupo.

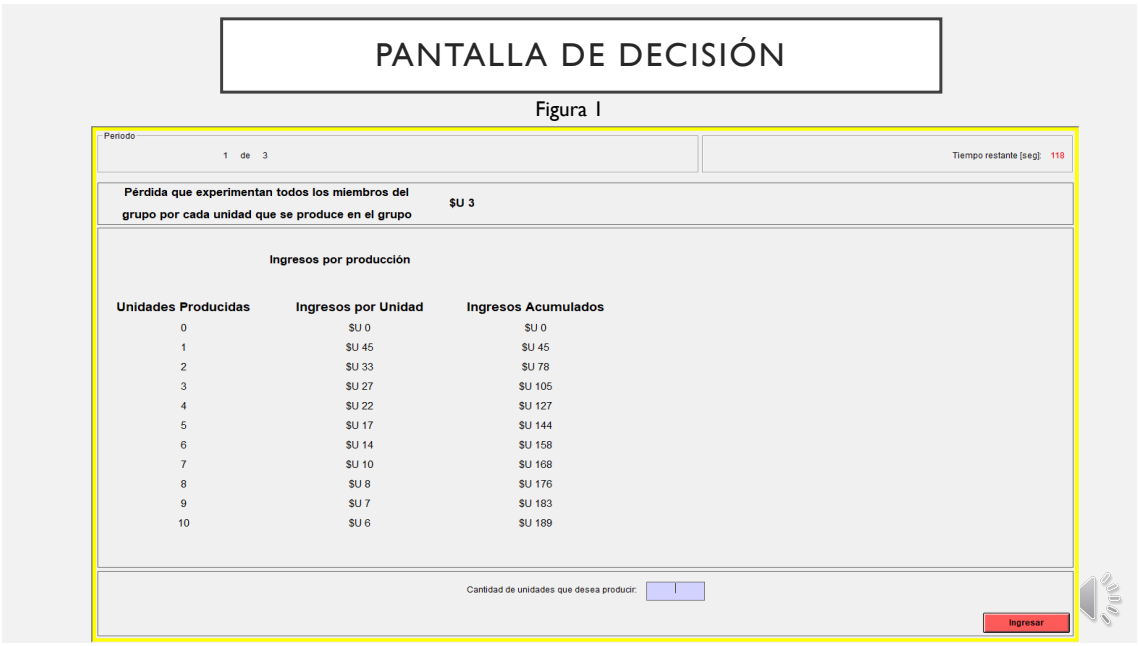
Usted

PRIMER EJERCICIO

- Usted y el resto de los integrantes de su grupo son productores de un bien inespecífico
- Todos pueden producir de **0 a 10 unidades**
- Por cada unidad que Usted produzca, obtendrá un ingreso económico, expresado en pesos uruguayos ($U).




PRIMER EJERCICIO
Su producción reduce las ganancias de los integrantes de su grupo, incluido usted
La producción del resto de integrantes de su grupo reduce sus ganancias
Usted
Usted
Usted
Su Producción
ganancias de los integrantes de su grupo, incluido usted
La producción del resto del grupo
Sus ganancias


PANTALLA DE DECISIÓN
Figura 1
Periodo
1 de 3
Tiempo restante [seg]: 118
Pérdida que experimentan todos los miembros del grupo por cada unidad que se produce en el grupo
$U 3
Ingresos por producción
Unidades Producidas
Ingresos por Unidad
Ingresos Acumulados
0 $U 0 $U 0
1 $U 45 $U 45
2 $U 33 $U 78
3 $U 27 $U 105
4 $U 22 $U 127
5 $U 17 $U 144
6 $U 14 $U 158
7 $U 10 $U 168
8 $U 8 $U 176
9 $U 7 $U 183
10 $U 6 $U 189
Cantidad de unidades que desea producir:
Ingresar

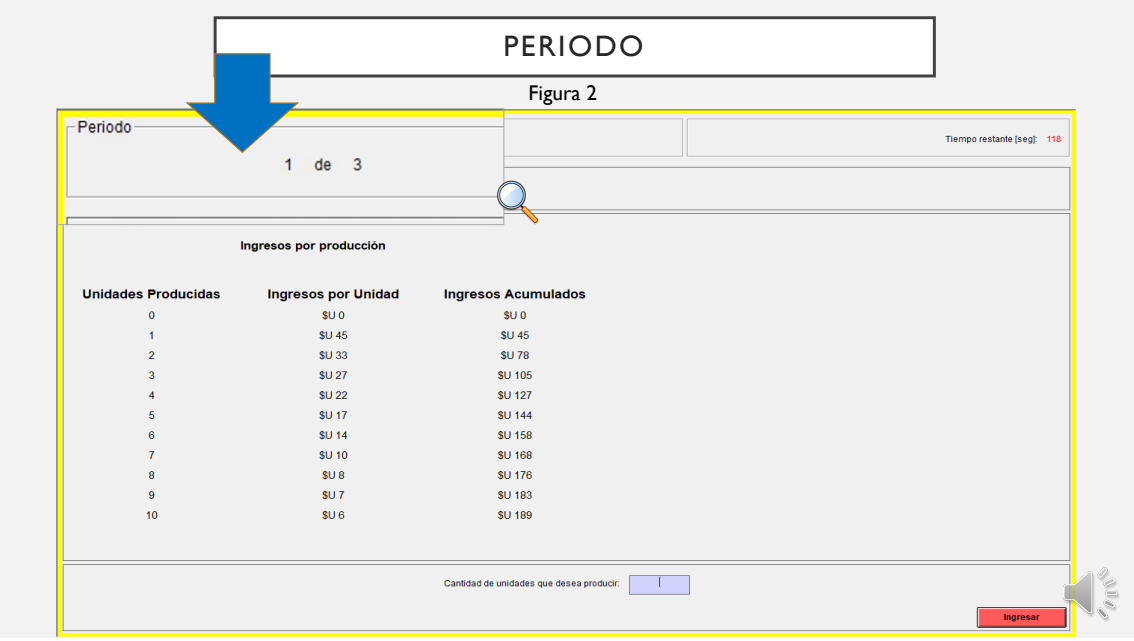


Figura 2

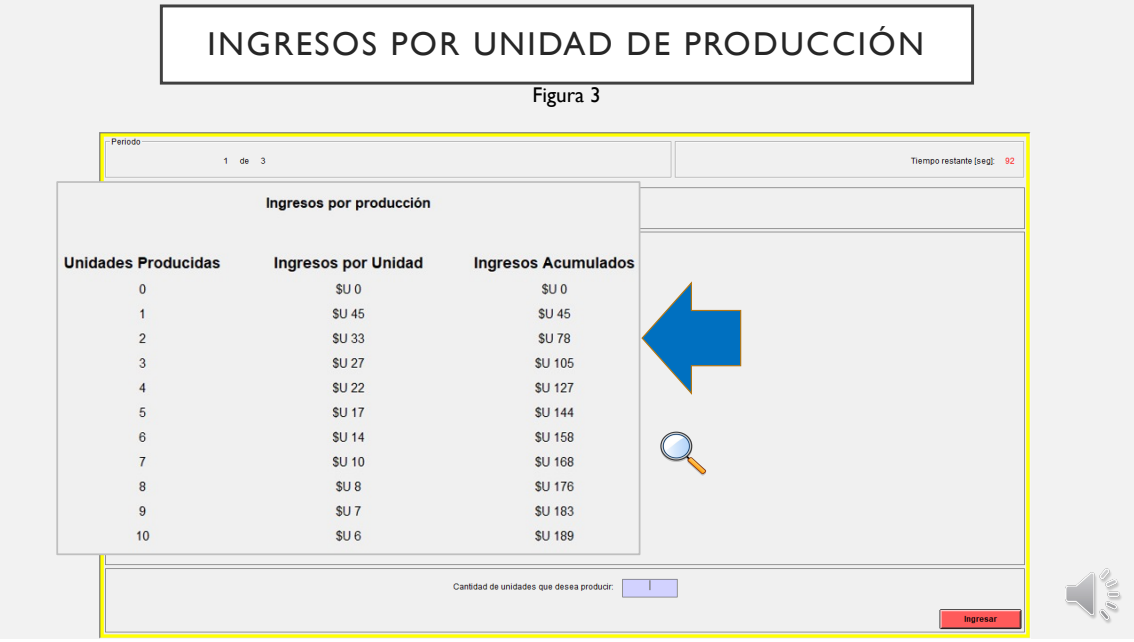
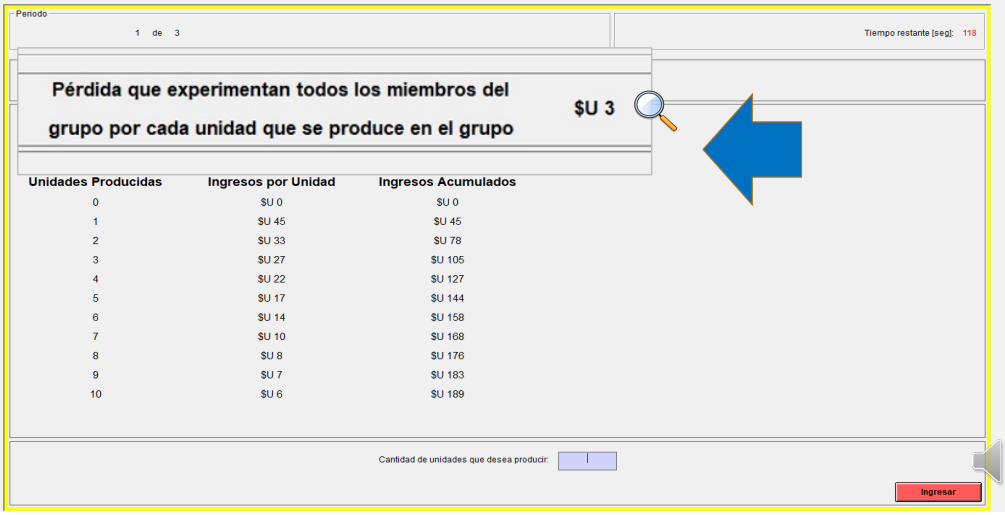


Figura 3

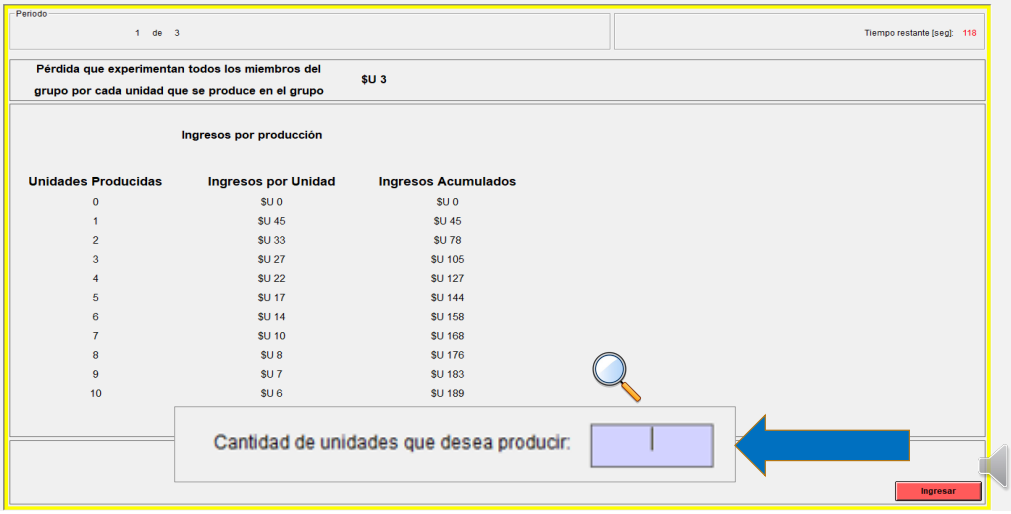
PÉRDIDAS POR UNIDAD DE PRODUCCIÓN

Figura 4



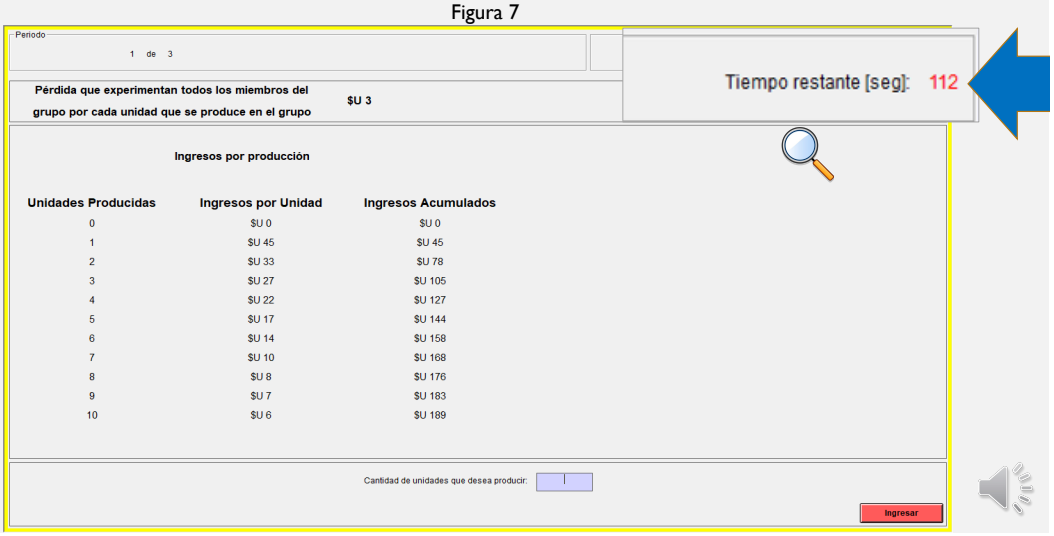
SU DECISIÓN: LA CANTIDAD DE UNIDADES PRODUCIDAS

Figura 6



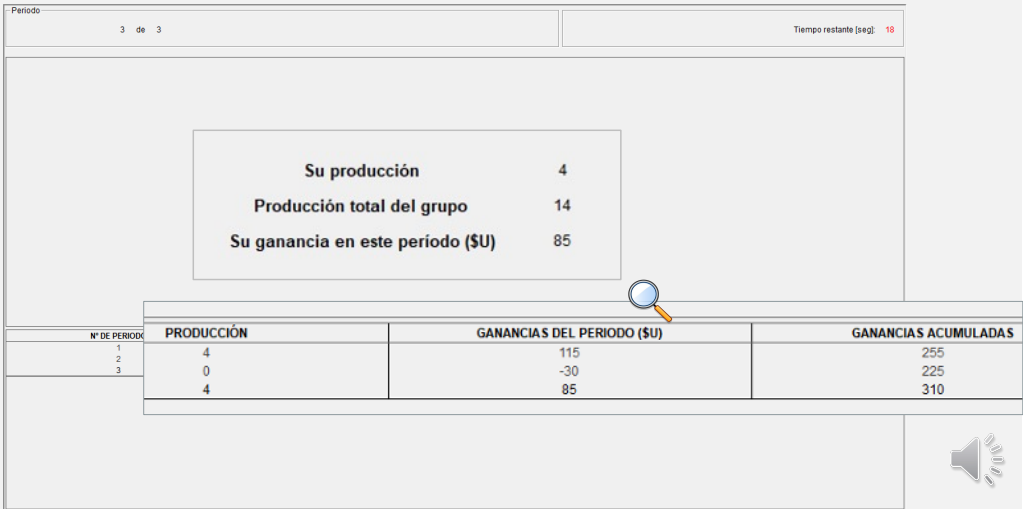
TIEMPO RESTANTE

Figura 7



PANTALLA DE INFORMACIÓN

Figura 8



EJEMPLO

Ronda 1:

- Suponga que Usted decide producir **4 unidades**.

Cantidad que decide producir del bien	Ingresos por unidad
0	0
1	45
2	33
3	27
4	22
5	17
6	14
7	10
8	8
9	7
10	6

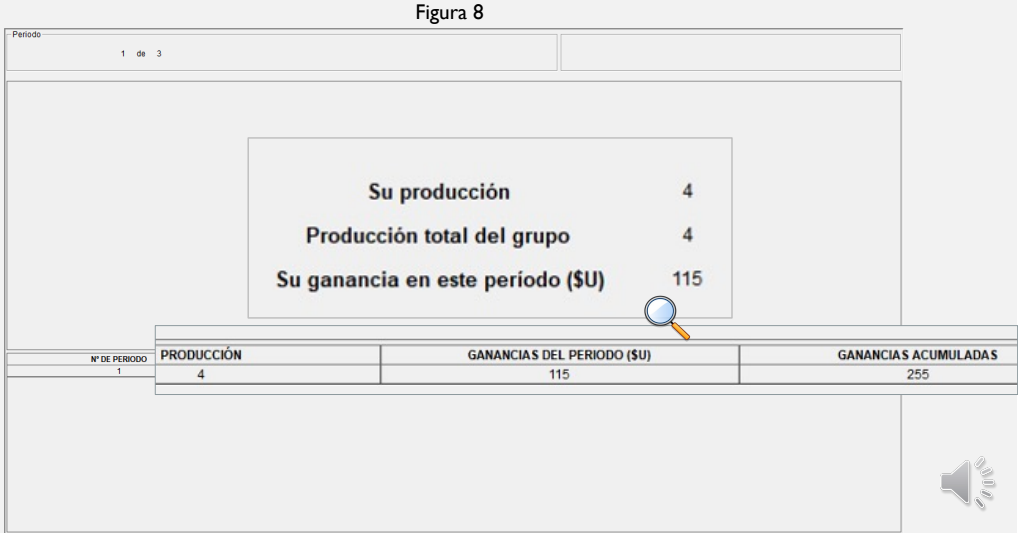
Sus ganancias por su producción serán: 45+33+27+22= **127**

EJEMPLO

A	B	C	D	E	F	G
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Total Producido por el Grupo	Mi Pérdida por producción del grupo (D*$3)	Mis ganancias del periodo (C-E)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	4	12	115	255
2						
3						
Total						

GANANCIAS PRIMERA RONDA

Figura 8



EJEMPLO

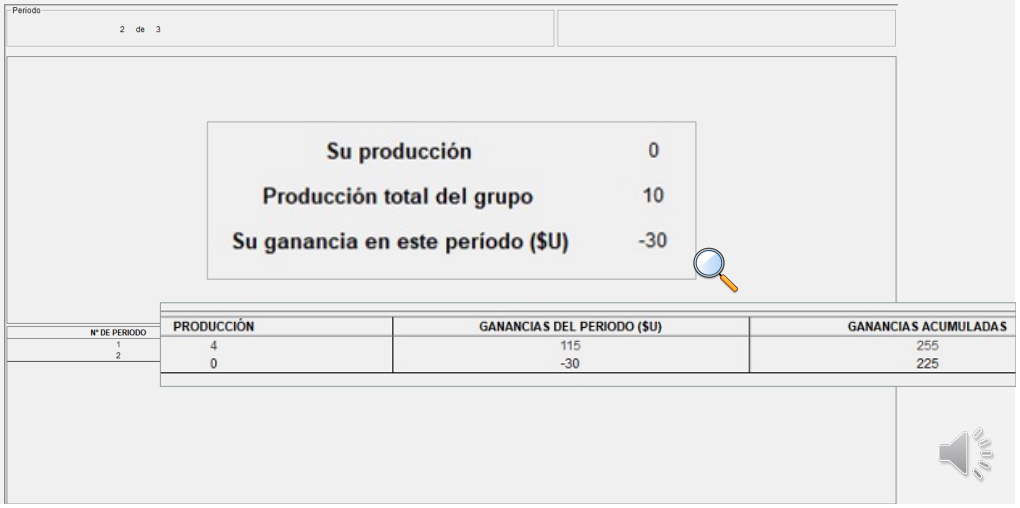
A	B	C	D	E	F	G
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Total Producido por el Grupo	Mi Pérdida por producción del grupo (D*$3)	Mis ganancias del periodo (C-E)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	4	12	115	255
2	0	0				
3						
Total						

EJEMPLO

A	B	C	D	E	F	G
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Total Producido por el Grupo	Mi Pérdida por producción del grupo (D*$3)	Mis ganancias del periodo (C-E)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	4	12	115	255
2	0	0	10	30	-30	225
3						
Total						

GANANCIAS SEGUNDA RONDA

Figura 9

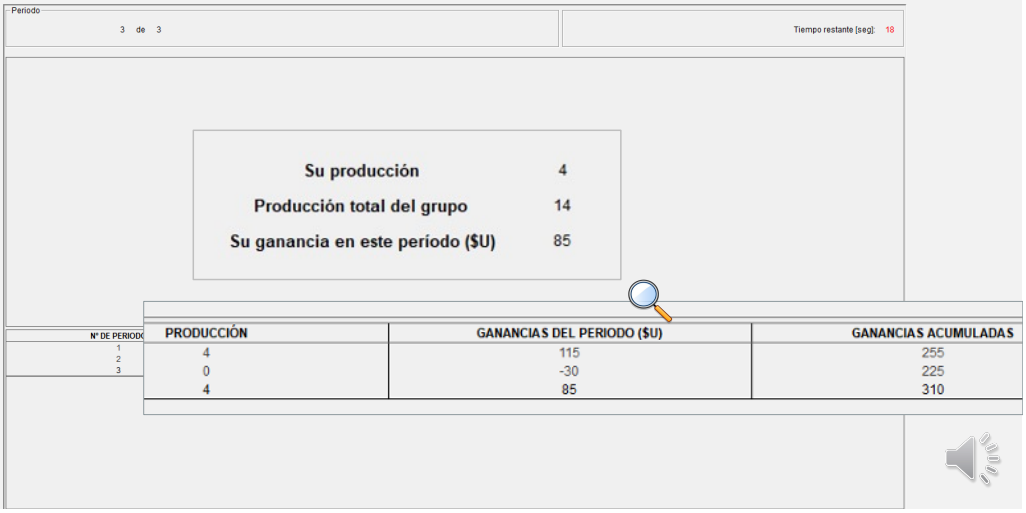


EJEMPLO

A	B	C	D	E	F	G
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Total Producido por el Grupo	Mi Pérdida por producción del grupo (D*$3)	Mis ganancias del periodo (C-E)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	4	12	115	255
2	0	0	10	30	-30	225
3	4	127	14	42	85	310
Total						

GANANCIAS TERCERA RONDA

Figura 10



PERÍODOS DE PRUEBA

- El primer ejercicio incluye dos períodos de prueba, en adición a los períodos reales.
- Las ganancias que Usted obtenga en estos períodos de prueba no serán computadas como parte de las ganancias reales

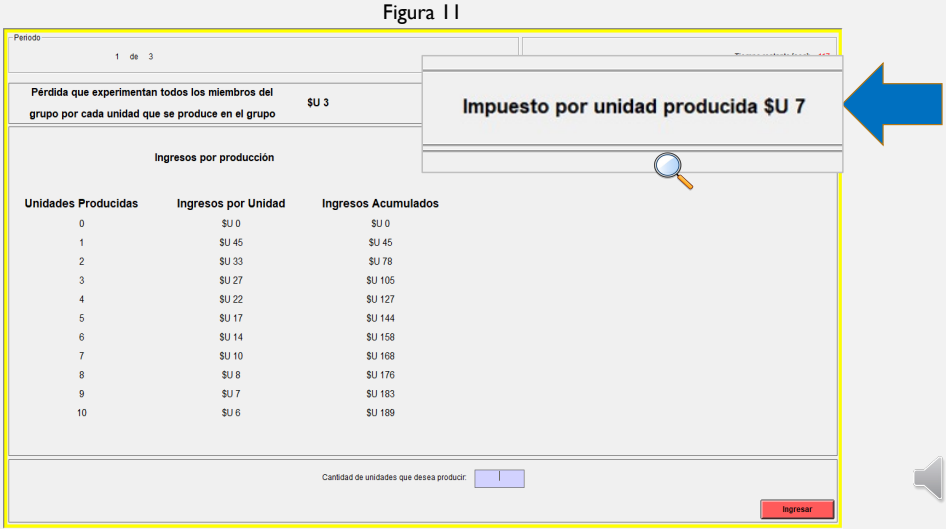
FIN DE EXPLICACIÓN DEL PRIMER EJERCICIO

SEGUNDO EJERCICIO

Usted y el resto del grupo deberán pagar un **impuesto por unidad producida**.

PANTALLA DE DECISIÓN SEGUNDO EJERCICIO

Figura 11



EJEMPLO

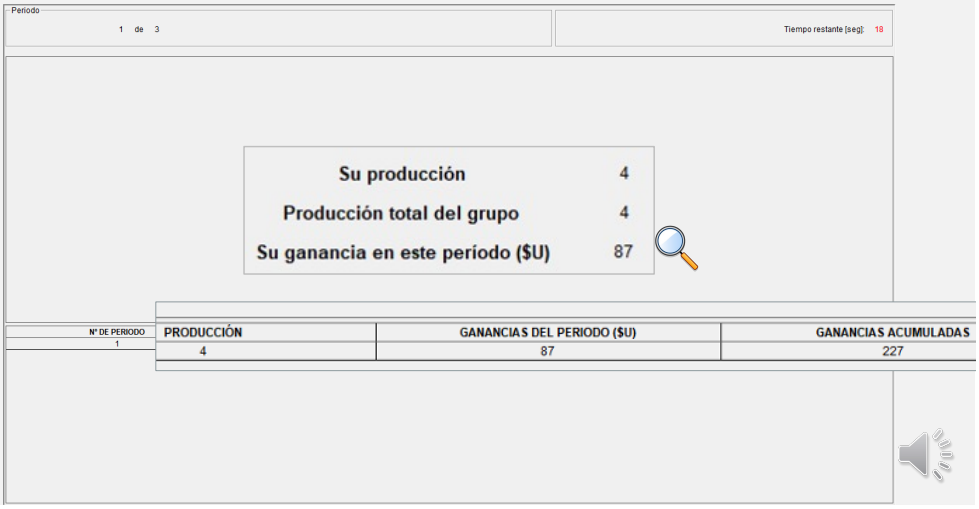
A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28				
2							
3							
Total							

EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2							
3							
Total							

GANANCIAS PRIMERA RONDA

Figura 12

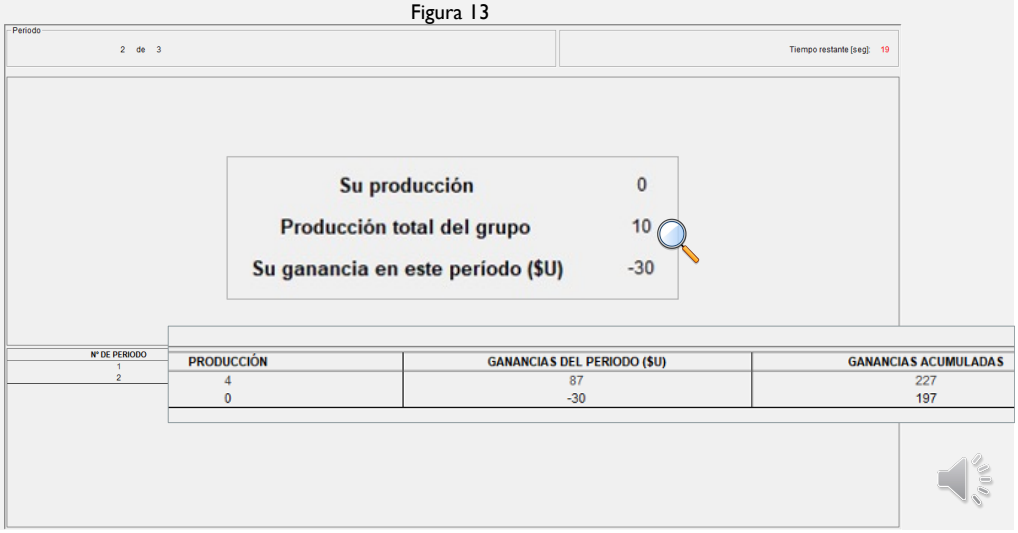


EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2	0	0	0	10	30	-30	197
3							
Total							

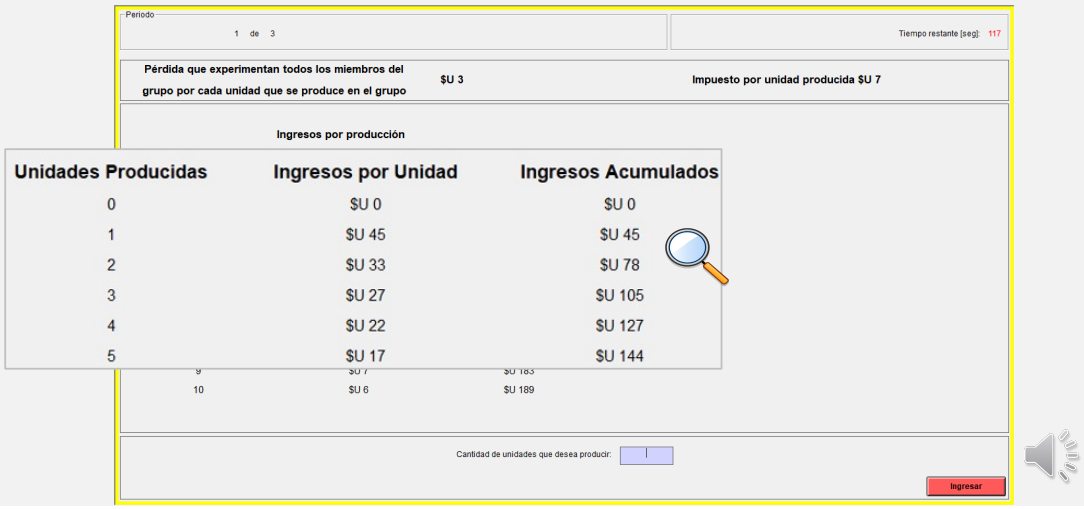
GANANCIAS SEGUNDA RONDA

Figura 13



EJEMPLO

Figura 14

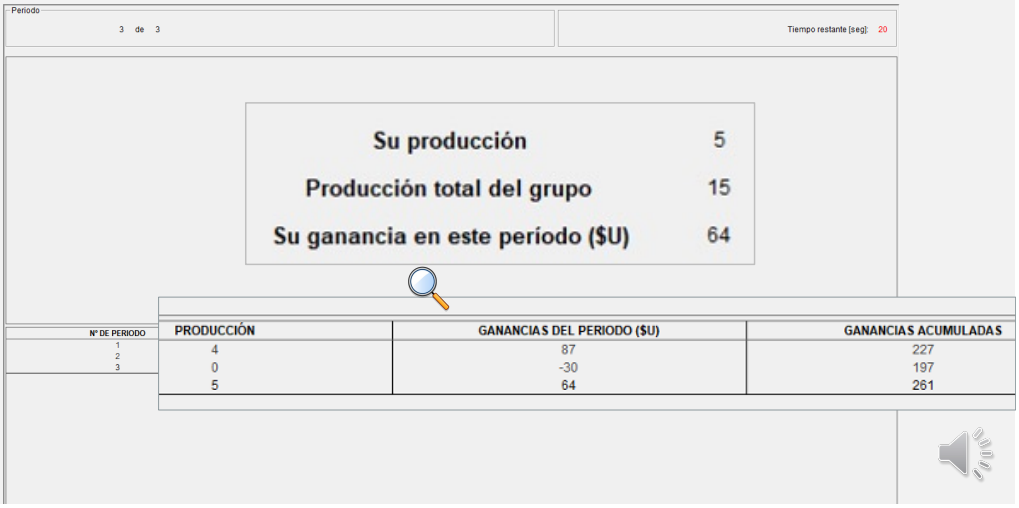


EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2	0	0	0	10	30	-30	197
3	5	144	35	15	45	64	261
Total							

GANANCIAS TERCERA RONDA

Figura 15



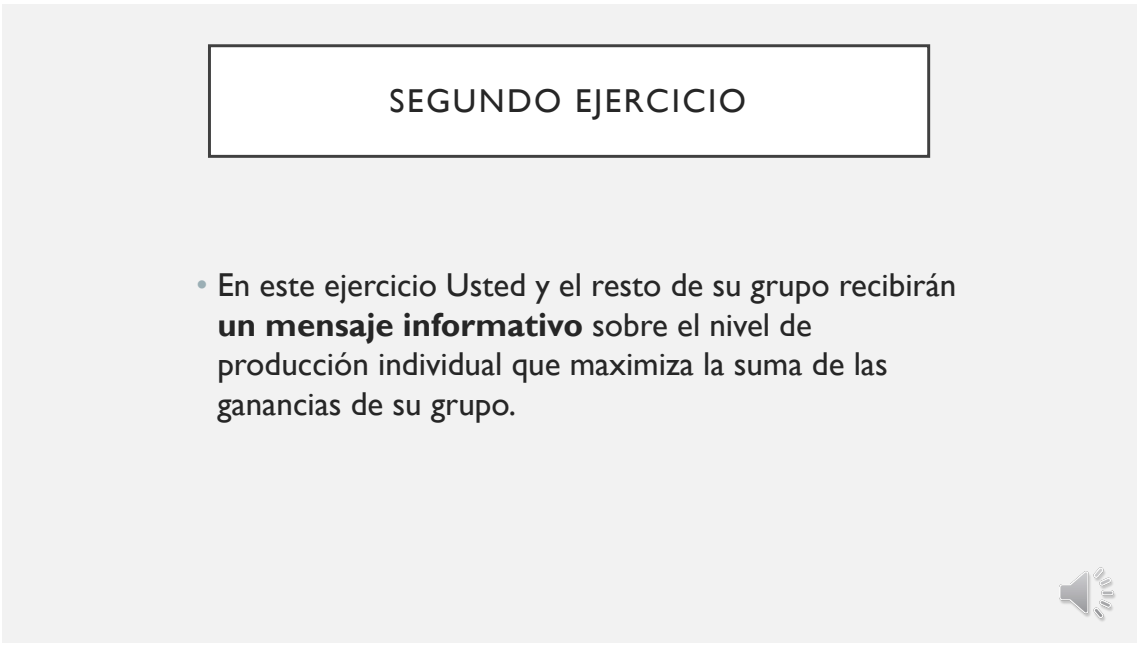
FIN DE EXPLICACIÓN DEL SEGUNDO EJERCICIO

CUESTIONARIO

- La última parte de la actividad de hoy consiste en responder un cuestionario.

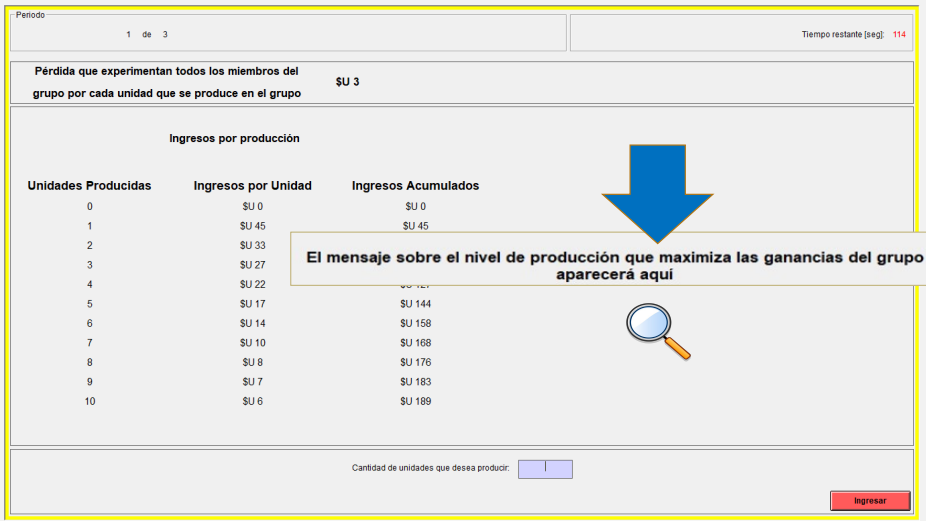
10.2 Baseline + Message

The introduction, the general description of the activity, the description of the first exercise and the questionnaire-completion of the activity are the same that in the Baseline + Tax instructions, what differs from the instructions is the description of the second exercise, which in this case deals with the message and is presented below.



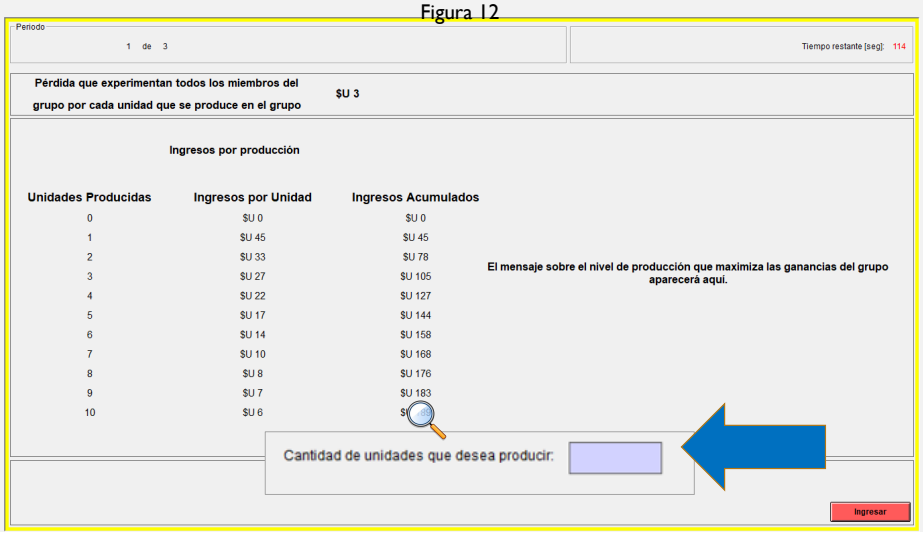
PANTALLA DE DECISIÓN SEGUNDO EJERCICIO

Figura 11



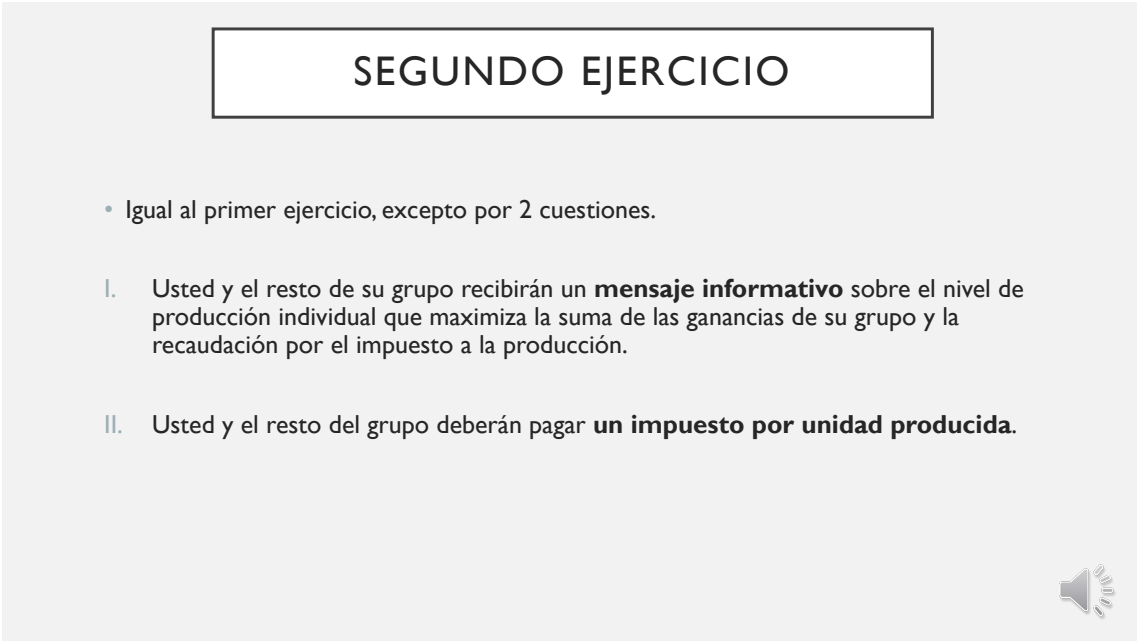
INGRESAR DECISIONES EN LA COMPUTADORA

Figura 12



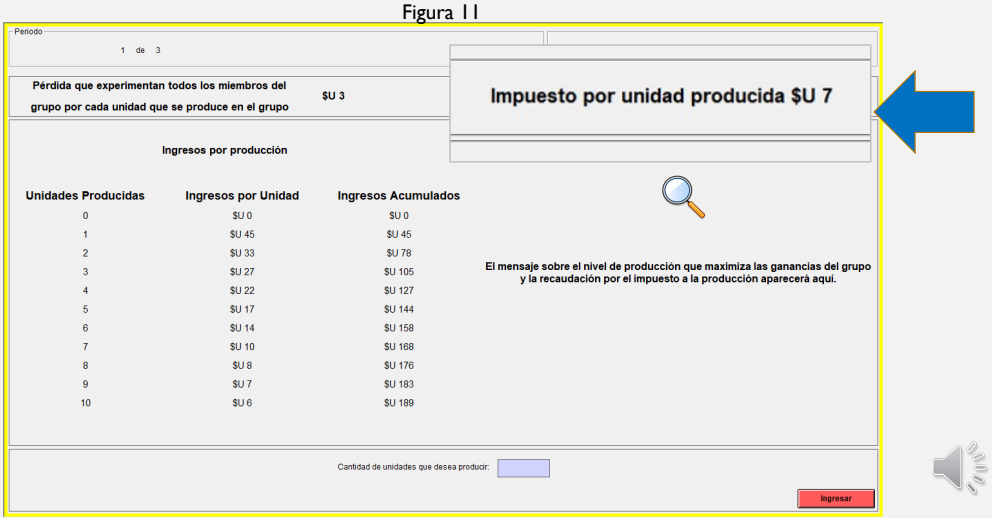
10.3 Baseline + Message+Tax

The introduction, the general description of the activity, the description of the first exercise and the questionnaire-completion of the activity are the same that in the Baseline + Tax instructions, what differs from the instructions is the description of the second exercise, which in this case deals with the message and the tax as a whole and is presented below.



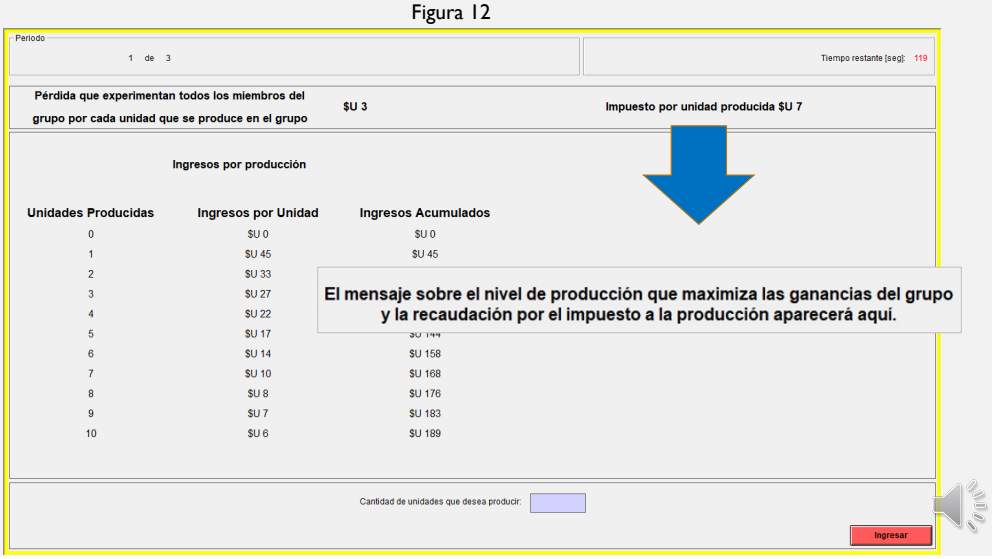
PANTALLA DE DECISIÓN SEGUNDO EJERCICIO

Figura 11



MENSAJE INFORMATIVO

Figura 12



EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28				
2							
3							
Total							

EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2							
3							
Total							

EJEMPLO

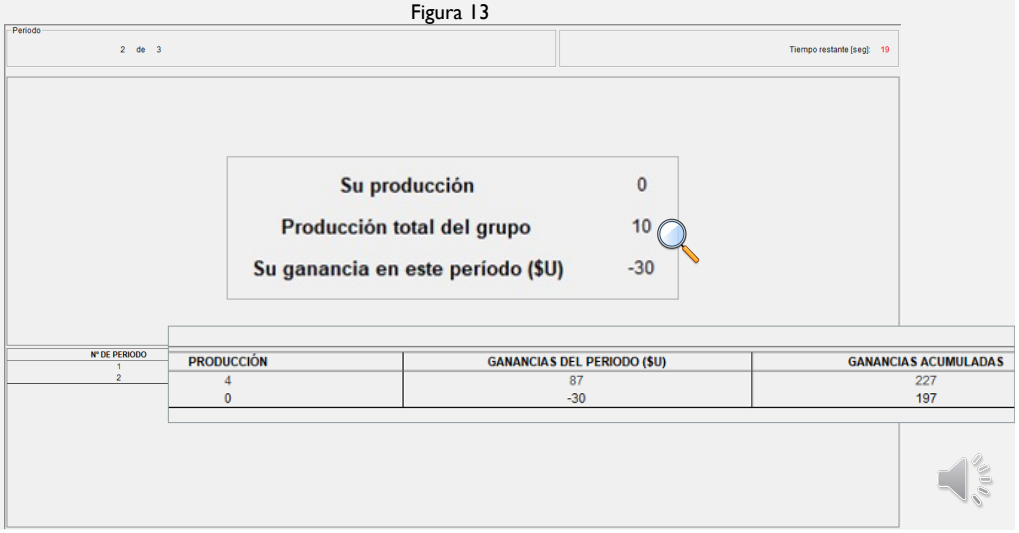
A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2							
3							
Total							

EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2	0	0	0	10	30	-30	197
3							
Total							

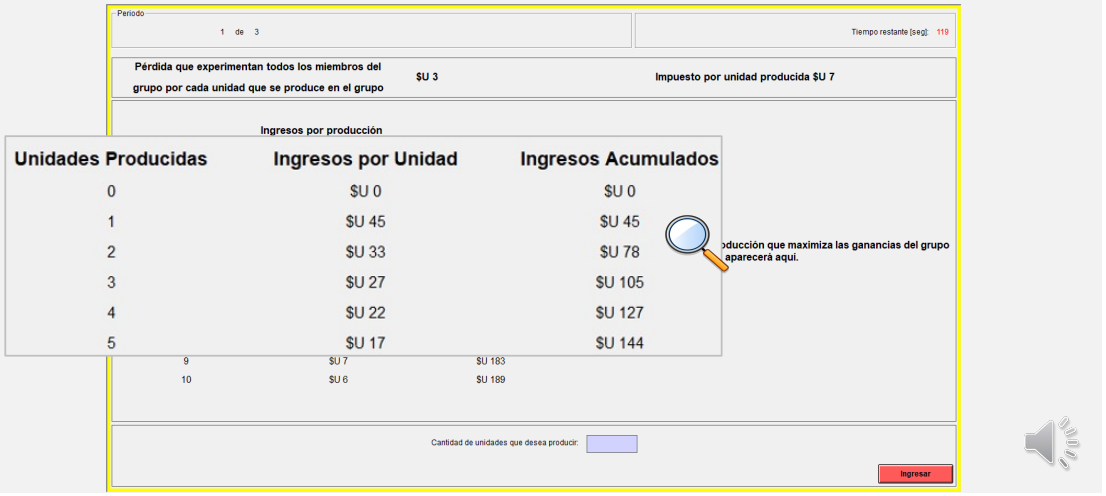
GANANCIAS SEGUNDA RONDA

Figura 13



EJEMPLO

Figura 14

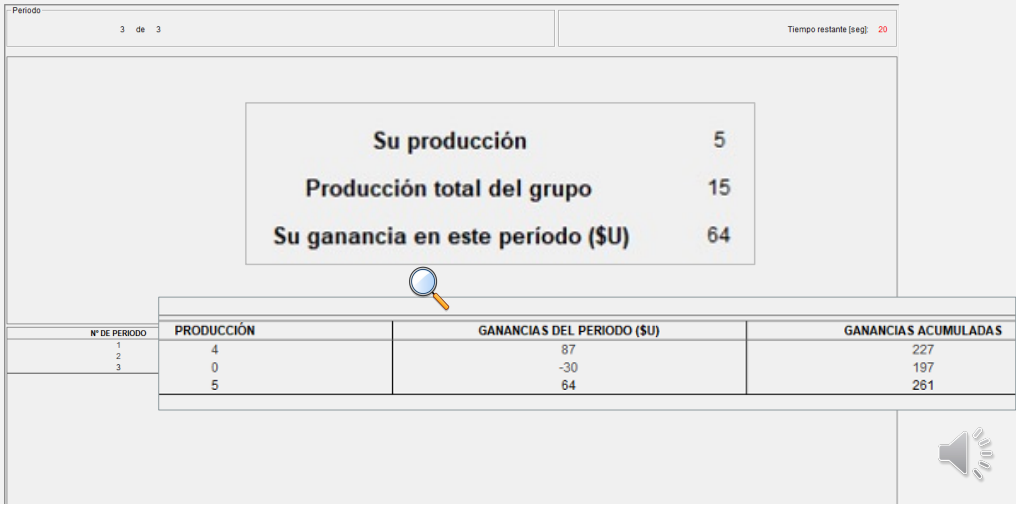


EJEMPLO

A	B	C	D	E	F	G	H
Ronda	Mi Nivel de Producción (Mi decisión) (0 al 10)	Mi ingreso por producción	Mis pagos de impuesto (B*$U7)	Total Producido por el Grupo	Mi Pérdida por producción del grupo (E*$U3)	Mis ganancias del periodo (C-D-F)	Mis ganancias netas finales (Incluye el capital de trabajo de $140)
1	4	127	28	4	12	87	227
2	0	0	0	10	30	-30	197
3	5	144	35	15	45	64	261
Total							

GANANCIAS TERCERA RONDA

Figura 15



11 Appendix C: Questionnaire

Post-experimental questionnaire

Description:

The following survey aims to collect information from the participants. The survey is anonymous. The information will be treated confidentially and for academic purposes only. Please answer the following questions:

1. Age (in completed years)

2. Gender

Man	
Woman	
Other	
I prefer not to answer	

3. Identify the University or Institution of higher education in which you are currently studying:
   a. Universidad de Montevideo
   b. Universidad ORT
   c. Universidad Católica
   d. Universidad de la República
   e. Universidad de la Empresa
   f. Universidad Tecnológica (UTEC)
   g. Universidad CLAEH
   h. Other: _____
4. What year of the degree are you studying?
   a. First
   b. Second
   c. Third
   d. Fourth
   e. Fifth
   f. Sixth

5. What career / academic program are you currently studying?
   a. Economy
   b. Business Administration
   c. Public accountant
   d. Engineering
   e. Law
   f. Biology
   g. Medicine
   h. Architecture
   i. Veterinary
   j. Agronomy
   k. Art
   l. Politic science
   m. Media Science
   n. Psychology
   o. Others: _____
6. On a scale of 1 to 10, where 1 represents as far to the left as possible and 10 as far to the right as possible, in what political position would you place yourself?

Más izquierda	1	2	3	4	5	6	7	8	9	10	Más Derecha

7. On a scale of 0 to 6, where 0 indicates that you are not willing to take risks and 6 indicates that you are willing to take any type of risk, where would you place your willingness to take risks

Not willing to take risks	0	1	2	3	4	5	6	You are willing to take any kind of risk

8. On a scale of 0 to 6, where 0 indicates that no one can be trusted and 6 indicates that anyone can be trusted completely, where would you place your level of trust in people?

Don't trust anyone	0	1	2	3	4	5	6	Trust totally anyone

9. Indicate in which of the following ranges the monthly income of your household falls, adding all the income of the people who contribute to the income of your household. Please include wages, pensions or social security, child support, allowances, business income or deposits, and any other income. (Check indicated range). If you are not sure, give us your best estimate.

1. Less than $U 29.500 __
2. Between $U 29.501 y $U 38.000 __
3. Between $U 38.001 - $U 45.900 __
4. Between $U 45.901 - $U 53.600 __
5. Between $U 53.601 - $U 62.200 __
6. Between $U 62.201 - $U 72.000 __
7. Between $U 72.001 - $U 83.300 __
8. Between $U 83.301 - $U 99.000 __
9. Between $U 99.001 - $U 121.400 __
10. Between $U 121.401 - $U 163.000 __
11. More than $U 163.000 __

10. How many people make up the family group that lives in your home (including you)? ___

11. Are you a believer/ profess any religion?
Yes __ No __

12. If you have answered "Yes" in the previous question, on a scale of 1 to 7, what is the intensity of your religious faith?

Minimal faith	1	2	3	4	5	6	7	Maximum faith

Indicate how often do you:

13. Use single-use bags for your purchases.

never seldom sometimes often always

14. Buy a drink in a non-returnable container.

never seldom sometimes often always

15. Bring empty bottles to a recycling bin.

never seldom sometimes often always

16. Use public transportation, ride a bike, or walk.

never seldom sometimes often always

17. On a scale of 0 to 6, where 0 indicates "strongly disagree" and 6 "strongly agree", indicate how much you agree with the following statement: "It is necessary to apply taxes to polluters".

Strongly disagree 0 1 2 3 4 5 6 Totally agree

18. Have given money to a charity.

never seldom sometimes often always

19. Have donated goods or clothes to a charity.

never seldom sometimes often always

20. Have done volunteer work for a charity.

never seldom sometimes often always

21. Have donated blood.

never seldom sometimes often always

22. Have offered to help a person with a disability or elderly stranger across a street.

never seldom sometimes often always

23. Have offered your seat to a stranger who was standing.

never seldom sometimes often always

24. In the activity instructions you were informed that the experimenter will not be able to know how many units you decided to produce during the exercises. Was it clear to you?

Yes __ No __

25. In general, did you take into account the effect of your decision on the earnings of the other members of your group when deciding how much to produce?

Yes __ No __

26. What is the level that you think a person should produce in each of the rounds of the activities in which he has just participated?

1 2 3 4 5 6 7 8 9 10

27. If you chose to produce a larger quantity than indicated in your previous answer in some of the rounds, did you feel guilty when making that decision?

Yes __ No __

28. If you have answered "Yes" in the previous question, on a scale of 1 to 5, how guilty did you feel?

No guilt 1 2 3 4 5 Lot of guilt

NOTE: ONLY FOR SUBJECTS WHO PLAYED THE TREATMENT OF NUDGE

29. Did the informational message shown to you in the experiment affect what you thought was correct to produce?

Yes__ No__

30. Did receiving the message motivate you to change your decisions?

Yes__ No__

31. In case the message did not motivate you to change your decisions, why?

_____

_____

_____

_____

_____

32. What percentage of people in the room do you think reduced their production level by showing the informational message? Give us your best estimate.

_____ %

33. Do you approve of the use of messages by a government to inform the population about the environmental effects of our behaviors?

Yes__ No__

34. Do you approve of the use of messages by a government to inform the population about the social effects (on other people) of our behavior?

Yes__ No__

12 Appendix D: Descriptive statistics - Questionnaire

The last stage of the experiment consisted of answering some questions about the participant and his beliefs. An example of this questionnaire can be found in the annexes.

Table 9: Descriptive statistics for sociodemographic questions

	Variable	Subjects
Gender	F	46%
	M	54%
Age	17-18	8%
	19	25%
	20	20%
	21	13%
	22	7%
	23-30	21%
	31-49	6%
Home Income	Less than $38000	12%
	Between $38000 and $53600	16%
	Between $53600 and $72000	17%
	Between $72000 and $99000	15%
	Between $99000 and $163000	18%
	More than $163000	22%
University	UDELAR	74%
	Universidad de Montevideo	21%
	Universidad Católica	1%
	ORT	1%
	Other	3%
University Degree	Economics	57%
	Business Administration	7%
	Engineering	6%
	Public Accountant	6%
	Medicine	3%
	Media Science	2%
	Law	2%
	Statistics	2%
	Psychology	2%
	Data Science	1%
	Finance	1%
	International Business	
	other	11%
Political Preferences	Right	26%
	Left	21%

Middle	53%	

Table 10: Descriptive statistics for environmental questions

		Mean (q)	%
"It is necessary to apply taxes to polluters"	Disagree (0,1 or 2)	7,22	6%
	Neutral (3)	7,31	9%
	Agree (4, 5 or 6)	6,99	85%
"Indicate how often do you use single-use bags for your purchases"	Never	7,28	13%
	Seldom	6,95	53%
	Sometimes	7,11	16%
	Often	6,83	13%
	Always	7,53	5%
"Indicate how often do you buy a drink in a non-returnable container"	Never	7,38	6%
	Seldom	6,97	21%
	Sometimes	6,96	28%
	Often	6,87	31%
	Always	7,51	14%
"Indicate how often do you bring empty bottles to the recycling bins"	Never	6,97	29%
	Seldom	6,70	27%
	Sometimes	7,19	17%
	Often	7,25	18%
	Always	7,41	11%
"Do you approve of the use of messages by a government to inform the population about the environmental effects of our behaviors?"	Yes		94%
	No		6%

Table 11: Descriptive statistics for pro-social questions

		Mean (q)	%
"Indicate how often do you use public transportation, ride a bike, or walk."	Never	8,60	1%
	Seldom	7,54	3%
	Sometimes	7,83	10%
	Often	7,23	31%
	Always	6,71	55%
"Indicate how often do you have given money to a charity"	Never	7,45	11%
	Seldom	7,10	25%
	Sometimes	7,01	49%
	Often	6,69	13%
	Always	6,82	2%
"Indicate how often do you have donated goods or clothes to a charity"	Never	6,80	3%
	Seldom	6,98	15%
	Sometimes	7,15	38%
	Often	6,95	34%
	Always	6,96	10%
"Indicate how often do you have done volunteer work for a charity"	Never	6,86	29%
	Seldom	7,13	24%
	Sometimes	7,02	28%
	Often	7,41	14%
	Always	6,59	5%
"Indicate how often do you have donated blood"	Never	7,16	66%
	Seldom	7,01	12%
	Sometimes	6,85	11%
	Often	6,85	6%
	Always	5,98	5%
"Indicate how often do you have offered to help a person with a disability or elderly stranger across a street."	Never	7,94	7%
	Seldom	7,13	21%
	Sometimes	7,18	33%
	Often	6,81	23%
	Always	6,53	16%
"Indicate how often do you have offered your seat to a stranger who was standing"	Never	7,53	2%
	Seldom	7,19	8%
	Sometimes	7,04	30%
	Often	7,16	40%
	Always	6,67	20%

Table 12: Descriptive statistics for non-environmental – non prosocial questions

		Mean (q)	%
"On a scale of 0 to 6, where 0 indicates that you are not willing to take risks and 6 indicates that you are willing to take any type of risk, where would you place your willingness to take risks"	None risk (0)	-	0%
	Less risk (1-2)	6,89	21%
	Neutral risk (3-4)	7,16	62%
	More risk (5-6)	6,78	18%
"On a scale of 0 to 6, where 0 indicates that no one can be trusted and 6 indicates that anyone can be trusted completely, where would you place your level of trust in people?"	Less confidence (0-2)	6,97	42%
	Neutral confidence (3-4)	7,19	50%
	More confidence (5-6)	6,38	8%
"Are you a believer/ profess any religion?"	No	6,91	63%
	Yes	7,24	37%
In general, did you take into account the effect of your decision on the earnings of the other members of your group when deciding how much to produce?	No	7.94	26%
	Yes	6.74	74%
What is the level that you think a person should produce in each of the rounds of the activities in which he has just participated?	1	8.3	1%
	2	-	0,0%
	3	5.3	1%
	4	5.6	10%
	5	6.9	50,5%
	6	7.2	17,5%
	7	7.8	9,5%
	8	7.7	5,5%
	9	8.8	1,5%
	10	8.6	3,5%
If you chose to produce more than the amount you indicated you believed was correct in your previous response, did you feel guilty?	No	7.3	62%
	I did not choose to produce a larger amount	5.6	10%
	Yes	6.9	28%
Did the informational message shown to you in the experiment affect what you thought was correct to produce?	No	7.58	30,7%
	Yes	6,96	69,3%
Did receiving the message motivate you to change your decisions?	No	7.78	36%
	Yes	6.79	64%

Of the 46 subjects who answered that the message did not motivate them to change their decision, 38 gave their reasons, which are broadly divided into the following reasons:

Short answer	%
Personal interests	47,37%
He changed his decision based on what he saw the group produce	36,84%
Choose 5	2,63%
He already knew that the optimal amount was 5	5,26%
Unfinished answer	7,89%

How many participants do you think reduced their level of production after seeing the informative message?	
	%
0-10%	34%
11-20%	5%
21-30%	6%
31-40%	4%
41-50%	10%
51-60%	6%
61-70%	13%
71-80%	16%
81-90%	4%
91-100%	2%

13 Appendix E: Parametric test with controls

13.1 Household Income

Table 13. Linear Random Effect regressions results controlling for self-reported household income

Dependent variable: Level of production	(1)
20-40%	-0.16
	(0.35)
40-60%	-0.34
	(0.34)
60-80%	0.08
	(0.35)
80-100%	0.70**
	(0.30)
Message (M)	-0.50***
	(0.13)
Tax (T)	-1.00***
	(0.13)
Message and Tax (M+T)	-0.97***
	(0.13)
Constant	7.31***
	(0.23)
Chi-squared	140.82
N	2000.00

Notes: This table presents estimates from random-effect GLS regressions, estimated on a panel data structure grouped in two levels: a concatenation of session and subject as the group identifier, and round as the time identifier. In each column, the dependent variable is the production level chosen by subject i of group g, in round t. There is one independent variable for each treatment, recalling that a subject only participated in one of them. Each treatment is represented by a dummy variable, which is 1 if subject i has faced that treatment, and 0 in other case. We also include a categorical variable to control for the several household incomes brackets a subject i can be part of. In total, we have 5 categories for household income, being the lowest one the chosen to take out of the regression and avoid multicollinearity. The constant reflects the average production level in baseline for subjects in the lowest household income bracket, as well as the coefficient related to the treatment effects, which are the effects only for these subjects. Significance levels are indicated as * 0.10 ** 0.05 *** 0.01

13.2 University

Table 14. Linear Random Effect regressions results controlling for university.

Dependent variable: Level of production	(1)
Public University (PU)	-0.79***
	(0.25)
Message (M)	-0.46**
	(0.21)
Tax (T)	-1.01**
	(0.46)
Message and Tax (M+T)	-1.01***
	(0.23)
PU*M	-0.06
	(0.27)
PU*T	0.03
	(0.48)
PU*(M+T)	0.02
	(0.28)
Constant	8.03***
	(0.22)
Chi-squared	137.98
N	2000.00

Notes: This table presents estimates from random-effect GLS regressions, estimated on a panel data structure grouped in two levels: a concatenation of session and subject as the group identifier, and round as the time identifier. In each column, the dependent variable is the production level chosen by subject i of group g, in round t. There is one independent variable for each treatment, recalling that a subject only participated in one of them. Each treatment is represented by a dummy variable, which is 1 if subject i has faced that treatment, and 0 in other case. We also include a dummy variable called Public University (PU), which takes values 1 if subject i assists to a public university and 0 otherwise. We also include the interaction of this variable with the treatment ones, to check if there are differential effects of each instrument over production depending on university. The constant reflects the average production level in baseline for subjects which assist to a private university. Significance levels are indicated as * 0.10 ** 0.05 *** 0.01

14 Appendix F: Non-Parametric Tests

Statistical significance of the test

Treatment	Test		
	Median Test	Ranksum	Ttest
B vs M	+	+	+
B vs T	+	+	+
B vs M+T	+	+	+
M vs T	-	+	+
M vs M+T	+	+	+
T vs M+T	-	-	-

14.1 Baseline vs Message

Non-parametric tests Baseline vs Message

	Baseline	Message	P-value
Median Test	.	.	.0019
T-test (Mean)	7.446	7.1354	.0343
Ranksum Test	.	.	.0055

14.2 Baseline vs Tax

	Baseline	Tax	P-value
Median Test	.	.	0.000
T-test (Mean)	7.446	6.3086	0.000
Ranksum Test	.	.	0.000

14.3 Baseline vs Message and Tax

Non-parametric tests Baseline vs Message and Tax

	Baseline	Message and Tax	P-value
Median Test	.	.	0
T-test (Mean)	7.446	6.4308	0
Ranksum Test	.	.	0

14.4 Message vs Tax

Non-parametric tests Message vs Tax

	Message	Tax	P-value
Median Test	.	.	.2419
T-test (Mean)	7.1354	6.3086	0
Ranksum Test	.	.	.0001

14.5 Message vs Message and Tax

Non-parametric tests Message vs Message and Tax

	Message	Message and Tax	P-value
Median Test	.	.	.0118
T-test (Mean)	7.1354	6.4308	0
Ranksum Test	.	.	.0001

14.6 Tax vs Message and Tax

Non-parametric tests Tax vs Message and Tax

	Message and Tax	Tax	P-value
Median Test	.	.	.1625
T-test (Mean)	6.4308	6.3086	.4469
Ranksum Test	.	.	.9445