// NO.21-010 | 02/2021

DISCUSSION PAPER

// MARIUS ALT AND CARLO GALLIER

Incentives and Intertemporal Behavioral Spillovers: A Two-Period Experiment on Charitable Giving

Leibniz Association

ZEW

Incentives and intertemporal behavioral spillovers: a two-period experiment on charitable giving

Marius Alt[1,2]*, Carlo Gallier[1]

[1]ZEW - Leibniz Centre for European Economic Research

[2]University of Copenhagen, Dept. of Food and Resource Economics.

This version: January, 2021

Abstract: We test whether and, if so, how incentives to promote pro-social behavior affect the extent to which it spills over to subsequent charitable giving. To do so, we conduct a two-period artefactual field experiment to study repeated donation decisions of more than 700 participants. We vary how participants' first pro-social behavior is incentivized by a wide range of fundraising interventions ranging from soft to hard paternalism. Our design allows us to decompose spillover effects into a pure spillover effect, which identifies the impact of previous pro-social behavior on subsequent donation decisions and a crowding effect, which captures the extent to which the spillover effects are affected by the incentives exerted on the previous pro-social behavior. We find evidence for negative spillover effects. Participants donate less if they completed a pro-social task prior to the donation decision. Most importantly, we find that the spillover effects depend on how the initial pro-social behavior has been incentivized. Especially participants who are incentivized to donate through social comparisons are more willing to give to charity thereafter compared to participants whose initial pro-social behavior is incentivized by monetary rewards. The variations in spillover effects are driven by participants' perceived external pressure in the first pro-social decision.

JEL classification: C91, C93, D01, H41, D04

Keywords: Charitable giving, Social preferences, Experimental economics, Behavioral spillovers, Policy-making, Economic incentives

Acknowledgements: My thanks to Wolfgang Habla, Martin Kesternich, and Toke Reinholt Fosgaard and seminar and conferences audiences in SPI 2020, EAERE 2020, CNEE-workshop, 5th EEFE-workshop and AUROE-Nachwuchsworkshop for constructive and very helpful comments. We also thank Lisa Jeuck, Marcel Pommer, Ekin Demirtas and Niklas Schoch for their research assistance. The project was made possible through funding from the German Federal Ministry of Education and Research (FKZ 01UT1701B). This experiment has been registered and reviewed by the institutional review board of the Ethical Committee of SCIENCE and SUND at the University of Copenhagen.

*Correspondence: marius.alt@zew.de; ZEW, L7,1, 68161 Mannheim, Germany.

1 Introduction

Charitable organizations can choose from a wide range of different fundraising interventions to attract new donors and create a habit of giving. A large literature in economics has intensively studied measures such as altruistic appeals (e.g., List et al. 2019; Fielding et al. 2020), social comparisons (e.g., Frey and Meier 2004; Croson and Shang 2011), and rewards (e.g., Carpenter and Matthews 2017; Lacetera et al. 2014), which have proven to be effective in attracting new donors and increasing charitable giving.[1] These studies typically focus on the instantaneous effects of fundraising interventions, i.e., how much is raised by the charity at the time of the intervention. However, many individuals donate on a regular basis. In Germany, for instance, the number of individual donors has been decreasing for years, but the frequency of giving of those who donate increases. On average, donors give more than six times per year to charity (GfK, 2019). This emphasizes that it is not only important to understand the instantaneous effects of fundraising interventions, but also to take into account how decisions spill over to donations at a later point in time. It is exactly this question of how interventions influences behavioral spillover effects between pro-social decisions that we address in this paper.

How do previously applied mechanism to attract donors affect the extent to which their behavior spills over to subsequent decisions in a charitable appeal? Although behavioral spillover effects, defined as the extent to which individuals‘ pro-social behavior in the past affects their current pro-social decisions, have been intensively studied (e.g., Gee and Meer 2019; Scharf et al. 2017; Sass et al. 2015; Schmitz 2019; Adena and Huck 2019), little is known about how the extent to which pro-social behavior spills over depends on the intervention applied in the initial decision. This strikes us as an important but understudied questions as the extent to which pro-social behavior spills over determines whether the effect of an intervention is persistent. The experimental literature shows that nudges tend to have rather small effects on subsequent behavior, being larger when they involve social pressure or non-monetary rewards (e.g., Gallus 2017; Ghesla et al. 2019; Kesternich et al. 2019; Adena and Huck 2020). In contrast, monetary incentives, in the form of rewards or gifts for pro-social behavior, show relatively large effects on subsequent giving (e.g., Landry et al. 2006; Gallier et al. 2017; Krieg and Samek 2017). In line with this, (d'Adda et al., 2017) find that minimum contribution levels to public goods substantially affect subsequent pro-social behavior. To summarize, there is evidence that incentive schemes to behave pro-social in a given situation have an effect on subsequent

[1] Other related work focuses, for instance, on matching donations (e.g., Karlan and List 2007, Kesternich et al. 2019) or door to door fundraising campaigns (e.g., DellaVigna et al. 2013).

pro-social behavior.

In the abovementioned studies, however, it remains unclear what is driving the extent to which pro-social behavior spills over, given it has been incentivized. There are at least two potentially different effects that could affect subsequent pro-social behavior, which cannot be disentangled in the previous literature. At first, incentive schemes that exert some sort of external pressure on individuals to behave pro-social in a given situation can, of course, instantaneously affect their decision at a current point in time. This incentive induced behavioral change can directly spill over and, thereby, affect individuals pro-social behavior at a later point in time. Second, the incentive schemes can also have a more indirect effect on individuals' subsequent behavior by mediating the extent to which pro-social decisions spill over.

We contribute to the literature by investigating whether and, if so, how incentives to promote pro-social behavior affect the extent to which donations spill over to subsequent charitable giving. Most importantly, our experimental design allows us to disentangle spillover effects and clearly identify to which extent they depend on the applied fundraising intervention, while controlling for the induced change in initial pro-social behavior.

We begin the paper by developing a conceptual framework that extents common models of impure altruism (e.g., Andreoni 1989, 1990; Landry et al. 2006, 2010) to a situation in which agents face repeated donation decisions. In doing so, we outline whether and, if so, how incentive schemes that exert external pressure on agents' initial donation decision affects their subsequent pro-social behavior. We derive that spillover effects - in general - are negative such that agents' subsequent donations are decreasing with their initial pro-social engagement. Furthermore, by building upon the psychological concept of 'diagnosticity' (Kelley 1973; Bem 1972), we derive that additional incentives which exert pressure on agents' initial donation decision have the potential to crowd-out the spillover levels, e.g., mitigate the extend to which donation decisions are reduced by past pro-social behavior.

Guided by our theoretical considerations, we implemented our experimental design into a field setting of repeated donation decisions.[2] The experiment is embedded in a survey on a topic unrelated to pro-social behavior. All in all, we observe the repeated donation decisions of more than 700 participants. Within the survey, we implement two pro-social behavior tasks, one at the beginning and one at the end of the experiment. The second task consists of the option to donate a share of the participation fee to charity. This task is held constant across all treatments. The first task has the form of a real ef-

[2]Following the taxonomy of Harrison and List (2004), we rely on a framed field experiment to observe participants' pro-social behavior in an natural environment while maintaining a high level of control.

fort task, which we systematically vary across treatment. Our treatments differ across two dimensions. First, we alter whether the performance in the real-effort task has a pro-social consequence. Second, if the real-effort task contains a pro-social component, we vary whether participants' performance is additionally incentivized. In our *NoSocial* treatment, participants' initial decision has no pro-social consequences. In *Baseline*, the initial decision has pro-social consequences that are, however, not additionally incentivized. In our *Nudge*, *Monetary Incentive*, and *Punishment* treatments, participants' initial decisions have pro-social consequences that are additionally incentivized via either social comparisons, monetary incentives, or a time punishment, respectively. Based on the experimental design, we are able to decompose the spillover effects into the pure effect of acting pro-socially in a previous decision and the effect of how this initial decision has been additionally incentivized.

Our central findings are briefly summarized. First, we find a negative pure spillover effect. This means that participants, who completed a pro-social task that has not been additionally incentivzed right before the donation decision, donate significantly less than participants who did not complete a pro-social task previously. Second, the spillover effects depend on how participants' initial pro-social behavior has been additionally incentivized. While we find significant spillover effects in our *Monetary Incentive* and *Punishment* treatments, the spillover effect in *Nudge* does not reach significance at conventional levels of statistical inference. In order to investigate why participants in *Monetary Incentive* and *Punishment* show relatively strong spillover effects compared to participants in our *Nudge* treatment, we extend our analysis and investigate how participants perceive the external pressure exerted by our treatment interventions and how the level of perceived pressure affects the spillover level. We find that the level of perceived pressure is comparably low in our *Nudge* treatment. Not only lower than in *Monetary Incentive* and *Punishment*, but also lower than in our *Baseline* treatment. Furthermore, we find that participants' level of perceived pressure negatively affects the degree to which their initial pro-social behavior spills over to the subsequent donation decision. All this helps to explain the relatively weak spillover effects in our *Nudge* treatment, contributing to the understanding of the mechanisms behind the observed spillover effects when pro-social behavior is incentivized.

The rest of the paper is organized as follows. The next section provides a discussion of the closely related literature. Section 3 describes our theoretical considerations that guide our experimental design and provide structure to our analysis. Section 4 describes the experimental design and treatments, derives our hypotheses, and provides the experimental procedures. The results are presented in Section 5 and Section 6 concludes.

2 Previous Work

2.1 Behavioral spillovers

As closely related to behavioral spillovers, the concept of moral balancing originates from social psychological research (Funder and Colvin 1991; Monin and Miller 2001) and describes that past good deeds could legitimizes individuals to engage in behaviors that are bad, immoral, or unethical, or vice versa (e.g., Merritt et al. 2010). In economic studies, behavioral spillovers of pro-social behavior become relevant when assessing the substitutability of donation behavior with respect to time and with respect to charity space (Scharf et al., 2017). Spillovers within charity spaces refers to decisions of donors, who choose within a broad range of charities engaging in different pro-social causes. It investigates the influence of a donation to a certain charity on the probability to donate to other charities (Krieg and Samek, 2017). Behavioral spillovers with respect to time analyze whether a charitable contribution in a certain period represents a substitute or a complement compared to donation decisions in other time periods. Since the latter is also subject to the investigation in this study, we provide a brief overview of former studies on this topic. If donations are complements from a cross-periodical perspective, we refer to them as positive spillovers, as donations in a certain period positively influence the likelihood to give in another period of time. Donations, which are substitutes with respect to time, are described as negative spillovers, as a contribution to charity reduces the probability of contributing in the other periods. Negative spillovers are observed in the field by Adena and Huck (2019), who investigate the effect of announcing another upcoming donation request on current donation behavior. The results show that donors reduce charitable contributions on the basis of such announcements. Evidence from the laboratory on repeated giving in dictator games also show a tendency to negative behavioral spillovers across periods. In Sass et al. (2015) decisions are made repeatedly on a weekly basis over the course of four weeks within the context of a laboratory study. The authors provide evidence that particularly in weeks two and three reductions in giving are noticeable. Schmitz (2019) also observes negative behavioral spillovers, as he compares differences in donation amounts given the second choice takes place on the same day or with a delay of a week. He finds a slight fading of negative behavioral spillovers given an increase in the time gap between decisions. Mixed evidence on the cross-periodical dependencies of giving are found by Gee and Meer (2019), who screen the literature applying a broader context of pro-social behavior and spillovers. They establish the concept of 'altruism budget', taking into account also other components of pro-social behavior like in-kind gifts or volunteering. Positive spillovers of pro-social

behavior can be observed, for instance, in Baca-Motes et al. (2013), who find guests committing to reuse their towels to be less likely to leave on lights in an unattended room. Further studies on positive behavioral spillovers emphasize the importance of identifying with the cause (Clot et al. 2016; Effron and Monin 2010; Meijers et al. 2014; Kang and Glassman 2010; Brañas-Garza et al. 2011). A comprehensive assessment of the occurence of spillover effects is provided in Blanken et al. (2015). In a meta-analysis of 91 studies dealing with behavioral spillovers, the authors find a general tendency of negative behavioral spillovers and a large distribution across effect sizes, suggesting heterogeneity in behavioral spillover effects across individuals and settings.

2.2 Behavioral spillovers and external pressure on pro-social behavior

Due to the focus on economic policy assessment within the economic literature, a large strand of studies concerning behavioral spillovers implement an intervention to the initial pro-social task and analyze its effect on subsequent behaviors not directly targeted by the intervention Truelove et al. (2014). For instance, Landry et al. (2010) use a door-to-door fundraising campaign to investigated the effects of mechanism compared to nonmechanism incentives on long-run donation behavior. The results show that donors who are initially exposed to an economic mechanism were more prone to subsequently continue giving. In an online experiment, Gallus (2017) assesses the effect of awards to volunteers on Wikipedia, finding that the awards raise performance of the targeted group with a persisting effect over time. Similar to this, Kesternich et al. (2019) vary whether carbon offsetting behavior differs if the active pro-social choice is enforced or skippable, when purchasing a bus ticket via an online platform. Among others, they find persistent effects in the decisions of customers who were forced to take the active offsetting choice. Positive spillovers of intervention induced pro-social behavior are also found in the laboratory by Gallier et al. (2017), as participants tended to decide on larger subsequent donations given the presence of a lottery or a tax rebate in a previous dictator game. Contrasting evidence is provided by Adena and Huck (2020), who show adverse long-run effects of fundraising campaigns on an opera booking platform, as customers exposed to more insisting online fundraising in the past purchased fewer opera tickets online subsequently.

While the provided evidences show that interventions are likely to have an effect on subsequent giving, we introduce external pressure on donors as a possible mediator behind the behavioral spillovers of incentive induced pro-social behavior. According to the concept of 'diagnosticity' (Kelley 1973; Bem 1972), which originates from the attribu-

tion theory and self-perception theory, external pressure to perform an initial behavior decreases the respective impact on moral self-regard and affects not only subjects' initial decisions but also subsequent moral behavior. The reasoning is based on the psychological mechanism that external pressure to conduct a moral behavior leaves subjects ambiguous about their true intentions and thereby diminishes the opportunity to attribute the respective behavior as an own moral achievement. This diminishes possible effects on the moral self-regard of individuals (Mullen and Monin, 2016). Therefore, external pressure has the potential to mitigate cross-periodic dependencies of moral behavior. Khan and Dhar (2006), for instance, ask participants to imagine to have worked for community service over a period of six weeks, either voluntarily or under duress. In a subsequent task, significantly less participants state to choose a luxury good than a necessity good, if they have imagined to do community service under duress. In a closely related study, Clot et al. (2013) test whether paying participants to perform a pro-environmental task affects their willingness to donate to charity subsequently. They find that significantly more participants are willing to donate, if they previously have imagined to perform a pro-environmental behavior and getting paid for it, compared to those imagining a purely voluntary action. d'Adda et al. (2017) investigate the persistence effect of information nudges and rebates as well as a minimum contribution rule on subsequent behavior in either a dictator game or a prisoners dilemma game. They find stronger persistence effects of the latter two measures on subsequent behavior, given participants face the same game twice. The nudges had no effect on subsequent behavior. To summarize, there is evidence for a causal relation between external pressure and the occurrence of behavioral spillover effects.

Our work differs from the previous studies, in particular Khan and Dhar (2006), Clot et al. (2013), and d'Adda et al. (2017). In addition to our consequential design, which differs from Khan and Dhar (2006) and Clot et al. (2013), we fix moral behavior in the initial pro-social task in order to disentangle the effect of different incentives on subsequent behavior. d'Adda et al. (2017) do not hold moral behavior constant in the initial pro-social task. Therefore, they cannot disentangle the effect of different incentives on subsequent behavior, as behavior is confounded by the moral achievements from the initial pro-social task. In addition, our approach intends to diverge from the abstract setting of dictator games, using a real-effort task in a real life scenario.

3 Theoretical Considerations

To provide a framework for our experimental design and a structure to our analysis, we introduce the concept of behavioral spillovers into a model of charitable giving. To do so, we extend common models of impure altruism (e.g., Andreoni 1989, 1990; Landry et al. 2006, 2010) to a situation in which agents face repeated donation decisions and vary whether and, if so, how the first donation decision has been incentivized. By the means of our model, we identify the effect of incentivizing the initial donation decision on subsequent pro-social behavior. We demonstrate that incentives do not only affect agents' immediate public good contributions that spill over to the subsequent decision, but also affect the extent to which the initial pro-social behavior spills over.

The most simple case has two time periods, $t \in \{1, 2\}$. In each period, agent $i \in \{1, \ldots, n\}$, can decide how to allocate her endowment, $m_{i,t}$, on private consumption, $c_{i,t}$, or to a public good, D_t, using own contributions, $d_{i,t}$. Agent i derives utility from private consumption, $c_{i,t}$, the overall public good provision level, $D_t = \sum_{j=1}^{n} d_{j,t}$, and from her own contribution to the public good, $d_{i,t}$. Furthermore, she faces the budget constraint $c_{i,t} + d_{i,t} \leq m_{i,t}$. In $t = 2$, agent i derives utility according to the following additively separable utility function:

$$U_{i,2} = u(c_{i,2}) + h(D_2) + \alpha_i(\gamma(k_1)d_{i,1}(k_1))f(d_{i,2}).$$

The functions $u(\cdot)$, $h(\cdot)$, and $f(\cdot)$ are strictly increasing and concave. $u(\cdot)$ captures the utility from private consumption and $h(\cdot)$ the utility from the the overall public good provision level. $f(\cdot)$ describes the utility from contributing to the public good and $\alpha_i(\cdot)$ is the spillover parameter, which links public good contributions across periods. This spillover parameter depends on public good contributions in $t = 1$, $d_{i,1}$, which in turn depends on the external pressure, k_1, on agents to make the corresponding contribution. We assume that the spillover parameter is positive, $\alpha_i(\cdot) > 0$, and that external pressure directly increases agents' contribution levels in the first period such that $d'_{i,1}(\cdot) > 0$. Furthermore, we allow for the possibility that the external pressure exerted on agents' public good contributions in $t = 1$ affects the extent to which the contributions spill over to period two. Therefore, we introduce $\gamma(\cdot)$ as a weighting factor for $d_{i,1}(\cdot)$ in the spillover parameter with $\gamma(\cdot) > 0$, $\gamma'(\cdot) < 0$, and $\gamma(k_1) = 1$ if $k_1 = 0$. Consequently, agent i gives according to the following first-order condition in period two:

$$u'(m_{i,2} - d^*_{i,2}) = h'(\sum_{i=1}^{n} d^*_{i,2}) + \alpha_i(\gamma(k_1)d_{i,1}(k_1))f'(d^*_{i,2}),$$

where $d^*_{i,2}$ is the optimal individual public good contribution level in $t = 2$. The relationship between $d^*_{i,2}(\cdot)$ and $\alpha_i(\cdot)$ is characterized by:

$$\frac{\mathrm{d}\ d^*_{i,2}}{\mathrm{d}\ \alpha_i} = -\frac{f'(\cdot)}{u''(\cdot) + h''(\cdot) + \alpha_i(\cdot)f''(\cdot)} := \phi.$$

The concavity of utility functions implies that agents' public good contribution levels in $t = 2$ are increasing in $\alpha_i(\cdot)$, i.e., $\phi > 0$, since the spillover parameter increases the marginal utility of public good contributions in $t = 2$, $d_{i,2}(\cdot)$.

Pro-social behavior across periods — In our model, public good contribution levels in both periods are cross-periodically linked through the behavioral spillover parameter. The relationship between contributions in period one and two is given by the marginal spillover effect:

$$\frac{\mathrm{d}\ d^*_{i,2}}{\mathrm{d}\ d_{i,1}} = \alpha'_i(\cdot)\gamma(\cdot)\phi.$$

Here, $\alpha'_i(\cdot)$ determines the sign of the marginal spillover effect. Based on a variety of empirical studies, finding both negative and positive spillover effects (e.g., Blanken et al. 2015; Tiefenbeck et al. 2013; Jessoe et al. 2017; Gneezy et al. 2012; Gneezy et al. 2014; Mazar and Zhong 2010), we model the sign of $\alpha'_i(\cdot)$ to be heterogeneously distributed across agents, i.e., $\alpha'_i(\cdot) \gtreqless 0$. If $\alpha'_i(\cdot) > 0$, the effect is positive, indicating that the optimal public good contribution levels in $t = 2$, are increasing with contributions in $t = 1$, $d_{i,1}(\cdot)$. If $\alpha'_i(\cdot) < 0$, contributions in $t = 2$ are decreasing in contributions in $t = 1$. By extrapolating from an individual level to an aggregated level of decision making, we are able to determine the sign of $\alpha'_i(\cdot)$ for an average agent. The meta-study by Blanken et al. (2015) suggests that public good contribution levels in a given period on average reduces the marginal utility of contributions in a subsequent period. Therefore, we assume that for an average agent $\bar{i}$ the marginal spillover effect is less than zero, such that $\alpha'_{\bar{i}}(\cdot) < 0$ and $\frac{\mathrm{d}\ d^*_{\bar{i},2}}{\mathrm{d}\ d_{\bar{i},1}} < 0$. This implies that agents' public good contribution levels in period two are decreasing in their contributions in period one.

External pressure and pro-social behavior across periods — Also the external pressure to contribute to the public good in $t = 1$, k_1, affects contributions in $t = 2$, $d_{i,2}$, via the spillover parameter $a_{i,2}(\cdot)$. The effect is twofold. First, the external pressure directly affects public good contribution levels in period one, since $d_{i,1}(\cdot)$ is a function of k_1 in the sense that contributions in $t = 1$ depend on the external pressure exerted on agents. Via the spillover parameter, $\alpha_{i,2}(\cdot)$, these contributions affect agents' decisions in period two. Second, the external pressure affects the weighting factor, $\gamma(\cdot)$, that separately captures

the extent to which contributions spill over from the first period to the second. This can be shown analytically by deriving the effect of k_1 on $d_{i,2}$, which represents the spillover effect in case of external pressure being assigned on agents' public good contribution levels in the first period.:

$$\frac{\mathrm{d}\ d^*_{i,2}}{\mathrm{d}\ k_1} = [\gamma'(\cdot)d_{i,1}(\cdot) + d'_{i,1}(\cdot)\gamma(\cdot)]\alpha'(\cdot)\phi.$$

This expression contains two separate effects: a pure spillover effect as well as a crowding effect. The pure spillover effect, $d'_{i,1}(\cdot)\gamma(\cdot)\alpha'_i(\cdot)\phi$, is given by the marginal spillover effect, $\gamma(\cdot)\alpha'_i(\cdot)\phi$, multiplied by the magnitude to which the external pressure affects public good contributions in the first period, $d'_{i,1}(\cdot)$. It relies on the assumption that the more we move towards hard paternalistic measures to foster public good contributions in $t = 1$, the larger the contributions (e.g., d'Adda et al. 2017), such that $d'_{i,1}(\cdot) > 0$. Thus, external pressure applied on agents to contribute in $t = 1$ directly increases their public good contribution levels in this period. Since public good contributions in period one and two are linked by the marginal spillover effect, the external pressure assigned in period one also indirectly affects contributions in period two via an increase in $d_{i,1}(\cdot)$ that spills over to public good contributions in period two. Given that the marginal spillover effect is negative for an average agent and that pressure increases contributions within $t = 1$, $d'_{i,1}(\cdot) > 0$, the spillover effect is, on average, negative, i.e., $d'_{i,1}(\cdot)\gamma(\cdot)\alpha'_i(\cdot)\phi < 0$.
The crowding effect, $\gamma'(\cdot)d_{i,1}(\cdot)\alpha'(\cdot)\phi$, is based on the concept of diagnosticity (Kelley 1973; Bem 1972). This concept describes that the extent to which individuals attribute a certain action to their own self-regard is decreasing in the degree of external pressure applied on them. According to this, external pressure imposed on pro-social tasks attenuates the attribution to an individual's moral self-regard and thereby diminishes the tendency that the action affects subsequent pro-social behavior. Adapted to our model, this means that the external pressure applied to agents' public good contributions in $t = 1$ reduces the extent to which these contributions spill over to the next period. More formally, this is captured in our model by introducing the term $\gamma(\cdot)$ as a weighting factor of $d_{i,1}(\cdot)$ in the spillover parameter, such that $\alpha_i(\gamma(k_1)d_{i,1}(k_1))$. Thereby, $\gamma(\cdot)$ moderates the extent to which $d_{i,1}(\cdot)$ impacts the spillover parameter $a_{i,2}(\cdot)$ and, thereby, the public good contribution levels in period two. We assume that that the external pressure assigned on contributions in period one diminishes the influence of $d_{i,1}(\cdot)$ on $a_{i,2}(\cdot)$. Therefore, for a given contribution level in $t = 1$, the contribution weighs less for the second contribution decisions, such that the crowding effect is positive, i.e., $\gamma'(\cdot)d_{i,1}(\cdot)\alpha'(\cdot)\phi < 0$.

The spillover effect and the crowding effect have opposing impacts on $d_{i,2}(\cdot)$. Concerning the spillover effect, it is well documented that external pressure on a pro-social task increases pro-social behavior in that task (e.g., Charness and Gneezy 2009; Goette and Stutzer 2020; Ito et al. 2018; Ferraro et al. 2011) and that moral deeds spill over to the subsequent pro-social behavior (e.g., Schmitz 2019; Effron and Monin 2010; Gneezy et al. 2012). Yet, little is known about the intertemporal dimension of the crowding effect, i.e., how external pressure affects the degree to which public good contributions spillover to subsequent periods. To isolate the crowding effect, we have to separate it from the spillover effect. To do so, we hold public good contribution levels in $t = 1$ constant. This simplifies the expression of the cross-periodical impact of k_1 on contributions in $t = 2$ to:

$$\frac{\mathrm{d}\ d^*_{i,2}}{\mathrm{d}\ k_1} = \gamma'(\cdot) d_{i,1}(\cdot) \alpha'(\cdot) \phi$$

Therefore, for an average agent and a given level of $d_{i,1}(\cdot)$, we get $\frac{\mathrm{d}\ d^*_{i,2}}{\mathrm{d}\ k_1} > 0$. This shows that on average for a given public good contribution level in period one, an increase in the external pressure applied on agents' public good contributions increases the crowding effect and, thereby, has a positive impact on public good contributions in period two.

4 Experimental Design and Treatments

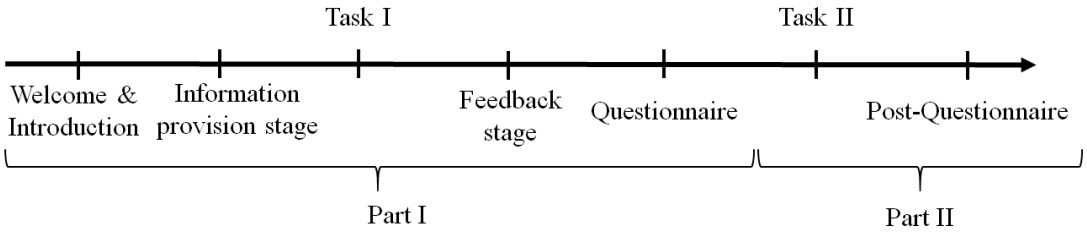
Following our theoretical considerations, we implemented our experiment into a field setting of repeated pro-social decisions within the general framework of a survey in order to experimentally identify behavioral spillover effects and test how these depend on whether the initial pro-social decisions has been incentivized. In the following we describe the general framework of our experiment, explain our treatments, derive predictions, and, finally, provide the procedural details.

4.1 General Framework

Our experiment is incorporated into an environment common for most participants. We use the general framework of a computer assisted survey on a thematically unrelated topic to confront participants with two subsequent pro-social decisions.[3] Figure 1 illustrates our setting.

[3]The survey covers technical attributes of mobile phones and inquires participants' preferences via a conjoint analysis. A detailed description of the survey is provided in the Supplementary Materials (1.2).

Figure 1: Timeline of the experiment



The survey consists of two parts, Part I and II. The core of our experimental design is given by two subsequent tasks, Task I and II, which are integrated in the general framework of our survey. At the beginning of Part I, participants are welcomed and receive a general introduction to the procedure as well as the topic. The introduction contains information on the fix participation fee of 7€ and that 5€ can be additionally earned in the course of the survey.[4] At the end of the introduction, they are made aware of the following step: At first, they receive further information on the topic of the survey. Second, the provided information is subject to a comprehension check in the form of a quiz, which takes place after the information provision. As announced, directly after the introduction the information provision stage starts.[5] Participants are asked to read information about five different technical attributes of mobile phones and their corresponding nuances carefully. Thereafter, Task I starts. It is operationalized via a quiz as real-effort task. Participants had to answer five comprehension questions on the information provided in the previous stage.[6] Participants are not able to proceed with the survey, until they have answered all comprehension questions correctly. Given that participants did not answer a question correctly, they are automatically redirected to the information provision stage and asked to read the content again. Afterwards they had the chance to retake the question. This procedure ensures that all participants, who complete Task I, have answered the five comprehension questions correctly. After Task I, participants access the feedback stage. At this stage, participants receive feedback in

[4]Since the perception of voluntariness is important for spillover effects to occur (Mullen and Monin, 2016), all participants, who are supposed to make a pro-social contribution in Part I (unless they are meant to be forced to the pro-social contribution by treatment variation), receive the information that they will have the possibility to contribute to a pro-social cause and have the option to stop the survey at this stage if they do not want to proceed. In this case, they receive 3€ for participation. All in all, only three out of 421 participants decided to stop the survey at this stage.

[5]A screenshot of the information provided to participants at the information provision stage is shown in the Supplementary Materials (2.1.3).

[6]The questions are provided in the Supplementary Materials (2.1.4).

the form of information on the resulting consequences of successfully completing Task I of the experiment.[7] Thereafter, the questionnaire starts, which serves as a filler task in our experiment. Here, participants are asked to make six consecutive choices, selecting from the option of two different mobile phones or a no choice option.[8] Once the six decisions regarding mobile phones are completed, Part I of the study is finished. At this point, it is assured that all participants will have received a reward of 5€ additional to the participation fee of 7€ within the course of Part I. Afterwards, Part II starts.
Part II consists of the second task and a post-questionnaire. Task II is a post-survey donation appeal where participants can choose whether to donate all or a share of their reward to charity.[9] They have the choice between three different charities, i.e., the World Wildlife Fund (WWF), the Nature and Biodiversity Conservation Union (NABU), and the Foodbank Mannheim.[10] Participants can either give to only one of these charities or split their donation across the different charities. Importantly, this second donation decision occurs unexpected and could not have been anticipated by the participants. The survey ends with a post-questionnaire, in which we inquire a range of individual preferences and characteristics from participants.[11]

4.2 Treatments

Our experimental design varies whether and, if so, how participants' performance in Task I of the experiment is incentivized. Task II is held constant across all treatments. Our treatments differ across two dimensions. First, we influence whether participants' performance in the first task has a pro-social component. Second, if this is the case, we vary how the pro-social consequences of completing Task I are additionally incentivized. Table 1 provides an overview.
In our control treatment, *NoSocial*, participants' performance in Task I has no pro-social consequences that are, consequently, also not additionally incentivzed. Hence, participants take the five comprehension questions in the quiz only for the purpose of proceeding with the survey. Within the quiz, the amount of questions already answered

[7] This can comprise of pro-social contributions and additional monetary payoff, depending on treatment assignment. The feedback in the respective treatments can be retrieved from the Supplementary Materials (2.1.4.2)

[8] The inquired choice-set can be retraced in the Supplementary Materials (2.1.5).

[9] At the end of the survey, participants have the option to receive a receipt on aggregated donations raised by all participants in the survey via e-mail once the transfers to the charities are conducted. This option is provided to assure participants that the donations are actually carried out.

[10] Details regarding the charities and the donation option are provided in the Supplementary Materials (2.1.6).

[11] The variables surveyed in the post-questionnaire are listed in Table 6.

Table 1: Overview treatment interventions

Treatment	Pro-social component		Additional incentives		Observ.
NoSocial	No		No incentive		150
Baseline	Yes	Pro-social treatments	No incentive		144
Nudge	Yes		Social comparison	Interventional treatments	127
Monetary Inc.	Yes		Monetary rewards		150
Punishment	Yes		Time punishment		137

correctly are provided to participants by a progress bar.[12] After completing the quiz, the treatment deviates from the general procedure as we skipped the feedback stage.
In all other treatments, completing Task I of the experiment has pro-social consequences. In *Baseline*, participants' performance in Task I of the experiment has pro-social consequences that are, however, not additionally incentivized. Participants are informed that each correctly answered question in the quiz results in a donation in form of a financial transfer to a community afforestation project in Nepal. The donation amount is sufficient to finance the plantation of a tree within the project.[13] Hence, the completion of Task I translates into the well-being of others and, thereby, entails a pro-social incentive to provide effort to carefully answer the inquired questions in the quiz correctly. We ensure that the pro-social incentives still hold, although the size of the pro-social contributions in Task I is predetermined. By the provided information, participants are aware that a correctly answered question in Task I generates a donation. By design, participation in Task I leads to five correct questions.[14] During the quiz, the progress on the yet generated donations is constantly displayed to participants by a progress bar. Directly after the quiz, participants receive feedback on the generated donations at the feedback stage.[15]
In addition to the pro-social component, participants' performance in Task I of the experiment is additionally incentivized in our three remaining treatments. In our *Nudge* treatment, the pro-social consequences of Task I are additionally incentivized by informing participants on their relative performance in the quiz compared to other participants. Participants are informed about the modalities of the *Nudge* treatment in the

[12] A screenshot of the progress bar is provided in the Supplementary Materials (2.1.4).
[13] Details regarding the afforestation project are provided in the Supplementary Materials (2.2).
[14] While participants are informed that completing Task I leads to donations, they do not know at this stage that they will have answered all five questions correctly at the end of Task I.
[15] The progress bar and the explicit content of the feedback can be retraced in the Supplementary Materials (2.1.4).

introduction phase. It is conveyed that they receive real-time feedback on their relative performance in the quiz. The feedback is applied during the process of answering the comprehension questions in the first task of the experiment. As in the *Baseline* treatment, at each question the yet generated donations are immediately visualized through progress bars. In addition, there are two additional progress bars provided. The second bar shows the generated donation at the corresponding attempt for the average participant. The third bar reports the donations generated by the best performing 25 percent of participants at the corresponding attempt. For a more intuitive visualization of the relative performance, emojis were provided, indicating, whether participants' performance is within the upper 25 percent by a happy green face, within the average by a neutral orange face or lower than average performance by a frowny red face.[16] Exactly as in our *Baseline* treatment, participants received the information on the generated donations by correctly answering the comprehension questions, on the feedback stage that automatically follows Task I.

In our *Monetary Incentive* treatment, successfully completing Task I of the experiment does not only lead to a contribution to an afforestation project, participants also receive an additional payment for each correctly answered question. At the end of the introduction, participants are informed that they receive 1€ for each correct answer in the quiz. For better visualization of the payments per correct question, within the quiz stage, an additional progress bar reports the additionally earned remuneration generated by correct answers in the quiz. After having completed the quiz, participants receive feedback on the generated donations as well as their additional earnings generated by answering the comprehension questions correctly.[17] In order to avoid endowment effects, participants in all other treatments receive 5€ en bloc for completing the questionnaire stage, instead of 1€ for each correctly answered question in Task I separately, as it is the case in the *Monetary Incentive* treatment.

Finally, also in our *Punishment* treatment Task I contains a pro-social component through donations generated in the quiz. In addition, the treatment manipulation exerts pressure on the participants such that they feel highly pressured to answer all five comprehension questions in the first task of the experiment correctly. To do so, participants are informed that there is a time punishment for each incorrectly answered question. We impose a penalty of 15 seconds for each incorrectly answered question. Participants who failed to answer a question correctly must wait for 15 seconds to be able to proceed with Task I. In addition, the wording is adjusted towards a more compelling tone. In

[16] A screenshot is provided in the Supplementary Materials (2.1.4).

[17] The graph containing the progress bars and the the feedback provided to participants at the end of the quiz is depicted in the Supplementary Materials (2.1.4).

particular, instead of asking participants to take part in the quiz, we indicate that they are obliged to do the quiz, in order to proceed with the survey. In addition, while in *Baseline*, *Nudge* and *Monetary Incentive* participants are provided with the option to stop the survey after the introduction and receive 3€ as compensation if they wish not to contribute to a pro-social cause, this option is not provided in the *Punishment* treatment, in order to increase participants' notion of being pressured to act pro-social. Similar to the *Baseline* treatment, participants receive the information on the generated donations at the feedback stage of the experiment directly after the comprehension questions.

4.3 Hypotheses

The following section merges our theoretical considerations and the experimental design. First, we show that our design reflects the essential elements of the theory. Second, we derive a set of clear hypotheses for our different treatment effects.

Based on our theoretical consideration, the core of our experimental design consists of two subsequent decisions, i.e., Task I and II. Task II is a donation appeal with a clearly defined pro-social component. For Task I, we vary whether its completion has pro-social consequences and, if so, how these are additional incentivized. Thereby, we can experimentally investigate how donations in the second task of the experiment depend on the pro-social consequences of a previous task and how these has been additionally incentivized.

Our theoretical considerations predict negative spillover effects. In general, participants' pro-social behavior reduces the marginal utility of subsequent donations such that participants' donations are decreasing with their previous pro-social engagement. Transferred to our experimental design, this means, that participants' donations in Task II of the experiment are expected to be lower when Task I contains a pro-social component than when completing Task I entails no pro-social consequences. Accordingly, participants in our *Baseline* treatment, where Task I has pro-social consequences that are, however, not additionally incentivized, donate less in the second task of our experiment than participants in our *NoSocial* treatment, where Task I has no pro-social consequences. To provide an overview, we refer to participants' donations in Task II in *Baseline* and *NoSocial* as $d_{Baseline}$ and $d_{NoSocial}$, respectively, and define the pure spillover effect (PSE) as $PSE = d_{Baseline} - d_{NoSocial}$. This is summarized in our first hypotheses.

HYPOTHESIS 1. (Negative pure spillover effect): *In Task II, participants' average donations are smaller in Baseline than in NoSocial, i.e.,* $PSE < 0$.

On this basis, our theoretical considerations further predict that the extent to which

participants' pro-social behavior negatively spills over to a subsequent donation decision is crowded out by external pressure assigned on the initial decision. In our interventional treatments, we impose a varying degree of external pressure on participants to behave pro-social in Task I of the experiment via additional incentive schemes. One experimental difficulty in identifying and measuring the crowding effect is that the external pressure assigned on participants' initial pro-social decision can cause two separate effects: a pure spillover as well as a crowding effect. External pressure could directly affect participants' performance in Task I of the experiment and, thereby, their initial pro-social contribution that spills over to their decisions in Task II. In addition, the external pressure is also expected to crowd out the extent to which participants' performance in Task I spills over to their donation in Task II. One special feature of our experiment is that we are able to isolate the crowding effect. To do so, we fix the pro-social component generated by participants in Task I of the experiment for all participants across all our pro-social treatments. This means, independently of how strongly we additionally incentivize the pro-social component of Task I, all participants make exactly the same donation to a community afforestation project in Nepal. Thereby, we eliminate spillover effects caused by the mark-up in pro-social behavior, which is induced through incentives. This has the advantage that we can isolate crowding effects by simply manipulating whether and, if so, how we assign external pressure to participants' performance in the first task of the experiment.

In our *Baseline* treatment, successfully completing Task I of the experiment has pro-social consequences that are, however, not additionally incentivized. In our interventional treatments, in contrast, participants' performance in the first task is additionally incentivized. Here, we increase the external pressure on participants' performance in Task I compared to our *Baseline* treatment. Also within our interventional treatments, we vary the degree of external pressure assigned to participants in the first task of the experiment. In ascending order of external pressure, our treatments *Nudge*, *Monetary Incentive*, and *Punishment* reflect different interventions ranging from liberal to hard paternalistic measures. Our theoretical considerations predict that an increase in the external pressure assigned on participants in Task I increases the extent to which the negative spillovers are crowded out. Consequently, we expect that the crowding effect are larger, the higher the external pressure assigned on participants' performance in Task I of the experiment. Experimentally, we capture the crowding effects (CE) as the difference in participants' donations in our interventional treatments and the donations in our *Baseline* treatment. Following our notation, we refer to participants' donations in Task II of the experiment in our interventional treatments *Nudge*, *Monetary Incentive*, and

Punishment as d_{Nudge}, $d_{MonetaryIncentive}$, and $d_{Punishment}$ such that the crowding effects are given by $CE_T = d_T - d_{Baseline}$ with $T \in \{Nudge, Monetary\ Incentive, Punishment\}$. This leads to our second hypothesis.

HYPOTHESIS 2. (Positive crowding effect): *An increase in the degree of external pressure imposed on participants in Task I crowds out the extent to which participants' initial pro-social behavior spills over to their donation behavior in Task II, i.e.,* $CE_{Nudge} < CE_{MonetaryIncentive} < CE_{Punishment}$.

Since the spillover effects in our interventional treatments are composed of two effects, the pure spillover effect as well as the crowding effect, they are expected to be lower in our *Baseline* treatment than in *Nudge*, *Monetary Incentive*, and *Punishment*. Put differently, the spillover effects in *Nudge*, *Monetary Incentive*, and *Punishment* are given by $SE_T = d_T - d_{NoSocial}$ with $T \in \{Nudge, Monetary\ Incentive, Punishment\}$. Consequently, in absolute terms the spillover effects in our interventional treatments are expected to be smaller than in our *Baseline* treatment and decreasing with the degree of external pressure, i.e., $SE_{Baseline} < SE_{Nudge} < SE_{MonetaryIncentive} < SE_{Punishment}$.

4.4 Procedure & Descriptive Statistics

Following a strand of experimental literature that shows that behavioral spillover effects are highly variable and depend on individual characteristics and attitudes (e.g., Blanken et al. 2015; Mullen and Monin 2016), we address a non-student sample in order to capture these heterogeneous effects. To assure the necessary diversity of the sample and at the same time remain within a natural environment, which is common to most of our participants, we implement our experiment into a field setting. To do so, we set up a stall in the pedestrian zone in Mannheim, Germany, in August 2019, which is comparable to those of many other organizations like NGOs, political parties or corporate institutions. As in many other cities, in Mannheim stalls serve as a common vehicle to reach out to the general public and represent an established part of the city's landscape. The purpose of such stalls is diverse, varying from purely informational campaigns, opinion polls, discussion platforms to recruiting new members for certain organizations. Our stall was located at a square which is situated directly at the main station. This location was chosen as it minimizes neighborhood effects and attracts the attention of a wide cross-section of society ranging from employees commuting to younger adults and elderly people.[18] The stall was staffed on eight days within the time from 6th of August to

[18] In addition, it had the advantage that individuals with layovers between trains were attracted to participate.

the 28th of August, between 11am and 5pm. It hosts our mobile laboratory, which is suitable to conduct computed-based economic experiments under controlled conditions. Our recruiting strategy relies on two posters next to the stall to attract participants among passing pedestrians.[19] The posters provide general information on the average participation fee of 12 € , the time needed to complete the survey, as well as the institution conducting the survey. The content of the poster was well-readable from larger distances and was kept as brief as possible. The posters do not provide any information about the topic or pro-social component of the survey.

Directly in front of the stall we set up our reception. Interested pedestrians had the option to register as participants in our experiment here. To register, participants were asked to give their names and show their ID such that we can rule out that they participate in the experiment multiple times.[20] It was conveyed that the personal information serve the purpose of recording participation only and cannot be linked to the indications made in the survey. After the registration, participants were guided to our mobile laboratory. In case of questions concerning the procedure, a research assistant was eligible for clarification.

Our laboratory consisted of six identical separated cubicles, all equipped with a portable computer, a pen and paper.[21] To guarantee the anonymity of participants and their data, partitioning walls served as sight protection to other participants and passing pedestrians. To protect participants from other factors, such as solar radiation or rain, a pavilion has been set up. Before taking part in the survey, participants were asked not to talk to any of the other people taking part in the survey nor to use their mobile phones. Questions about the content and the technical details of the survey could be addressed at each time to the experimenter, who supervised participation in the mobile laboratory. Since further clarification on the procedure was provided within the survey, participants were asked to follow the instructions on the computer screen. We used the software o-Tree developed by Chen et al. (2016) for programming. As no interaction between participants in the survey was necessary, it was run on the local servers of the portable computers.[22] The software puts the survey on full screen mode and prevents participants from using the computer for any other purposes except of answering the questions provided in the survey. Once participants completed the survey, they were informed about the remuneration on the computer screen and were asked to raise their hand to inform the experimenter. Finally, participants were asked to fill in a receipt

[19] A copy of the poster is shown in the Supplementary Materials (1.1).

[20] Only a two people intended to participate several times.

[21] A sketched plan and a photo of our mobile laboratory is shown in the Supplementary Materials (1).

[22] The data was saved, collected, and merged from each portable computer at the end of each day.

containing the amount to be received and their signature to receive their payment. Afterwards, the remuneration was payed out in cash by the experimenter. The payment procedure was conducted at the cubicles, while the participants remained seated, to guarantee that the amount received was exclusively known by the experimenter and the respective participant and did not influence the decision of other subjects.
In total 711 participants took part in the experiment. The number of participants per day fluctuated between 54 and 121. On average, participation in the experiment lasted about 15 minutes and participants received an individual payment of 10.07€ . The average amount donated was at 1.85€ . To provide an overview on the characteristics of the sample, the descriptive statistics in Table 2 show the mean, median, minimum and maximum values of participants' core demographic variables. On average, participants are 32.6 years old and have a mean income of 1230€. Furthermore, the average years of education is 13.7 with 37% women and 63% men.[23] Among participants, 37% are enrolled in a university and had donated on average 114€ during the previous year, with a median value of 10€. Also, 63% of participants voted for a left wing party at the last election.

Table 2: Descriptive sample statistics

	Mean	Median	Min	Max
Payoff	10.11	12	0	13.9
Donation (Task II)	1.81	0	0	12
Completion time (in min)	15	13	5	65
Age	32.5	27	15	79
Income	1250€	0-1000€	0-1000€	>5000€
Years of Education	13.85	13	7	18
Female	0.36	0	0	1
Student	0.37	0	0	1
Donations last year	128€	10€	0€	10,000€
Left wing voters	0.63	1	0	1

Note: Further information on the definition and measurement of the variables can be retrieved from Table 6 in Appendix A.

[23]This compares to the average population as the sample is younger given the mean age in the city of Mannheim is 42.6 (City of Mannheim, 2020). This is not surprising as participation in the survey required the navigation of a computer. The average income is also below average given a value of 1716€ a month for the city of Mannheim (Seils and Baumann, 2019). The years of education are in line with the average within the German population, which is at 13.14 years (Federal Statistical Office, 2020).

5 Results

We investigate how charitable giving is affected by previous pro-social behavior and, most importantly, how different incentives to previous behavior affect the extent to which this behavior spills over. At first, we present the average amounts donated in each treatment. Thereafter, we compare donations across treatments and analyze behavioral spillover effects, which are further decomposed into sub-effects. Lastly, we directly assess the influence of perceived pressure on subsequent donations and compare these results with the obtained spillovers from the treatment effects.

5.1 Treatment and Spillover Effects

Figure 2 shows the average donations in Task II of the experiment per treatment. Participants in *NoSocial* donate the largest amounts with 2.48€ on average. This is followed by donations in *Nudge* and *Punishment*, in which participants on average decide to donate 2.16€ and 1.65€ , respectively. In the *Baseline* treatment, participants give 1.50€ on average, while the fewest donations are made in the *Monetary Incentive* treatment with 1.29€ on average.[24]

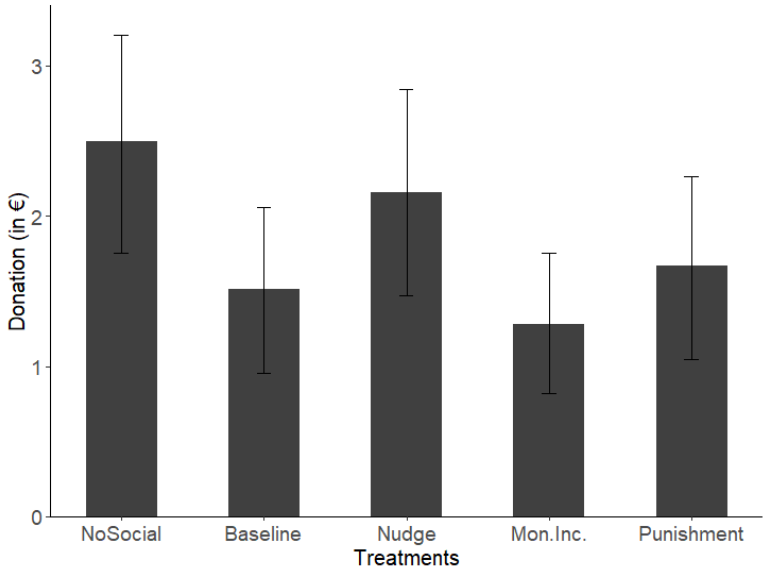
To investigate how charitable giving is affected by previous pro-social behavior and, most importantly, how this has been incentivized, we start with analyzing the spillover effects. These capture the difference in donations in our pro-social treatments where Task I of the experiment contains a pro-social component, i.e., *Baseline*, *Nudge*, *Monetary Incentive*, and *Punishment*, and donations in the case that the completion of Task I has no pro-social consequences, i.e., *NoSocial*. More formally, spillover effects are given by $SE_T = d_T - d_{NoSocial}$ with $T \in \{Baseline, Nudge, Monetary\ Incentive, Punishment\}$. The corresponding results are shown in Table 3.

Table 3 – Column 1 contains the results of an OLS-regression of participants' donations in Task II of the experiment on the set of indicators variables for our four pro-social treatments. Participants in *NoSocial* are our control group. The treatment coefficients capture whether participants alternate donation decisions given the pro-social consequences in (*Baseline*) and additional incentives, i.e., via social comparisons (*Nudge*), monetary incentives (*Monetary Incentives*) or time-punishments (*Punishment*), in the previous task.

Following our theoretical considerations, we estimate the pure spillover effect (PSE) by comparing participants' donation behavior in *Baseline* and *NoSocial*. It measures

[24] To illustrate participants' donation behavior at the intensive and extensive margin, we show the corresponding cumulative distributions of donations in Appendix B.2 Figure 5.

Figure 2: Donations in Task II across treatments



Note: Average donations in Task II of the experiment by treatment. Confidence intervals at the 95%-level.

the difference in participants' donations given that they have previously completed a pro-social task that has not been additionally incentivized ($d_{Baseline}$) and donations in case the previous task has no pro-social component ($d_{NoSocial}$), i.e., $PSE = d_{Baseline} - d_{NoSocial}$. In line with our first hypothesis and the literature on moral balancing (e.g., Gneezy et al. 2012, Mazar and Zhong 2010), we find a negative pure spillover effect. Participants donate significantly less when they have completed a pro-social task that has not been additionally incentivized than when the previous task did not lead to pro-social consequences. While participants in *Baseline* donate on average 1.50€, participants in *NoSocial* donate 2.48€. With a difference of 0.98€ (p-value=0.0237, Table 3 – Column 1), participants in *Baseline* donate on average 40% less than participants in *NoSocial*. To summarize,

OBSERVATION 1. *The pure spillover effect is negative. Participants, who completed a pro-social task that has not been additionally incentivized prior to the donation decision, donate significantly less than participants who did not complete a pro-social task previously.*

Table 3: OLS-regression of treatment effects on donation behavior

	Dependent variable: Donations in Task II							
	(1)	(2)			(3)	(4)		
	OLS	Hurdle			OLS	Hurdle		
	donations	Two stage			donations	Two stage		
		Extensive Margin		Intensive Margin		Extensive Margin		Intensive Margin
		Coefficients	Average margina effects			Coefficients	Average margina effects	
Baseline	−0.973*	−0.137	−0.048	−2.227*	−1.005*	−0.168	−0.056	−2.359*
	(0.429)	(0.155)	(0.054)	(0.943)	(0.434)	(0.163)	(0.054)	(0.924)
Nudge	−0.319	0.094	0.033	−1.651^{+}	−0.302	0.100	0.033	−1.686^{+}
	(0.443)	(0.156)	(0.054)	(0.908)	(0.440)	(0.162)	(0.054)	(0.873)
Monetary Incentive	−1.190**	−0.095	−0.033	−3.251***	−1.047*	−0.028	−0.009	−3.129***
	(0.424)	(0.152)	(0.053)	(0.919)	(0.425)	(0.158)	(0.052)	(0.890)
Punishment	−0.823^{+}	−0.121	−0.042	−1.794^{+}	−0.739^{+}	−0.102	−0.034	−1.651^{+}
	(0.438)	(0.157)	(0.055)	(0.957)	(0.436)	(0.164)	(0.054)	(0.934)
Constant	2.477***	−0.468***		7.740***	0.818	−0.904^{+}		4.976^{+}
	(0.300)	(0.107)		(0.632)	(1.259)	(0.469)		(2.814)
Control Variables					X	X	X	X
Observations	703	703	703	212	676	676	676	209
R^2	0.015			0.060	0.092			0.186

Note: Models in columns 1 and 5 are OLS regressions. The dependent variable is the amount donated in Task II of the experiment. Models in columns 2 to 4 and 6 to 8 are two-stage hurdle models. First stage, Probit regression model, where the dependent variable is equal to one for positive donations and zero otherwise. Coefficients in columns 2 and 6. Average marginal effects in columns 3 and 7. Second stage, columns 4 and 8, truncated regression models, where the dependent variable is the amount donated in Task II of the experiment, conditional on donations being positive. Controls: last year's donations, female, age, native speaker, student, years of schooling, part. time, part day (see Table 6 in Appendix A for definitions). Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001.

The presence of a strong pure spillover effect is robust to controlling for a wide range of covariates (see Table 3 – Model 3). In addition, in Table 3 – Model 2 and 4, we use a two-stage estimation procedure to separately estimate the effects on participants' propensity to give and the amount donated conditional on a positive donation. While we do not find a significant pure spillover effect at the extensive margin (see Table 3 – Model 2), average donations of donors differ between *Baseline* and *NoSocial*. In *Baseline*, donors give 2.22€ less than in *NoSocial* (p-value=0.019, Table 3 – Model 2). The results at the extensive and intensive margin are also robust to controlling for covariates (see Table 3 – Model 4).

In our interventional treatments, i.e., *Nudge*, *Monetary Incentive*, and *Punishment*, participants' pro-social behavior in Task I is additionally incentivized via a set of respective treatment interventions. The results in Table 3 – Model 1 show that the spillover effects depend on the type of the intervention. Although donation levels in all our interven-

tional treatments are lower than donations in *NoSocial*, the difference reaches statistical significance only in the *Monetary Incentive* treatment (1.21 vs. 2.48, p-value=0.005, Table 3 – Model 1) and the *Punishment* (1.65 vs. 2.48, p-value=0.061, Table 3 – Model 1) treatment. We do not find significant differences at conventional levels of statistical inference in *Nudge* (2.16 vs. 2.48, p-value=0.472, Table 3 – Model 1). While the spillover effect in *Punishment* amounts to 0.83€, the average donations of 1.29€ in *Monetary Incentive* are considerably lower than the donations of 2.48€ in *NoSocial* (p-value=0.005, Table 3 – Model 1), leading to a statistically significant spillover effect of 1.19€. This leads to our second observation.

OBSERVATION 2. *Participants, whose initial pro-social behavior has been additionally incentivized via monetary incentives or punishments, donate significantly less than participants who did not complete a pro-social task previously. Participants in the Nudge, in contrast, treatment do not show significant spillover effects.*

The regression in Table 3 – Model 3 supports this observation. The results in our *Monetary Incentive* treatment show a significant spillover effect, which is robust to controlling for covariats (see Table 3 – Column 2). At the extensive margin (see Table 3 – Model 2), we do not find a higher propensity to give in the *Monetary Incentive* treatment. At the intensive margin (see Table 3 – Model 2), however, we find that donors give significantly more in *Monetary Incentive* than in *NoSocial*.

Next, we empirically disentangle the spillover effect in our interventional treatments. Following our theoretical considerations, the spillover effects in our interventional treatments contain two separate components: A pure spillover component as well as a crowding component. Our experimental design allows us to disentangle the spillover effects and isolate the crowding effect from the pure spillover effect. The crowding effect (CE) captures the difference in donations in our interventional treatments and donations in our *Baseline* treatment, such that $CE_T = d_T - d_{Baseline}$ with $T \in \{Nudge, Monetary\ Incentive, Punishment\}$. Following this, we can rewrite the spillover effects in our interventional treatments into pure spillover effect plus the crowding effect, i.e., $SE_T = PSE + CE_T$. The results of our decomposition analysis are summarized in Table 4.[25] Additionally, we decompose the average spillover effects (see Table 4 – Panel A) as well as the spillover effects at the extensive (Table 4 – Panel B) and intensive margin (Table 4 – Panel C).

The results in Table 4 – Panel A column three shows the average crowding effects.

[25]The estimates are based on a series of regression models reported in Table 9 and Table 10 in Appendix B.2

Table 4: Decomposition Analysis

Treatment	Spillover Effect	Pure Spillover Effect	Crowding effect
Panel A. Average effects			
Nudge	-0.319	-0.973*	0.654
	(0.424)	(0.429)	(0.448)
Monetary Incentives	-1.19**	-0.973*	-0.217
	(0.425)	(0.430)	(0.429)
Punishment	-0.823+	-0.973*	0.151
	(0.438)	(0.429)	(0.443)
Panel B. Extensive margin			
Nudge	0.033	-0.048	0.08
	(0.054)	(0.054)	(0.055)
Monetary Incentives	-0.033	-0.048	0.014
	(0.053)	(0.054)	(0.054)
Punishment	-0.042	-0.048	0.006
	(0.055)	(0.054)	(0.056)
Panel C. Intensive margin			
Nudge	-1.651+	-2.227*	0.567
	(0.908)	(0.943)	(0.957)
Monetary Incentives	-3.251***	-2.227*	-1.024
	(0.917)	(0.943)	(0.968)
Punishment	-1.794+	-2.227*	0.433
	(0.957)	(0.943)	(1.004)

Note: Decomposed spillover effects in our interventional treatments, i.e., *Nudge*, *Monetary Incentives*, and *Punishment*, on average (Panel A), the extensive margin (Panel B), and the intensive margin (Panel C). The coefficients are retrieved from regression Table 9 and Table 10 in Appendix B.2. The extensive margins are given by average marginal effects from a probit model. Following our theoretical considerations, we can decompose the spillover effects into a pure spillover effect and the crowding effects. Robust standard errors in parentheses. Further information on the definition and measurement of the variables can be retrieved from Table 6 in Appendix A.Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001.

We find a crowding out of 0.65€ in *Nudge* and 0.15€ in *Punishment*. In *Monetary Incentive*, in contrast, we find a weak crowding-in of -0.22€ . The crowding effects are not significantly different from zero. However, we find a statistically significant difference by comparing crowding effects across treatments. The crowding effect in *Nudge* and *Monetary Incentive* differ significantly by 0.87€ (0.65 vs. -0.22, p-value=0.0497, Table 11 in Appendix B.2). This indicates that an incentivation of pro-social behavior via social comparisons induces substantially larger crowding effects compared to an incentivation

that is based on monetary rewards. Put differently, the spillover effects, in absolute terms, are significantly larger in *Monetary Incentive* than in *Nudge*. This is summarized in our third observation.

OBSERVATION 3. *The use of social comparisons to additionally incentivize pro-social behavior leads to a significantly larger crowding effect than using monetary incentives.*

Our analysis at the extensive (see Table 4 – Panel B) and intensive margin (see Table 4 – Panel C) reveals that this observation is largely driven by variations in the decisions of donors. We do not find statistically significant differences in participants' propensity to give. The largest difference to the propensity to donate in *Baseline* is observed in *Nudge*. We observe a crowding effect for the *Nudge* treatment, which outweighs the pure spillover effect, as the share of donors increases by about eight percentage points (0.273 vs. 0.354, p-value=0.147). In contrast, the pure spillover effect in *Nudge* diminishes the propensity to donate by about 5 percentage points (0.32 vs. 0.273, p-value=0.3752). In *Monetary Incentive* and *Punishment*, we observe marginal crowding effects of around one percentage points (0.273 vs. 0.287 , p-value=0.791; 0.273 vs. 0.278 , p-value=0.919), resulting in negative overall spillover effects due to the comparably larger negative pure spillover effect. In contrast to the effects on the extensive margin, on the intensive margin of donations, we observe the largest crowding effects in the *Monetary Incentive*. In this treatment, there is a crowding-in of -1.02€ . In contrast, a crowding-out of 0.48€ is reported for the *Punishment* treatment (5.51 vs. 4.49, p-value=0.291; 5.51 vs. 5.95 p-value=0.667) and of 0.64€ in *Nudge* (5.51 vs. 6.09, p-value=0.548). Although the crowding effects at the intensive margin in the interventional treatments are statistically indistinguishable from zero, provisions of donors to charity are weakly statistically significant in *Nudge* compared to donations in *Monetary Incentive* (6.09 vs. 4.49, p-value=0.088). The latter result further supports observation three.

5.2 Perceived Pressure and Spillover Effects

Our analysis of treatment and spillover effects reveals that spillovers clearly depend on how the pro-social component of the first task of the experiment has been incentivized. Deviating from our theoretical considerations, however, we find that an incentivation via monetary incentives and time punishments induces stronger spillover effects than an incentivation that is based on social comparisons. The analysis in Section 5.1 leaves open why participants in *Monetary Incentive* and *Punishment* donate relatively smaller amounts in Task II of the experiment compared to participants in *Nudge*. In this section

we extend our analysis and investigate how participants perceive the pressure induced by our different treatment interventions and how this perceived pressure affects their individual spillover levels. More precisely, at first, we test whether our treatments actually induce the intended degree of external pressure on participants' pro-social behavior in Task I of the experiment. Second, we investigate how participants' perceived external pressure affects their subsequent donation decisions in Task II of the experiment.

To assess the causes of participants behavior in Task I of the experiment, we measure perceived pressure by asking participants to which degree they perceive their pro-social behavior in Task I to be driven either by their intrinsic motivation or external pressure.[26] To do so, we inquired a corresponding question in the experimental post-questionnaire.[27] One methodological challenge in the empirical relationship between participants' perceived pressure in Task I of the experiment and their donation behavior in the second task is that participants with fewer pro-social attitudes could be more likely to feel highly pressured in Task I. At the same time, they are also expected to contribute lower amounts to charity in Task II. This causes selection effects in the measure of perceived pressure. In order to control for the effects of pro-social attitudes, we adjust participants' perceived pressure by weighing it with their pro-social attitudes.[28] As an indicator for participants' pro-social attitudes, we rely on their approval to a series of pro-social statements in the experimental post-questionnaire.[29]

Figure 3 shows participants' perceived pressure in Task I of the experiment across treatments.[30] Panel A shows the values of participants' perceived pressure across treatments, Panel B summarizes the adjusted perceived pressure. The results in Figure 3 (Panel A) reveal that participants perceive the least external pressure in our *Nudge* treatment (5.71), followed by *Baseline* (6.10), *Monetary Incentive* (6.51), and, finally, *Punishment* (6.88). The difference between participants' perceived pressure reaches statistical significance between *Baseline* and *Punishment* (p-value=0.0276, M-W U Test), *Nudge* and *Monetary Incentive* (p-value=0.0181, M-W U Test), and *Nudge* and *Punishment*

[26]The question on perceived pressure in Task I was only inquired if the treatment involved pro-social contributions in this firs task. Therefore, the comparisons of perceived pressure solely considers observations from the pro-social treatments.

[27]In the question, we asked participants to indicate whether they conducted the first task due to their intrinsic motivation or external factors on a Likert-scale from one to ten. Further details are provided in Table 6.

[28]Consequently, participants' adjusted perceived pressure is given by $adj.\ perceived\ press. = \frac{perceived\ press.}{reversed(pro-social\ attitudes)}$.

[29]Pro-social attitudes were inquired in the post-questionnaire by asking participants to state their accordance with four normative pro-social statements. More details are provided in Table 6.

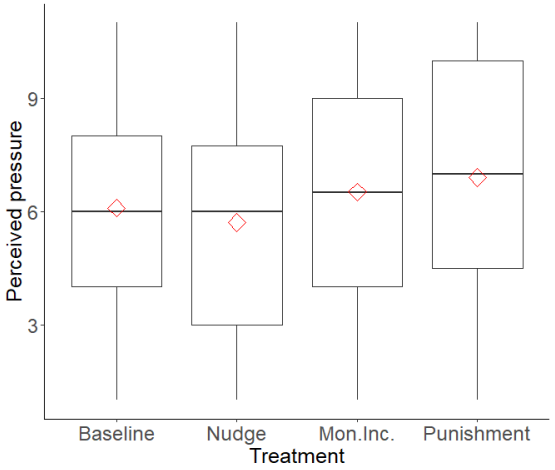
[30]Evidences on the effectiveness of the incentives in terms of influencing participants behavior in the real effort task in Part I is provided in Appendix B.1.

(p-value=0.0021, M-W U Test). Participants' adjusted perceived pressure (Figure 3 - Pane B) shows a very similar trend. Yet, concerning the differences between treatments, the pressure levels of *Baseline* and *Punishment* are not statistically distinguishable (p-value=0.1992, M-W U Test). Instead, the pressure level in *Nudge* is significantly lower than in *Monetary Incentive* and *Punishment* (p-value=0.0049, M-W U Test; p-value=0.0045, M-W U Test). This is summarized in our next observations.

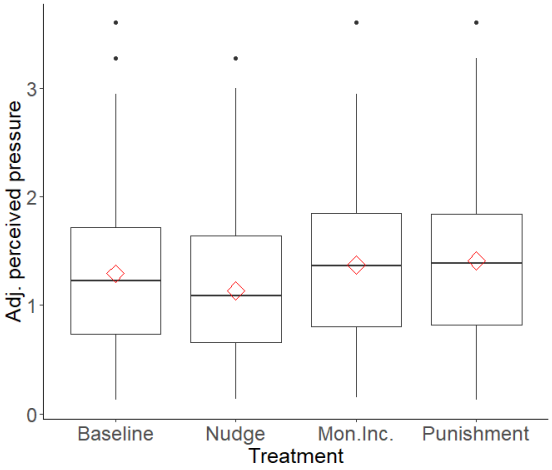
OBSERVATION 4. *Participants' perceived level of external pressure to act pro-socially in Task I of the experiment varies across treatments. It is particularly low in our Nudge treatment.*

Figure 3: Perceived pressure by treatments

(a) Panel A: Perceived pressure



(b) Panel B: Adjusted pressure



Note: The mean and median values of the perceived pressure variable and the adjusted perceived pressure variable can be observed in Table 13 in Appendix B.3. The perceived pressure variable measured in the post-questionnaire on a 0 to 10 scale is used in the left graph. Adjusted perceived pressure, which corrects for selection effects, is used in the right graph. Within the boxplots, the horizontal bar shows the median and the red hash represents the mean value of perceived pressure.

Since participants do not perceive the degree of external pressure as intended by our pro-social treatments, we also apply a more fine-grained decomposition analysis by using participants' perceived level of external pressure as an explanatory variable. To do so, we classify participants in our pro-social treatments into a low and high pressure group, depending on their reported perceived level of external pressure in Task I of the experiment. We determine a cutoff level at the 75th percentile of the perceived

pressure variable.[31] Complementary to the classification based on the perceived pressure variable's cutoff, we conduct the similar analysis using the adjusted perceived pressure variable. Based on our classification, we can specify three groups of participants in order to conduct our decomposition analysis. The first group, captures participants in the *NoSocial* treatment, as they did not contributed to charity in Task I of the experiment and, consequently, did not feel pressured to behave pro-social. The second group is given by participants in our pro-social treatments, who did not perceive a high degree of external pressure to act pro-social in the first task. This is complemented by our third group, which consists of participants in the pro-social treatments, who felt highly pressured to complete the pro-social task. Following the logic of our theoretical considerations and the analysis in Section 5.1, we estimate the spillover effects as the difference in subsequent donations of participants in our pro-social treatments, who report a relatively high level of perceived external pressure, and participants in our *NoSocial* treatment. The pure spillover effect is captured by the difference in donations in Task II of participants in our pro-social treatments, who state a low level of perceived external pressure, and participants in *NoSocial*. Finally, we identify the crowding effects by comparing participants in our pro-social treatments, who state a relatively high level of external pressure, to those, who report a low level of perceived pressure.
The results are summarized in Table 5. The coefficients for the average spillover effect are shown in the first column of Panel A. We observe a significant negative spillovers of -1.58€ (2.48 vs. 0.88, p-value=0.0015, Table 15), which looses some of its strength when we adjust the participants' perceived pressure levels for potential effects of self-selection (2.48 vs. 1.15, p-value=0.004, Table 15). On average, participants reduce donations by 63% (56% applying the adjusted perceived pressure variable) if they previously felt a high level of external pressure to contribute to a pro-social cause in the first task of the experiment compared to participants in *NoSocial*. To summarize,

OBSERVATION 5. *Participants who felt highly pressured to act pro-social in Task I of the experiment donate significantly less in Task II compared to those in our NoSocial treatment.*

The spillover effects on the extensive margin and the intensive margin in Table 5 confirm the influence of perceived pressure on subsequent donation levels. Concerning the propensity to donate, we observe a significant negative spillover effect of -0.16 (0.32 vs.

[31] The cutoff level at the 75th percentile was selected based on an analysis of the correlation between external pressure and donations. Figure 6 in Appendix B.3 shows the donation behavior of participants distinguished by perceived external pressure. It indicates that a comparably high level of perceived pressure (above the 75th percentile) is require to notice an effect on subsequent donations.

Table 5: Decomposition Analysis

	Spillover Effect	Pure Spillover Effect	Crowding Effect
Panel A. Average effects			
Unadjusted	-1.598^{**}	-0.652^{+}	-0.946^{*}
	(0.498)	(0.362)	(0.381)
Adjusted	-1.329^{**}	-0.738^{*}	-0.591^{+}
	(0.457)	(0.363)	(0.335)
Panel B. Extensive margin			
Unadjusted	-0.163^{**}	0.013	-0.19^{***}
	(0.056)	(0.045)	(0.050)
Adjusted	-0.113^{*}	-0.007	$-$-0.126**
	(0.057)	(0.045)	(0.047)
Panel C. Intensive margin			
Unadjusted	-2.302^{+}	-2.253^{**}	-0.049
	(1.324)	(0.731)	(1.129)
Adjusted	-2.204^{*}	-2.427^{**}	-0.22
	(1.090)	(0.739)	(0.89)

Notes: Following our theoretical considerations, we can decompose the spillover effects into a pure spillover effect and the crowding effects. Spillover effects are decomposed by the means of the perceived pressure variable, showing average effects (Panel A), effects at the extensive margin (Panel B), and effects at the intensive margin (Panel C). The decomposition is conducted by the perceived pressure variable as inquired in the questionnaire (unadjusted) and the adjusted perceived pressure variable, which corrects for sample selection effects (adjusted). The coefficients are retrieved from regression Table 15, Table 16 and Table 17. The extensive margins are given by average marginal effects from a probit model. Further information on the definition and measurement of the variables can be retrieved from Table 6 in Appendix A. Robust standard errors in parentheses.

0.16, p-value=0.003, Table 15), indicating that the probability of donating to charity is reduced by 16 percentage points if participants previously perceived pressure to behave pro-social compared to participants in *NoSocial*. With a reduction of 11 percentage points, the effect is slightly weakened when controlling for selection effects (0.32 vs. 0.21, p-value=0.028, Table 15). Analyzing spillover effects on the intensive margin, we observe that the donation behavior of donors involves sizable spillover effects of -2.30€ in the unadjusted case (7.74 vs. 5.44, p-value=0.087, Table 15) and of -2.20€ for adjusted external pressure (7.74 vs. 5.54, p-value=0.0468, Table 15).
Our decomposition allows us to disentangle the spillover effects into a pure spillover effect and a crowding effect. Concerning the pure spillover effect, we observe a negative effect of -0.65€ (Table 5 – Panel A). This indicates that participants' who felt relatively low degrees of external pressure to contribute to a pro-social cause in Task I of the experiments donate 26% less than participants in *NoSocial* (2.48 vs. 1.83, p-value=0.0427, Table 16). Controlling for selection effects raises the pure spillover effect to 0.75€ , which is equivalent to a reduction of donations in Task II of 30% (2.48 vs. 1.74, p-value=0.0427, Table 16). This leads to our next observations

OBSERVATION 6. *Participants who completed a pro-social task, but perceived only a low level of external pressure to act pro-social, donate significantly less than participants who did not complete a pro-social task previously.*

The analysis on the extensive and intensive margin of donations reveals that the pure spillover effect is largely driven by the decisions of donors. Since the difference in the propensity to donate of participants, who previously behaved pro-social in Task I and those who did not, tends towards zero (0.32 vs. 0.33, p-value=0.777, Table 16), we find no evidence for a pure spillover effect on the extensive margin. In contrast, analyzing the behavior of donors shows that donors, who previously performed a pro-social task give significantly less to charity subsequently compared to participants in *NoSocial*, amounting to a difference in donations of 2.25€ (unadjusted: 7.74 vs. 5.48, p-value=0.0024; adjusted: 7.74 vs. 5.31, p-value=0.0012, Table 16).
As the second component of the spillover effect, we observe the pressure induced changes in subsequent donations by analyzing the crowding effect. The respective coefficient reaches a statistically significant level of -0.95 (1.82 vs. 0.88, p-value=0.0133, Table 17). Adjusting for selection effects renders the coefficient to -0.65€ (1.74 vs. 1.15, p-value=0.0785, Table 17). This is summarized in our last observation

OBSERVATION 7. *Participants in our pro-social treatments, who perceived a high level of external pressure to act pro-social in Task I of the experiment, donate significantly*

less in Task II than those participants, who did perceive only a low level of external pressure in Task I.

The analysis of the crowding effect on the extensive and intensive margin of donations shows that, in contrast to the pure spillover effect, the crowding effect is largely driven by participants' propensity to donate. The propensity to contribute to charity increases by 19 percentage points on average if participants report low degrees of perceived pressure compared to participants perceiving high degrees of pressure in Task I (0.33 vs. 0.16, p-value=0.0005, Table 17). However, when controlling for selection effects, the coefficient reduces by almost a halve (0.32 vs. 0.21, p-value=0.0068, Table 17). This indicates that a part of the crowding effect on the extensive margin is caused by participants sorting themselves into the high and low pressure group by their pro-social attitudes. In contrast, pressure to behave pro-social has only small effects on the level of donations provided by donors, as can be retrieved from the coefficients for the crowding effect in Panel C (5.49 vs. 5.44, p-value=0.966, Table 17). Yet, the crowding effect on the intensive margin becomes larger in absolute values if we correct for selection effects, but remains insignificant (5.31 vs. 5.54, p-value=0.802, Table 17).

The findings suggests a substantial crowding-in of the pure spillover effect due to an increase in participants' perceived external pressure to act pro-social in the first task of the experiment. This observation runs counter our hypothesized effect of external pressure on subsequent donations as crowding rather amplify negative spillover effects instead of attenuating them.

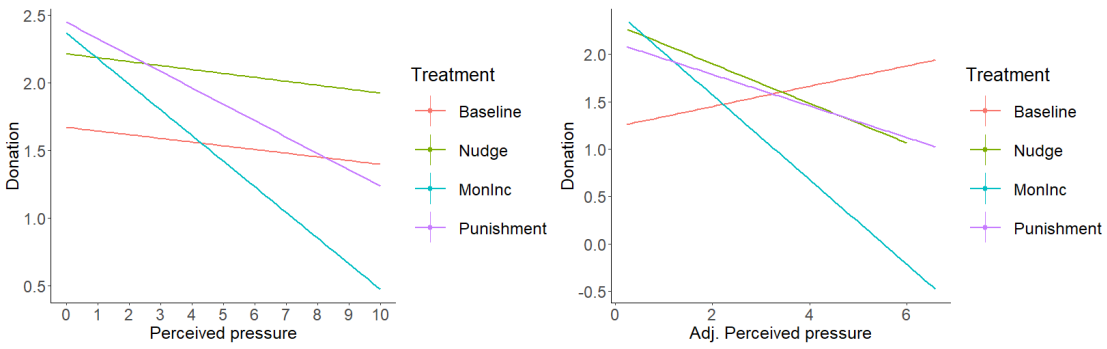
All these findings contribute to our understanding of the treatment and spillover effects in Section 5.1. Among all pro-social treatments, we observe the smallest spillover effects in absolute terms in *Nudge*. The analysis of participants' perceived pressure to act pro-social in Task I provides insights on the reasons for the small spillover effect in *Nudge*. First, the revealed negative correlation between perceived pressure in Task I and donations in Task II suggests that the level of subsequent donations increases if an incentive imposes low degrees of external pressure on participants. Second, participants perceive the lowest degrees of pressure when the pro-social task in Task I is incentivized by a *Nudge*. This even holds, when comparing pressure in *Nudge* to *Baseline*, in which no additional pressure through incentive schemes is applied on participants. This indicates that the reason for the low spillover effects observed in *Nudge* is given by the low degree of external pressure in this treatment, which inhibits the pressure related adverse effects on subsequent donations. In the remaining two interventional treatments, the stronger spillover effects are likely to be driven by intensified pressure levels. Yet, the observed degrees of perceived pressure seem not to directly translate into spillover effects observed

in *Monetary Incentive* and *Punishment*, as observed pressure levels are the highest in *Punishment*, whereas the strongest spillover effects are observed in *Monetary Incentive*. This can be explained by assessing the effect of perceived pressure on subsequent donations within each treatment separately. Figure 4a and Figure 4b show the fitted lines of the regression of donations in Task II on perceived pressure and adjusted perceived pressure by treatments.[32] The steep fitted regression line for the *Monetary Incentive* treatment in both graphs indicates that in *Monetary Incentive* a marginal increase in perceived pressure has stronger negative effects on subsequent donations than in other treatments, leading to a stronger crowding-in of negative spillover effects for the same levels of external pressure compared to other treatments. Hence, also the high level of negative spillover effects in the *Monetary Incentive* treatment can partly be explained by external pressure, as participants' subsequent donation decision is highly sensitive to pressure involving monetary stakes in a previous pro-social task. This amplifies negative spillover effects in the *Monetary Incentive* treatment.

Figure 4: Perceived pressure by treatments

(a) Panel A: Perceived pressure

(b) Panel B: Adjusted pressure



Note: The graphs show the fitted regression lines of the estimates of donations in Task II on perceived pressure and adjusted pressure. The perceived pressure variable measured in the post-questionnaire on a 0 to 10 scale is used in the left graph. Adjusted perceived pressure, which corrects for selection effects, is used in the right graph. The corresponding estimations can be retrieved from Table 19 and Table 18

[32]The corresponding estimates can be retrieved from Table 19 and Table 18 in the Appendix B.3

6 Conclusion

We study repeated pro-social behavior in a two-period experiment on charitable giving. Most importantly, our experiment allows us to identify whether and, if so, how the extent to which participants' pro-social behavior spills over from the first to the second decision depends on how the first decision has been incentivized. These are our central results: First, we find that pure spillover effects are negative, i.e., participants who contribute to a pro-social cause in the first task of the experiment give comparably low amounts to charity in the second task compared to those who did not contribute to a pro-social cause in the first task. On average, participants reduce their donations by 39 percent, if the first task contains a pro-social component. Second and most importantly, we find that the extent to which participants' pro-social decisions spill over from the first to the second task depends on how the decision in the first task has been incentivized. Spillover effects are especially pronounced in our *Monetary Incentive* treatment. If the initial pro-social contributions are incentivzed by monetary rewards, participants reduce their subsequent contributions to charity by 48 percent on average compared to participants, whose initial task has no pro-social component. Similar to this, in the *Punishment* treatment, in which time punishments are used to additionally incentivize Task I, participants reduce subsequent donations by 33 percent on average compared to participants who do not behave pro-social previously. In our *Nudge* treatment, which displays other participants' pro-social performance in the first task, in contrast, we find low and statistically insignificant spillover effects.

In line with our theoretical considerations, we find that participants' pro-social behavior negatively affects their subsequent donation decisions and that the extent of the spillover effects vary across our interventional treatments. However, our theoretical considerations cannot explain why the spillover effects are particularly weak in our *Nudge* treatment. Contrary to our theoretical predictions, we find that the spillover effects in our *Nudge* treatment are weaker compared to the effects in *Punishment* and *Monetary Incentives*. Our analysis shows, that this can be explained by how participants perceive external pressure exerted in our different interventional treatments and how this degree of perceived pressure affects the extent to which pro-social behavior spills over. At first, we find that the level of perceived pressure is particularly low in our *Nudge* treatment. Secondly, we find that participants' level of perceived pressure negatively affects donation levels in the second task of the experiment. In addition, we find that these adverse effects of external pressure on subsequent donations tends to be attenuated in Nudge.

Our experiment cannot directly answer the question why we observe the particularly low

level of perceived pressure in our *Nudge* treatment. Possible reason for the low pressure levels and the subsequently large amounts donated could be found in two underlying mechanisms of social comparisons. First, they change the perception of social norms (Bartke et al. 2017; Bicchieri and Dimant 2019). Second, social comparisons introduce non-rewarded competition between agents which increases pro-social performance (Duffy and Kornienko 2010). Social comparisons, as applied in the experiment, convey to participants that others do engage in pro-social causes and often to a higher degree than themselves. Thereby, social comparisons update participants' beliefs on the social norm of pro-social behavior. Participants intending to act in accordance with this social norm could have driven the responses on low pressure levels and the corresponding high levels of subsequent donations in *Nudge*. Another explanation for the low degree of perceived pressure in *Nudge* might be found in the competitive environment of the social comparison nudge. According to the concept of gamification (e.g., Wee and Choong 2019; Ouariachi et al. 2020), a competitive environment to solve a task might represent an enjoyable diversion for a certain share of participants. This also could have contributed to the low degree of perceived pressure.

In this line, two different effects could help to explain why participants' perceived level of external pressure increases the extent to which their pro-social behavior spills over: an annoyance effect and salience effect. The annoyance effect (e.g. Meer and Rosen 2012, Damgaard and Gravert 2018) captures the idea that external pressure leads to discomfort of participants while performing the pro-social behavior, which drives the negative effect of pressure on subsequent donations. Participants, who perceive discomfort through pressure, associate this with the institution carrying out the pressure. This makes them less prone to respond to a renewed appeal by this institution. Another explanation could be based on participants' attention during Task I.[33] As high degrees of pressure force participants to pay attention to the procedure of Task I, the related pro-social contribution in the first task becomes more salient in the subsequent donation decision. This pressure induced salience might have amplified the negative impact of the pro-social contributions in Task I on the subsequent donation levels, as the former pro-social behavior weighs larger for pressured individuals in the second decision.

Our results complement the existing literature on spillovers of pro-social behavior by showing that the extent of negative spillovers increases with external pressure imposed in the initial pro-social task. This suggests that the positive effect of incentives on subsequent pro-social decisions, as observed in e.g., Landry et al. (2010); Gallier et al. (2017);

[33] salience of a former choice can influence current choices as shown by the model of Bordalo et al. 2020

Gallus (2017); d'Adda et al. (2017), are largely driven by the incentive induced mark-up in donation levels, whose spillovers are not entirely crowded out by the incentivation induced pressure on pro-social behavior. This analysis particularly complements insights raised in d'Adda et al. (2017), who show that pushing participants to behave pro-social has a persistent positive effect on donations. Building up on this, our results suggest that push interventions lead to larger negative spillovers than less insisting incentives, due to the adverse effects of external pressure. This implies that push interventions require relatively strong initial incentive induced effects in order to lead to persistent effects compared to less insisting interventions.

The results are of use for practitioners in the context of pro-social behavior and charitable giving. They suggest that if policy makers, charities or NGOs aim at increasing pro-social behavior of a certain target group, they must be careful when applying compelling measures, as they are likely to have negative effects on pro-social behavior in subsequent decisions. This is particularly the case if the intervention is not capable of increasing initial pro-social behavior substantially in the first place. In the absence of such initial effects, the net-effect of these incentives are likely to be negative taking into account the adverse effect of pressure on donation decisions in subsequent periods. A nudge, instead, might have smaller effects on initial pro-social behavior, but provides donors with the notion of having contributed on a voluntary basis and thereby making them more prone to return and donate higher levels in future.

To draw implications from the experimental results, however, one should bear in mind the empirical evidences which are pending to be collected. First and foremost, even though a better understanding of behavioral spillover effects is essential to guarantee the long-term effectiveness of incentives on pro-social behavior, the absolute effect size tends to be rather small. Consequently, more large-scale studies are clearly needed to investigate and disentangle the different channels of how incentives of pro-social behavior in the past affects pro-social decisions in the present. Also, in the study, we used the controlled setting of an artefactual field experiment to rigorously test and identify the effect of incentives in repeated pro-social decisions. Although the design comes close to observing behavior in a natural environment, we cannot assert that similar effects occur in the field, when NGOs and charities apply incentives to raise contributions. Therefore, this requires a more detailed analysis in the field.

References

Adena, M. and Huck, S. (2019). Giving once, giving twice: A two-period field experiment on intertemporal crowding in charitable giving. *Journal of Public Economics*, 172:127–134.

Adena, M. and Huck, S. (2020). Online fundraising, self-image, and the long-term impact of ask avoidance. *Management Science*, 66(2):722–743.

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.

Baca-Motes, K., Brown, A., Gneezy, A., Keenan, E. A., and Nelson, L. D. (2013). Commitment and behavior change: Evidence from the field. *Journal of Consumer Research*, 39(5):1070–1084.

Bartke, S., Friedl, A., Gelhaar, F., and Reh, L. (2017). Social comparison nudges—guessing the norm increases charitable giving. *Economics Letters*, 152:73–75.

Bem, D. J. (1972). Self-perception theory. *Advances in experimental social psychology*, 6(1):1–62.

Bicchieri, C. and Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public choice*, pages 1–22.

Blanken, I., van de Ven, N., and Zeelenberg, M. (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin*, 41(4):540–558.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention, and choice. *The Quarterly Journal of Economics*, 1399:1442.

Brañas-Garza, P., Bucheli, M., Espinosa, M. P., and García-Muñoz, T. (2011). Moral cleansing and moral licenses: experimental evidence. *Economics and Philosophy*, 29(2):199–212.

Carpenter, J. and Matthews, P. H. (2017). Using raffles to fund public goods: Lessons from a field experiment. *Journal of Public Economics*, 150:30–38.

Charness, G. and Gneezy, U. (2009). Incentives to exercise. *Econometrica*, 77(3):909–931.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

City of Mannheim (2020). Statistische Daten 2019, Mannheim. Accessed: 11/13/2020.

Clot, S., Grolleau, G., and Ibanez, L. (2013). Self-licensing and financial rewards: is morality for sale? *Economics Bulletin*, 33:2298–2306.

Clot, S., Grolleau, G., and Ibanez, L. (2016). Do good deeds make bad people? *European Journal of Law and Economics*, 42(3):491–513.

Croson, R. and Shang, J. (2011). Social influences in giving: Field experiments in public radio. *The science of giving: Experimental approaches to the study of charity*, pages 65–80.

Damgaard, M. T. and Gravert, C. (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics*, 157:15–26.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2013). The importance of being marginal: Gender differences in generosity. *American Economic Review*, 103(3):586–90.

Duffy, J. and Kornienko, T. (2010). Does competition affect giving? *Journal of Economic Behavior & Organization*, 74(1-2):82–103.

d'Adda, G., Capraro, V., and Tavoni, M. (2017). Push, don't nudge: Behavioral spillovers and policy instruments. *Economics Letters*, 154:92–95.

Effron, D. A. and Monin, B. (2010). Letting people off the hook: When do good deeds excuse transgressions? *Personality and Social Psychology Bulletin*, 36(12):1618–1634.

Federal Statistical Office, G. (2020). Bevölkerung nach Bildungsabschluss in Deutschland. Accessed: 11/13/2020.

Ferraro, P. J., Miranda, J. J., and Price, M. K. (2011). The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment. *American Economic Review*, 101(3):318–22.

Fielding, D., Knowles, S., and Robertson, K. (2020). Materialists and altruists in a charitable donation experiment. *Oxford Economic Papers*, 72(1):216–234.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. *American Economic Review*, 94(5):1717–1722.

Funder, D. C. and Colvin, C. R. (1991). Explorations in behavioral consistency: properties of persons, situations, and behaviors. *Journal of personality and social psychology*, 60(5):773.

Gallier, C., Reif, C., and Römer, D. (2017). Repeated pro-social behavior in the presence of economic interventions. *Journal of behavioral and experimental economics*, 69:18–28.

Gallus, J. (2017). Fostering public good contributions with symbolic awards: A large-scale natural field experiment at wikipedia. *Management Science*, 63(12):3999–4015.

Gee, L. K. and Meer, J. (2019). The altruism budget: Measuring and encouraging charitable giving. Working Paper 25938, National Bureau of Economic Research, Cambridge.

GfK (2019). Bilanz des helfens. Policy brief, Deutscher Spendenrat, Berlin.

Ghesla, C., Grieder, M., and Schmitz, J. (2019). Nudge for good? choice defaults and spillover effects. *Frontiers in psychology*, 10:178.

Gneezy, A., Imas, A., Brown, A., Nelson, L. D., and Norton, M. I. (2012). Paying to be nice: Consistency and costly prosocial behavior. *Management Science*, 58(1):179–187.

Gneezy, U., Imas, A., and Madarász, K. (2014). Conscience accounting: Emotion dynamics and social behavior. *Management Science*, 60(11):2645–2658.

Goette, L. and Stutzer, A. (2020). Blood donations and incentives: Evidence from a field experiment. *Journal of Economic Behavior & Organization*, 170:52–74.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4):1009–1055.

Ito, K., Ida, T., and Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, 10(1):240–67.

Jessoe, K., Lade, G. E., Loge, F., and Spang, E. (2017). Spillovers from behavioral interventions: Experimental evidence from water and energy use. Working Paper E2e 033, Iowa State University Digital Repository, Ames.

Kang, M. J. and Glassman, M. (2010). Moral action as social capital, moral thought as cultural capital. *Journal of Moral Education*, 39(1):21–36.

Karlan, D. and List, J. A. (2007). Does price matter in charitable giving? evidence from a large-scale natural field experiment. *American Economic Review*, 97(5):1774–1793.

Kelley, H. H. (1973). The processes of causal attribution. *American psychologist*, 28(2):107.

Kesternich, M., Römer, D., and Flues, F. (2019). The power of active choice: field experimental evidence on repeated contribution decisions to a carbon offsetting program. *European Economic Review*, 114:76–91.

Khan, U. and Dhar, R. (2006). Licensing effect in consumer choice. *Journal of marketing research*, 43(2):259–266.

Krieg, J. and Samek, A. (2017). When charities compete: A laboratory experiment with simultaneous public goods. *Journal of behavioral and experimental economics*, 66:40–57.

Lacetera, N., Macis, M., and Slonim, R. (2014). Rewarding volunteers: A field experiment. *Management Science*, 60(5):1107–1129.

Landry, C. E., Lange, A., List, J. A., Price, M. K., and Rupp, N. G. (2006). Toward an understanding of the economics of charity: Evidence from a field experiment. *The Quarterly journal of economics*, 121(2):747–782.

Landry, C. E., Lange, A., List, J. A., Price, M. K., and Rupp, N. G. (2010). Is a donor in hand better than two in the bush? evidence from a natural field experiment. *American Economic Review*, 100(3):958–83.

List, J. A., Murphy, J. J., Price, M. K., and James, A. G. (2019). Do appeals to donor benefits raise more money than appeals to recipient benefits? evidence from a natural field experiment with pick. click. give. Woriking Paper 26559, National Bureau of Economic Research, Cambridge.

Mazar, N. and Zhong, C.-B. (2010). Do green products make us better people? *Psychological science*, 21(4):494–498.

Meer, J. and Rosen, H. S. (2012). Does generosity beget generosity? alumni giving and undergraduate financial aid. *Economics of Education Review*, 31(6):890–907.

Meijers, M. H. C. et al. (2014). *On justifying eco-unfriendly behaviors*. Universiteit van Amsterdam [Host].

Merritt, A. C., Effron, D. A., and Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and personality psychology compass*, 4(5):344–357.

Monin, B. and Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of personality and social psychology*, 81(1):33.

Mullen, E. and Monin, B. (2016). Consistency versus licensing effects of past moral behavior. *Annual review of psychology*, 67.

Ouariachi, T., Li, C.-Y., and Elving, W. J. (2020). Gamification approaches for education and engagement on pro-environmental behaviors: Searching for best practices. *Sustainability*, 12(11):4565.

Sass, M., Timme, F., and Weimann, J. (2015). The dynamics of dictator behavior. Working Paper 5348, Center for Economic Studies & Ifo Institute, Munich.

Scharf, K. A., Smith, S., and Wilhelm, M. (2017). Lift and shift: the effect of fundraising interventions in charity space and time. Working Paper 12338, CEPR, London.

Schmitz, J. (2019). Temporal dynamics of pro-social behavior: an experimental analysis. *Experimental Economics*, 22(1):1–23.

Seils, E. and Baumann, H. (2019). WSI Verteilungsmonitor. Policy Brief 14582, WSI, Hans Boeckler Stiftung, Duesseldorf.

Tiefenbeck, V., Staake, T., Roth, K., and Sachs, O. (2013). For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, 57:160–171.

Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., and Vandenbergh, M. P. (2014). Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change*, 29:127–138.

Wee, S.-C. and Choong, W.-W. (2019). Gamification: Predicting the effectiveness of variety game design elements to intrinsically motivate users' energy conservation behaviour. *Journal of environmental management*, 233:97–106.

Appendix A Variable definitions

Table 6: Description of variables

Variable	Explanation	Scale
Donation	Aggregated contributions to charities in Task II.	0-12
Recall	Participants are asked whether they remember how many trees they planted in the initial pro-social task and whether they remember the name of the afforestation organisation.	1-Both wrong 2-one of two correct 3-both correct
Depletion	Participants shall evaluate if they perceived the comprehension questions as effortful.	1-5
Perceived Pressure	Assesses whether participants completed the comprehension task due to intrinsic or extrinsic motivation using a likert sclae ranging from Purely intrinsically motivated equalling 0 and purely driven by external factors equalling 10.	1-10
Trust	Three questions about trust were raised in the questionnaire comprising trusting other people in general, carefulness when encountering strangers and reliance on others. Each contains 4 possible agreement choices comprising of "fully disagree", "slightly disagree", "neither nor", "slightly agree", "fully agree" and a no statement option.	1-5
Environmental Attitudes	Contains 9 questions on environmental attitudes retrieved from the short version of the New Ecological Paradigm. Participants had the option to choose "fully disagree", "disagree", "neither nor", "agree", "fully agree" and a no statement option	1-5
Environmental Identity	Controls whether participants regard themselves as environmentalists by inquiring the application of the following statements to oneself: Environmentally friendly behaviour is an important part of what defines me; I try, through my actions (e.g. consumption), to minimize negative impacts on the environment;I define myself as an environmentally conscious person. Participants can choose from "Fully applies","Applies","Neither nor","Does not apply","Does not apply at all" and a no statement option	1-5
Environmental Efficacy	Tests for participants' self-efficacy perceptions in their pro-environmental behavior. Inquires agreement to the following four questions: "Environmental and climate measures can actively halt climate change.", "I think that I have the ability to take some measure for the protection of the climate and the environment.", "Even if it creates inconveniences, I am able to change my behavior in order to protect the environment and the climate.", "I can try to do my best in every imagenable way to protect the environment and the climate.". Participants can choose from "Fully applies","Applies","Undecided","Does not apply","Does not apply at all" and a no statement option	1-5

Pro-Social Attitudes	Inquires agreements with the following three pro-social sentences: "Every person should share their income with others"; "We must help people in need"; "Everybody should be socially committed". The options to answer comprised of "fully disagree", "disagree", "neither nor", "agree", "fully agree" and a no statement option.	1-5
Former donation	Asks whether participants have donated to charity in the year 2018. The question states: "Have you made donations of over 10€in the last year?	Yes-no
Past donation	How much participants have donated to charity in the year 2018. The question states: "How high is the sum of these donations approximately?". Appearance of question conditional on a "Yes" in *Former donation*.	$0-\infty$
Risk revealed	Assesses participants' risk attitudes via Lottery. Participants were asked for their willingness to pay for a lottery with 50% chance of winning two Euros or nothing. The choice is consequential. For further details see Appendix **??**.	0-2
Risk Perception	Assesses participants' risk attitudes via own judgement. The question states: "How do you evaluate yourself personally: Are you in general a risk-taking person, or do you try to avoid risks?". Participants are able to respond to a Likert scale ranging from 0="Not willing to take risks at all" to 10="Very willing to take risks".	0-10
Ambiguity revealed	Assesses participants' ambiguity attitudes via Lottery. Participants were asked for their willingness to pay for a lottery with unknown probability of winning two Euros or nothing. The choice is consequential. For further details see Appendix **??**.	0-2
Ambiguity Perception	Assesses participants' ambiguity attitudes via own judgement. The question states: "How do you evaluate yourself personally: How good can you cope with uncertainty, e.g. lack of information, when making a decision?". Participants are able to respond to a Likert scale ranging from 0="Not good at all" to 10="Very good".	0-10
Female	If participants' reported gender is female	1-3
Age	Age of Participant	16-99
Native speaker	If participant has German as their mother tongue	0 (no) 1 (yes)
Student	Whether the participants is currently enrolled at a university or a university of applied sciences.	Yes-No
LeftVoters	Having voted for a left wing party in the past elections (either "Bündnis 90, die Grünen" or "Die Linke")	1-7
VoteGreen	Having voted for a pro-environmental party in the past elections ("Bündnis 90, die Grünen")	1-7
Income	Income of participants in €(1000€Intervals)	1-6
Education	Asks for the highest level of education	1-7

Donation receipt	Asks whether participants wish to receive a donation receipt on the entire amounts donated within the course of the study.	Yes-No
Attempts	Attempts needed to answer the five comprehension questions in the quiz.	5-∞
Mistakes	Mistakes made by participants when answering the five comprehension questions in the quiz.	0-∞
Time Task1	Time it took participants to answer the questions in the quiz (in seconds)	0-∞
Time Instr.	Time spent on the information provision page (in seconds)	0-∞
Time Feedback	Time spent on the feedback page (in seconds)	0-∞
Part. Time	Time of the day the participant took part in the study, differentiated by hours	0-24
Part. Day	Day of Participation, differentiated by days at which the experiment took place	1-8

Appendix B Additional Results

B.1 Performance Indicators

Performance Indicators

The intervention treatments are designed to increase effort in the pro-social real effort task. In order to investigate, whether the different treatments lead to behavioral changes in the processing the real effort task, we analyze the influence of the intervention treatments on seven different parameters used as dependent variables in Table 7. In the first column the treatment indicators of the intervention treatments are regressed on mistakes, which serves as an accuracy measure and reflects the number of wrongly answered questions. We observe that all three treatments reduced the number of correctly answered questions, with the lowest and insignificant value of a reduction of 0.56 mistakes in *Nudge*. The treatments *Monetary Incentive* and *Punishment* report a reductions in mistakes made by participants by one question on average in *Monetary Incentive* and 1.4 questions on average in *Punishment*. Therefore, the results speak in favor of an increase in accuracy by participants induced by regulation. The dependent variables in column two concern the time spent to answer the control questions. There is mixed evidence on time required to answer the control questions. If regulation was in place, participants in *Nudge* and *Monetary Incentives* needed more time compared to participants, who did not receive any incentives (vice versa in *Punishment*). This is not surprising as more effort generally reduces the time to answer the questions, whereas processing the

questions with greater care and accuracy is more time consuming. Yet, we conjecture that incentives let to higher attention to the questions, causing more time for processing, while stricter regulation reduces the time needed for the control questions. The latter explanation is supported by the results in column three, in which the treatment dummies are regressed on the time participants read through the information necessary to answer the questions in the effort task. The coefficients for *Monetary Incentive* is significant and positive, while the coefficient in *Punishment* is the second largest, followed by *Nudge*. This indicates that participants read through the information with greater care in order to avoid negative consequences from the incentives in the real effort task. Similar results are observed in column four, five and six, showing the treatment effects on time spent on the feedback page (time feedback) and whether participants correctly remembered the amount of trees planted as well as the name of the afforestation organization (recall). These three dependent variables serve as proxies for attention provided in the initial pro-social task. We observe that participants spent more time on the feedback page in *Monetary Incentive* and *Punishment* than in *Baseline* and performed better in recalling the facts from the initial pro-social task in *Monetary Incentive* and *Punishment* compared to *Nudge*. Lastly, column seven tests for differences between incentive treatments in perceived depletion in the comprehension question.[34] We observe that depletion was the largest in *Monetary Incentives*, followed by *Punishment* and *Nudge*. Hence, the performance variables show little effect of the *Nudge* treatment. Although the coefficients of the treatment indicator in *Nudge* points towards larger accuracy provided, the coefficients are not significant. In contrast, the treatment indicators for *Monetary Incentive* and *Punishment* suggest that participants took greater care in answering the questions and performed significantly better in responding to the comprehension questions. Furthermore, there are indications of a greater provision of attention.

[34]Depletion was inquired in the post-questionnaire, see Table A

Table 7: OLS-regression of treatment indicators on various performance variables

	Dependent variable:						
	Mistakes	Time Task1	Time Instr.	Time Feedback	Recall		Depletion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baseline						0.048	0.077
						(0.100)	(0.098)
Nudge	−0.560	16.711	4.186	1.802	−0.048		
	(0.612)	(16.118)	(5.858)	(5.055)	(0.100)		
Monetary Inc.	−0.966	0.603	12.213*	8.001^{+}	0.286**	0.334***	0.255**
	(0.591)	(15.235)	(5.537)	(4.773)	(0.094)	(0.099)	(0.095)
Punishment	−1.428*	−5.738	8.026	7.254	0.151	0.199^{+}	0.101
	(0.611)	(15.800)	(5.742)	(4.943)	(0.098)	(0.102)	(0.098)
Control Variables	X	X	X	X	X	X	X
Observations	529	366	366	368	368	368	515
R^2	0.124	0.199	0.059	0.038	0.144	0.144	0.124

Note: Further information on the definition and measurement of the variables can be retrieved from Table 6 in Appendix A. Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01;***p<0.001

B.2 Donations to charity across treatments

Table 8: Donations by treatments

Treatment	NoSocial	Baseline	Nudge	Mon.Inc.	Punishment
Panel A. Average donations					
Mean (in €)	2.48	1.50	2.16	1.29	1.65
Median (in €)	0	0	0	0	0
Panel B. Extensive margin					
Mean (in €)	0.32	0.27	0.35	0.29	0.28
Median (in €)	0	0	0	0	0
Panel C. Intensive margin					
Mean (in €)	7.74	5.51	6.09	4.49	5.95
Median (in €)	10	3	5	2	5
N	49	40	45	43	39

Table 9: OLS-regression of treatment effects on donation behavior

	Dependent variable: Donations in Task II							
	(1)	(2)			(3)	(42)		
	OLS	Hurdle			OLS	Hurdle		
	donations	Two stage			donations	Two stage		
		Extensive Margin		Intensive Margin		Extensive Margin		Intensive Margin
		Coefficients	Average margina effects			Coefficients	Average margina effects	
Baseline	−0.973*	−0.137	−0.048	−2.227*	−1.005*	−0.168	−0.056	−2.359*
	(0.429)	(0.155)	(0.054)	(0.943)	(0.434)	(0.163)	(0.054)	(0.924)
Nudge	−0.319	0.094	0.033	−1.651+	−0.302	0.100	0.033	−1.686+
	(0.443)	(0.156)	(0.054)	(0.908)	(0.440)	(0.162)	(0.054)	(0.873)
Monetary Incentive	−1.190**	−0.095	−0.033	−3.251***	−1.047*	−0.028	−0.009	−3.129***
	(0.424)	(0.152)	(0.053)	(0.919)	(0.425)	(0.158)	(0.052)	(0.890)
Punishment	−0.823+	−0.121	−0.042	−1.794+	−0.739+	−0.102	−0.034	−1.651+
	(0.438)	(0.157)	(0.055)	(0.957)	(0.436)	(0.164)	(0.054)	(0.934)
Past Donations					−0.0001	−0.0001	−0.0001	−0.001
					(0.0002)	(0.0001)	(0.0001)	(0.001)
Female					0.932**	0.249*	0.082*	1.001+
					(0.292)	(0.108)	(0.035)	(0.594)
Age					0.032**	0.0004	0.0001	0.086***
					(0.012)	(0.004)	(0.001)	(0.024)
Native Speaker					0.136	0.216	0.072	−0.721
					(0.539)	(0.197)	(0.065)	(1.056)
Student					0.279	0.097	0.097	0.264
					(0.319)	(0.118)	(0.118)	(0.642)
Education					0.184***	0.074***	0.024***	0.171+
					(0.044)	(0.016)	(0.005)	(0.101)
Part. Time					−0.002*	−0.001*	−0.0002*	−0.002
					(0.001)	(0.0003)	(0.0001)	(0.002)
Part. Day					−0.025	−0.033	−0.011	0.168
					(0.070)	(0.026)	(0.009)	(0.143)
Constant	2.477***	−0.468***		7.740***	0.818	−0.904+		4.976+
	(0.300)	(0.107)		(0.632)	(1.259)	(0.469)		(2.814)
Observations	703	703	703	212	676	676	676	209
R^2	0.015			0.060	0.092			0.186

Note: We use participants' donations in Task II of the experiment as dependent variable. Participants in *NoSocial* are our control group and we use indicators for our pro-social treatments as regressors.The regression containing the full set of control variables is presented in Table 12 in Appendix B.2; Further explanation of the control variables are provided in Table 6 in Appendix A. Robust standard errors in parentheses. +p<0.1; *p<0.05; **p<0.01; ***p<0.001.

Table 10: Crowding effects: OLS-Regression of interventional treatment effects on donations in Task II

	Dependent variable: Donations in Task II							
	(1)	(2)			(3)	(42)		
	OLS	Hurdle			OLS	Hurdle		
	donations	Two stage Extensive Margin		Intensive Margin	donations	Two stage Extensive Margin		Intensive Margin
		Coefficients	Average margina effects			Coefficients	Average margina effects	
Baseline	0.973*	0.137	0.137	2.227*	1.005*	0.168	0.168	2.359*
	(0.429)	(0.155)	(0.155)	(0.943)	(0.434)	(0.163)	(0.163)	(0.924)
Nudge	0.654	0.231	0.08	0.576	0.704	0.268	0.268	0.673
	(0.448)	(0.160)	(0.055)	(0.957)	(0.450)	(0.168)	(0.168)	(0.946)
Monetary Incentive	−0.217	0.041	0.014	−1.024	−0.042	0.139	0.139	−0.770
	(0.429)	(0.156)	(0.054)	(0.968)	(0.435)	(0.164)	(0.164)	(0.943)
Punishment	0.151	0.016	0.006	0.433	0.266	0.065	0.065	0.707
	(0.443)	(0.161)	(0.056)	(1.004)	(0.449)	(0.171)	(0.171)	(1.008)
Past Donations					−0.0001	−0.0001	−0.0001	−0.001
					(0.0002)	(0.0001)	(0.0001)	(0.001)
Female					0.932**	0.249*	0.082*	1.001+
					(0.292)	(0.108)	(0.035)	(0.594)
Age					0.032**	0.0004	0.0001	0.086***
					(0.012)	(0.004)	(0.001)	(0.024)
Native Speaker					0.136	0.216	0.072	−0.721
					(0.539)	(0.197)	(0.065)	(1.056)
Student					0.279	0.097	0.032	0.264
					(0.319)	(0.118)	(0.039)	(0.642)
Education					0.184***	0.074***	0.024***	0.171+
					(0.044)	(0.016)	(0.005)	(0.101)
Part. Time					−0.002*	−0.001*	−0.0002*	−0.002
					(0.001)	(0.0003)	(0.0001)	(0.002)
Part. Day					−0.025	−0.033	−0.011	0.168
					(0.070)	(0.026)	(0.009)	(0.143)
Constant	1.503***	−0.605***		5.513***	−0.187	−1.071*		2.617
	(0.307)	(0.112)		(0.701)	(1.283)	(0.479)		(2.850)
Control Variables					X	X	X	X
Observations	703	703	703	212	676	676	676	209
R^2	0.015			0.060	0.092			0.186
Log Likelihood		−428.889	−428.889			−394.822	−394.822	

Note: We use participants' donations in Task II of the experiment as dependent variable. Participants in *Baseline* are our control group and we use indicators for our interventional treatments as regressors. Further explanation of the control variables are provided in Table 6 in Appendix A. Robust standard errors in parentheses. +p<0.1; *p<0.05; **p<0.01; ***p<0.001.

Table 11: OLS-Regression on differences in Task II donations compared to *Nudge*

	Dependent variable: Donations in Task II							
	(1)	(2)			(3)	(42)		
	OLS	Hurdle			OLS	Hurdle		
	donations	Two stage			donations	Two stage		
		Extensive Margin		Intensive Margin		Extensive Margin		Intensive Margin
		Coefficients	Average margina effects			Coefficients	Average margina effects	
NoSocial	0.319	−0.094	−0.033	1.651^{+}	0.290	−0.129	−0.041	1.500
	(0.443)	(0.156)	(0.156)	(0.908)	(0.430)	(0.189)	(0.061)	(1.075)
Baseline	−0.654	−0.231	−0.080	−0.576	−0.705	−0.318^{+}	−0.102^{+}	−0.684
	(0.448)	(0.160)	(0.055)	(0.957)	(0.434)	(0.192)	(0.061)	(1.146)
Monetary Incentive	−0.871*	−0.189	−0.066	−1.601^{+}	−0.780^{+}	−0.100	−0.032	−1.611
	(0.443)	(0.157)	(0.055)	(0.933)	(0.429)	(0.186)	(0.060)	(1.055)
Punishment	−0.503	−0.215	−0.075	−0.143	−0.433	−0.162	−0.052	0.330
	(0.456)	(0.163)	(0.056)	(0.971)	(0.442)	(0.193)	(0.062)	(1.098)
Constant	2.157***	−0.374**	−0.374**	6.089***	0.557	−0.872	−0.872	2.483
	(0.326)	(0.114)	(0.114)	(0.652)	(1.221)	(0.564)	(0.564)	(3.552)
Control Variables					X	X	X	X
Missings imputed					X	X	X	X
Observations	703	703	703	212	703	541	541	158
R^2	0.015			0.060	0.092			0.185
Adjusted R^2	0.009			0.042	0.077			0.118
Log Likelihood		−428.889	−428.889			−307.543	−307.543	

Note: We use participants' donations in Task II of the experiment as dependent variable. Participants in *Nudge* are our control group and we use indicators for our interventional treatments as regressors. Further explanation of the control variables are provided in Table 6 in Appendix A. Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001.

Table 12: OLS-regression of treatment effects on donation behavior, including the entire set of control variables

	Dependent variable:							
	Donations (in €)		Extensive Margin				Intensive Margin	
			Coefficients	Average marginal effects				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Baseline	−0.908	−0.952*	−0.025	−0.008	−0.143	−0.045	−2.552+	−2.729**
	(0.646)	(0.421)	(0.262)	(0.079)	(0.165)	(0.052)	(1.516)	(0.958)
Nudge	−0.977	−0.276	0.157	0.048	0.114	0.036	−3.461*	−1.798+
	(0.671)	(0.431)	(0.267)	(0.080)	(0.165)	(0.052)	(1.543)	(0.915)
Monetary Incentive	−0.975	−1.045*	0.145	0.044	−0.057	−0.018	−4.204**	−3.226***
	(0.632)	(0.412)	(0.253)	(0.076)	(0.160)	(0.051)	(1.467)	(0.939)
Punishment	−0.824	−0.715+	−0.228	−0.068	−0.111	−0.035	−0.542	−1.639+
	(0.638)	(0.425)	(0.261)	(0.078)	(0.165)	(0.052)	(1.610)	(0.970)
Recall		0.536**			0.202*	0.063*		0.519
		(0.202)			(0.079)	(0.025)		(0.450)
LeftVoters	0.589	−0.068	−0.029	−0.009	−0.056	−0.017	1.503	−0.269
	(0.473)	(0.295)	(0.192)	(0.058)	(0.115)	(0.036)	(1.125)	(0.674)
Risk Perception	−0.030	−0.065	0.003	0.0008	−0.040+	−0.013+	0.062	0.019
	(0.085)	(0.054)	(0.035)	(0.01)	(0.021)	(0.007)	(0.207)	(0.128)
Environ. Att.	−0.130	0.185	0.144	0.043	0.112	0.035	−1.942	−0.152
	(0.579)	(0.356)	(0.241)	(0.073)	(0.141)	(0.045)	(1.789)	(0.956)
Env. Ident.	−0.702+	−0.189	−0.122	−0.037	−0.014	−0.004	−0.761	−0.271
	(0.379)	(0.237)	(0.156)	(0.047)	(0.095)	(0.030)	(0.897)	(0.550)
Env. Efficacy	0.930*	0.400	0.332+	0.1+	0.168+	0.053+	1.116	0.216
	(0.411)	(0.247)	(0.173)	(0.051)	(0.098)	(0.031)	(1.037)	(0.604)
Deontologist	0.440	0.176	−0.098	−0.03	−0.050	−0.016	0.449	−0.241
	(0.473)	(0.294)	(0.194)	(0.058)	(0.115)	(0.036)	(1.206)	(0.661)
Trust	0.600*	0.213	0.183	0.055	0.040	0.013	1.054	0.511
	(0.301)	(0.191)	(0.121)	(0.036)	(0.075)	(0.024)	(0.792)	(0.465)
Attempts	−0.087+	−0.006	−0.056*	−0.017*	−0.017	−0.005	−0.064	0.124+
	(0.051)	(0.033)	(0.028)	(0.008)	(0.014)	(0.004)	(0.174)	(0.071)
Past Donations	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	−0.0001	0.0002	−0.001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.002)	(0.001)
Depletion	−0.174	−0.340+	−0.045	−0.014	−0.099	−0.031	−0.228	−0.289
	(0.304)	(0.186)	(0.127)	(0.038)	(0.074)	(0.023)	(0.719)	(0.430)
Female	0.357	0.601*	0.113	0.034	0.088	0.027	0.391	0.956
	(0.468)	(0.306)	(0.188)	(0.056)	(0.117)	(0.037)	(1.159)	(0.688)
Age	0.035+	0.027*	0.003	0.001	0.002	0.001	0.078+	0.062*
	(0.019)	(0.012)	(0.008)	(0.002)	(0.005)	(0.002)	(0.045)	(0.027)
Native Speaker	−0.529	0.227	0.049	0.015	0.277	0.088	−0.133	−1.102
	(0.992)	(0.526)	(0.392)	(0.118)	(0.200)	(0.063)	(2.685)	(1.125)
Student	0.078	0.325	0.121	0.022	0.121	0.038	−0.565	0.327
	(0.490)	(0.324)	(0.196)	(0.009)	(0.125)	(0.039)	(1.154)	(0.680)

VoteGreen	−0.028	0.476	0.018	0.006	0.105	0.033	0.709	0.787
	(0.691)	(0.504)	(0.284)	(0.085)	(0.198)	(0.063)	(1.660)	(1.148)
Donation Receipt	0.233	−0.184	−0.055	−0.017	−0.206^{+}	−0.065^{+}	1.591	0.636
	(0.435)	(0.292)	(0.174)	(0.052)	(0.112)	(0.035)	(1.065)	(0.625)
Education	0.153*	0.145**	0.073*	0.073*	0.064***	0.020***	0.132	0.157
	(0.075)	(0.047)	(0.031)	(0.031)	(0.019)	(0.006)	(0.208)	(0.111)
Income	−0.152	0.057	−0.060	−0.018	0.0003	0.0001	−0.570	0.224
	(0.243)	(0.154)	(0.101)	(0.03)	(0.061)	(0.019)	(0.606)	(0.352)
Part. Time	−0.001	−0.002*	−0.001	−0.0002	−0.001*	−0.0002*	0.001	−0.003^{+}
	(0.001)	(0.001)	(0.0004)	(0.0001)	(0.0003)	(0.0001)	(0.003)	(0.002)
Part. Day	−0.316*	−0.033	−0.069	−0.021	−0.040	−0.013	−0.404	0.196
	(0.131)	(0.068)	(0.052)	(0.052)	(0.026)	(0.008)	(0.297)	(0.150)
Constant	0.037	−0.020	−1.404	−1.404	−1.070	−1.070	4.874	4.438
	(3.567)	(2.185)	(1.459)	(1.459)	(0.856)	(0.856)	(10.383)	(5.096)
Imputed Data-Set		X		X		X		X
Observations	317	703	317	317	703	703	94	212
R^2	0.163	0.118					0.358	0.225
Adjusted R^2	0.094	0.085					0.134	0.121
Log Likelihood			−168.863	−168.863	−392.103	−392.103		

Note: Further explanation of the control variables are provided in Table 6 in Appendix A. Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001; standard errors robust towards heteroscedasticity

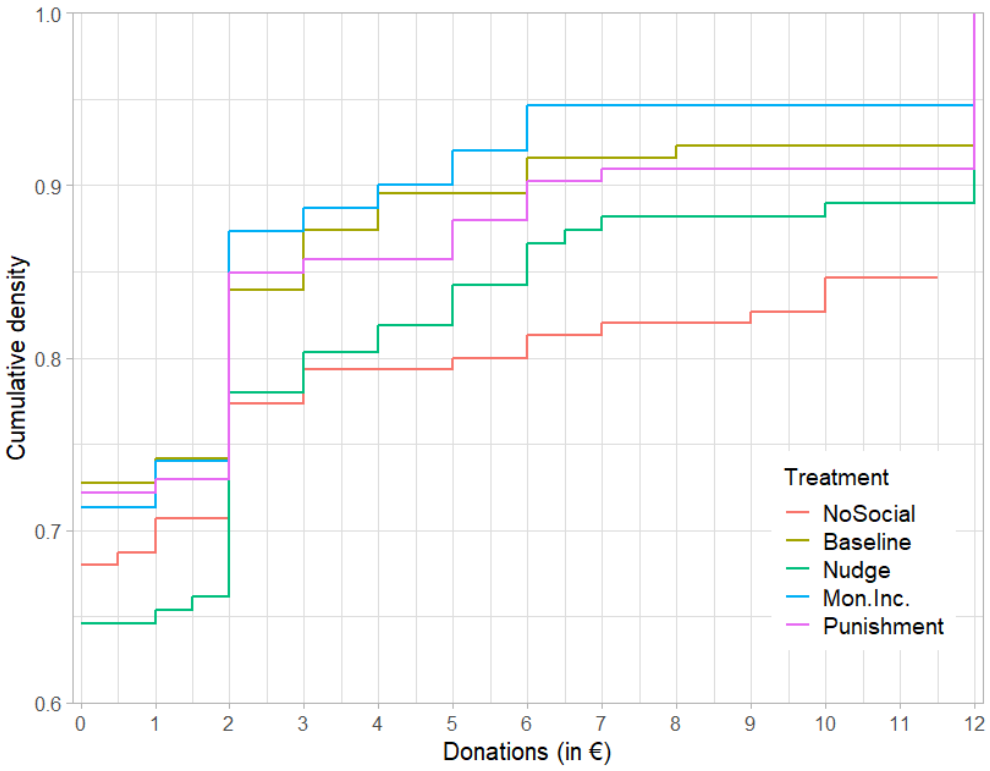
Cumulative distribution of donations by treatment

To provide an overview on donation behavior in *Baseline*, *Nudge*, *Monetary Incentive*, and *Punishment*, we show the cumulative distribution of donations for each of those treatments in Figure 5. The analysis at the extensive margin (see Table 8 –Panel B) reveals that there are minor differences in the propensity to give across treatments, which are not statistically different from each other. In *Nudge*, participants show the highest propensity to donate (35%). Participants in *NoSocial* show the second highest willingness to donate positive amounts to charity (33%). The propensity to give in *Baseline*, *Monetary Incentives*, and *Punishment* are clustered around 28%.

Analyzing the cumulative distribution functions at donation levels greater than zero reveals substantial treatment effects at the intensive margin (see Table 8 –Panel C). We find that participants in *NoSocial* donate their entire endowment twice as frequent compared to the pooled sample of participants in other treatments (*NoSocial* = 0.16, 'Others'=0.08, p-value=0.0198, M-W U test). In *NoSocial* (orange line), cumulative donations rise in a similar pattern compared to the other treatments up to donations of three Euros. Thereafter, the curve flattens and remains below the other cumulative

density functions, while increasing sharply at a value of 12€. In *Nudge* (green line), aggregated donation levels are larger compared to most other treatments. Participants are more willing to donate any positive amount and among those who donate, there is a higher frequency of participants donating large amounts of their endowment (d>9€, Nudge=0.11, Mon.Inc. or Baseline=0.06, p-value=0.1477, M-W U test). The differences in the course of the cumulative density function is relatively similar to the treatments *Baseline* and *Monetary Incentive*, but ranging on a constantly lower level. The course of the cumulative density function in *Punishment* (purple line) is more flat than in other treatments, indicating that those participants who donate, give larger amounts on average. The cumulative density function of the donations in *Baseline* (olive line) and *Monetary Incentives* (blue line) constantly range above the functions of donations in the other three treatments, indicating that participants in these treatments are not only the most hesitant to give any positive amount to charity, but also donating participants tend to give lower values to charity.

Figure 5: Cumulative distribution functions of donations by treatments



B.3 Perceived pressure and donations Part II

Table 13: Perceived Pressure by treatments

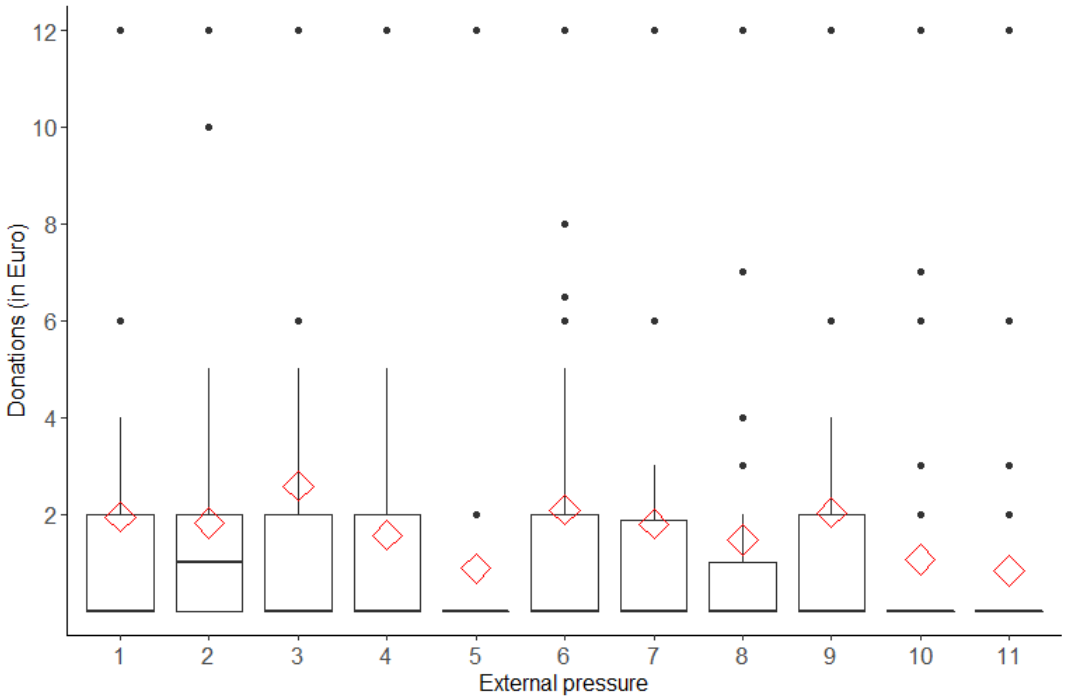
Treatment	Baseline	Nudge	Mon.Inc.	Punishment
Panel A. Perceived Pressure				
Mean (in €)	6.09	5.71	6.52	6.91
Median (in €)	6	6	7	7
Panel B. Adjusted Perceived Pressure				
Mean (in €)	2.37	2.07	2.50	2.58
Median (in €)	2.25	2	2.50	2.54

Table 14: OLS-regression on the effect of external pressure in Task I on donations in Task II

	Dependent variable:					
	Donation (in Euro)		Extensive Margin		Intensive Margin	
	(1)	(2)	(3)	(4)	(5)	(6)
External Pressure	−0.097*	−0.119*	−0.020**	−0.021**	0.046	−0.021
	(0.047)	(0.048)	(0.006)	(0.006)	(0.113)	(0.108)
Past donations		−0.0003		−0.00003		−0.001
		(0.0003)		(0.00004)		(0.001)
Female		0.921**		0.079+		1.290*
		(0.314)		(0.042)		(0.646)
Age		0.044***		0.002		0.096***
		(0.012)		(0.002)		(0.026)
Mother tongue		−0.431		0.013		−0.824
		(0.561)		(0.075)		(1.215)
Uni		0.165		0.049		−0.588
		(0.332)		(0.044)		(0.693)
Y.o.Education		0.140**		0.019**		0.107
		(0.046)		(0.006)		(0.111)
Time of day		−0.217**		−0.022*		−0.427*
		(0.079)		(0.011)		(0.189)
Day of part.		0.012		−0.007		0.202
		(0.075)		(0.010)		(0.158)
Constant	2.252***	1.783	0.425***	0.356+	5.221***	5.801+
	(0.333)	(1.384)	(0.044)	(0.185)	(0.724)	(3.402)
Observations	553	519	553	519	165	158
R^2	0.008	0.103	0.018	0.073	0.001	0.203

*Note:*Further information on the definition and measurement of the variables can be retrieved from Table 6 in Appendix A. Robust standard errors in parentheses. +p<0.1; *p<0.05; **p<0.01; ; **p<0.001

Figure 6: Donation behavior by degree of external pressure



Note: The vertical axis shows donations in Task II in Euros. The horizontal axis lists the reported perceived pressure ranging from 1 to 11. Mean values displayed by red rhombus. Median values are reported by the big horizontal bar. The boxes lower and upper ends show the 25th and 75th percentile.

Table 15: Determination of Spillover effects by the perceived pressure variable

	Dependent variable:							
	(1)	(2)		(3)	(4)	(5)		(6)
	Donations	Extensive margin		Intensive margin	Donations	Extensive margin		Intensive margin
	OLS	*probit*		*OLS*	*OLS*	*probit*		*OLS*
		Coefficients	Average margina effects			Coefficients	Average margina effects	
Pro-social treatments	-1.598^{**}	-0.520^{**}	-0.163^{**}	-2.302^{+}	-1.329^{**}	-0.348^{*}	-0.113^{*}	-2.204^{*}
	(0.498)	(0.185)	(0.056)	(1.324)	(0.456)	(0.162)	(0.051)	(1.090)
Constant	2.477^{***}	-0.468^{***}	-0.468^{***}	7.740^{***}	2.477^{***}	-0.468^{***}	-0.468^{***}	7.740^{***}
	(0.314)	(0.107)	(0.107)	(0.662)	(0.314)	(0.107)	(0.107)	(0.662)
Cutoff: Perceived Pressure	X	X	X	X				
Cutoff: Adj. Perceived Pressure					X	X	X	X
Observations	249	249	249	64	285	285	285	76
R^2	0.040			0.047	0.029			0.052
Adjusted R^2	0.036			0.031	0.026			0.040
Log Likelihood		−137.822	−137.822			−162.948	−162.948	

*Note:*The sample for the estimates is restricted to participants in the *NoSocial* treatment and participants in the pro-social treatments, who perceived a high degree of external pressure in Task I. In the regression models, donations in Task II is regressed on participants, who perceived a high degree of external pressure in Task I, with participants in the *NoSocial* treatment as a comparison group. In the first four columns, perceived pressure is used as measured in the post-questionnaire. In the last four columns, adjusted perceived pressure is used, which corrects for selection effects. The dependent variable is the amount donated in Task II of the experiment. Models (2) and (5) uses a probit model to estimate effects on the extensive margin, reporting the models' coefficients in column 2 and 6 and average marginal effects in column 3 and 7. Models (3) and (6) estimate effects on the intensive margin of donations, where the dependent variable is the amount donated in Task II of the experiment, conditional on donations being positive. Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table 16: Determination of the pure spillover effect by the perceived pressure variable

	Dependent variable:							
	(1)	(2)		(3)	(4)	(5)		(6)
	Donations	Extensive margin		Intensive margin	Donations	Extensive margin		Intensive margin
	OLS	*probit*		*OLS*	*OLS*	*probit*		*OLS*
		Coefficients	Average margina effects			Coefficients	Average margina effects	
Pro-social treatments	-0.682^{+}	0.017	0.006	-2.233^{**}	-0.729^{*}	0.021	0.008	-2.401^{**}
	(0.359)	(0.123)	(0.044)	(0.731)	(0.363)	(0.125)	(0.045)	(0.739)
Constant	2.477^{***}	-0.468^{***}	-0.468^{***}	7.740^{***}	2.477^{***}	-0.468^{***}	-0.468^{***}	7.740^{***}
	(0.311)	(0.107)	(0.107)	(0.636)	(0.309)	(0.107)	(0.107)	(0.632)
Cutoff: Perceived Pressure	X	X	X	X				
Cutoff: Adj. Perceived Pressure					X	X	X	X
Observations	604	604	604	196	547	547	547	178
R^2	0.006			0.046	0.007			0.057
Adjusted R^2	0.004			0.041	0.006			0.051
Log Likelihood		−380.642	−380.642			−345.078	−345.078	

Note: The sample for the estimates is restricted to participants in the *NoSocial* treatment and participants in the pro-social treatments, who perceived a low degree of external pressure in Task I. In the regression models, donations in Task II is regressed on participants, who perceived a low degree of external pressure in Task I, with participants in the *NoSocial* treatment as a comparison group. In the first four columns, perceived pressure is used as measured in the post-questionnaire. In the last four columns, adjusted perceived pressure is used, which corrects for selection effects. The dependent variable is the amount donated in Task II of the experiment. Models (2) and (5) uses a probit model to estimate effects on the extensive margin, reporting the models' coefficients in column 2 and 6 and average marginal effects in column 3 and 7. Models (3) and (6) estimate effects on the intensive margin of donations, where the dependent variable is the amount donated in Task II of the experiment, conditional on donations being positive. Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table 17: Determination of the crowding effect by the perceived pressure variable

	Dependent variable:							
	(1)	(2)		(3)	(4)	(5)		(6)
	Donations	Extensive margin		Intensive margin	Donations	Extensive margin		Intensive margin
	OLS	*probit*		*OLS*	*OLS*	*probit*		*OLS*
		Coefficients	Average margina effects			Coefficients	Average margina effects	
Perceived Pressure$_{High}$	-0.916^{*}	-0.537^{***}	-0.182^{***}	-0.069				
	(0.379)	(0.163)	(0.054)	(1.132)				
Adj. Perceived Pressure$_{High}$					-0.600^{+}	-0.369^{**}	-0.126^{**}	0.197
					(0.336)	(0.138)	(0.047)	(0.892)
Constant	1.795^{***}	-0.451^{***}	-0.451^{***}	5.507^{***}	1.748^{***}	-0.447^{***}	-0.447^{***}	5.338^{***}
	(0.160)	(0.061)	(0.061)	(0.354)	(0.169)	(0.065)	(0.065)	(0.375)
Observations	553	553	553	164	532	532	532	158
R^2	0.011			0.00002	0.006			0.0003
Adjusted R^2	0.009			-0.006	0.004			-0.006
Log Likelihood		-330.403	-330.403			-319.966	-319.966	

Note: The sample for the estimates is restricted to individuals in the pro-social treatments. In the regression models, donations in Task II is regressed on participants, who perceived a high degree of external pressure in Task I, with participants, who perceived a low degree of pressure as a comparison group. The dependent variable is the amount donated in Task II of the experiment. Models (2) and (5) uses a probit model to estimate effects on the extensive margin, reporting the models' coefficients in column 2 and 6 and average marginal effects in column 3 and 7. Models (3) and (6) estimate effects on the intensive margin of donations, where the dependent variable is the amount donated in Task II of the experiment, conditional on donations being positive. Robust standard errors in parentheses. $^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table 18: Effect of Perceived Pressure on donations in Part II by pro-social treatment

	Dependent variable:			
	Donations (in €)			
	(Baseline)	(Nudge)	(Mon. Inc.)	(Punishment)
Perceived Pressure	0.106	-0.228	-0.442^{*}	-0.167
	(0.204)	(0.287)	(0.175)	(0.221)
Constant	1.241^{*}	2.440^{***}	2.430^{***}	2.095^{**}
	(0.560)	(0.684)	(0.501)	(0.653)
Observations	136	122	146	128

Note: Robust standard errors in parenthesis. $^{+}$p<0.1; *p<0.05; **p<0.01; ; **p<0.001

Table 19: Effect of Adj. Perceived Pressure on donations in Part II by pro-social treatment

	Dependent variable:			
	Donations (in €)			
	(Baseline)	(Nudge)	(Mon. Inc.)	(Punishment)
Adj. Perceived Pressure	−0.028	−0.024	−0.189*	−0.117
	(0.095)	(0.114)	(0.078)	(0.098)
Constant	1.676*	2.293**	2.520***	2.461**
	(0.645)	(0.737)	(0.558)	(0.740)
Observations	143	127	150	133
R^2	0.001	0.0003	0.039	0.011

Note: Robust standard errors in parenthesis. $^{+}$p<0.1; *p<0.05; **p<0.01; ; **p<0.001

Supplementary Material available via:
http://ftp.zew.de/pub/zew-docs/dp/appendix/DPAltGallier2021_Supplemental_Material.pdf

Download ZEW Discussion Papers from our ftp server:

http://ftp.zew.de/pub/zew-docs/dp/

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de