NOTA DI LAVORO
108.2015

The not so Gentle Push: Behavioral Spillovers and Policy Instruments

Giovanna d'Adda, Politecnico di Milano
Valerio Capraro, Center for Mathematics and Computer Science (CWI)
Massimo Tavoni, Politecnico di Milano, Fondazione Eni Enrico Mattei (FEEM) and Euro-Mediterranean Center on Climate Change (CMCC)

Climate Change and Sustainable Development
Series Editor: Carlo Carraro

The not so Gentle Push: Behavioral Spillovers and Policy Instruments

By Giovanna d'Adda, Politecnico di Milano
Valerio Capraro, Center for Mathematics and Computer Science (CWI)
Massimo Tavoni, Politecnico di Milano, Fondazione Eni Enrico Mattei (FEEM) and Euro-Mediterranean Center on Climate Change (CMCC)

Summary

We examine whether spillovers of pro-social behavior depend on how behavioral changes are induced. We conduct a large experiment using economic games, with a Dictator Game (DG) followed by either an identical game or a Prisoner's Dilemma (PD). We influence initial behavior through widely used policy instruments, either behaviorally informed (default, social norms) or with an economic/regulatory rationale (incentives, regulation). Our results provide evidence of positive spillovers to subsequent economic games (which are not treated), but only for the traditional economic/regulatory interventions and within the same game type. Specifically, inducing higher giving in the first stage leads to subsequent higher altruism in the DG, but not to more cooperation in the PD. The carry over of pro-social behavior appears to be driven by an anchoring on the initial donation. We also measure observers' beliefs and we find that these results are not correctly anticipated by third parties, who systematically overestimate both the direct effect of behaviorally informed interventions on initial donations and their spillover to subsequent donations.

Keywords: Pro-social Behavior, Traditional and Behavioral Policies, Spillover Effects, Online Experiment

JEL Classification: H4, I3

Address for correspondence
Massimo Tavoni
Fondazione Eni Enrico Mattei
Corso Magenta 63
20123 Milano
Italy
E-mail: massimo.tavoni@feem.it

The not so gentle push: behavioral spillovers and policy instruments

Giovanna d'Adda*, Valerio Capraro[†]and Massimo Tavoni [‡]

October 12, 2015

Abstract

We examine whether spillovers of pro-social behavior depend on how behavioral changes are induced. We conduct a large experiment using economic games, with a Dictator Game (DG) followed by either an identical game or a Prisoner's Dilemma (PD). We influence initial behavior through widely used policy instruments, either behaviorally informed (default, social norms) or with an economic/regulatory rationale (incentives, regulation). Our results provide evidence of positive spillovers to subsequent economic games (which are not treated), but only for the traditional economic/regulatory interventions and within the same game type. Specifically, inducing higher giving in the first stage leads to subsequent higher altruism in the DG, but not to more cooperation in the PD. The carry over of pro-social behavior appears to be driven by an anchoring on the initial donation. We also measure observers' beliefs and we find that these results are not correctly anticipated by third parties, who systematically overestimate both the direct effect of behaviorally informed interventions on initial donations and their spillover to subsequent donations.

Keywords: Pro-social behavior; Traditional and behavioral policies; Spillover effects; Online experiment

*Department of Management and Economics, Politecnico di Milano, Italy

[†]Center for Mathematics and Computer Science (CWI), Amsterdam, The Netherlands

[‡]Department of Management and Economics, Politecnico di Milano. Fondazione Eni Enrico Mattei (FEEM) and Euro-Mediterranean Center on Climate Change (CMCC), Italy

1   Motivation

Evaluation of policy interventions is widely recognized as a valuable tool to inform policy design ([87]; [1]). Generally, rigorous evidence, on which the relevance, efficiency and effectiveness of policies is based, does not extend beyond the specific scope of the intervention under study. However, the extent to which the effects of public policy interventions spill across contexts can greatly influence their evaluation.

Pro-social behavior is one realm where the presence of spillovers is likely to be particularly relevant. Individuals interact strategically in many different settings and the same normative value can influence behavior in different contexts and groups. For instance, evidence from economic experiments shows that previously established cooperative precedents lead to higher cooperation in subsequent games ([60]; [30]; [17]). The presence of spillovers of pro-social behavior has large policy implications as well: for instance, a key policy question in the environmental economics literature is whether interventions that foster cooperative norms in the conservation of a specific environmental resource have positive or negative spillovers to other pro-environmental behaviors ([32, 90]).

The sign and magnitude of policy spillovers may depend on the instrument used to induce behavioral changes. In recent years, innovative policy interventions based on the insights of social psychology and behavioral economics have attracted increasing interest by legislators. These policy interventions complement traditional ones based on economic incentives or regulation, and have now seen application in many fields, including to a large extent pro-social and environmental behavior ([2]; [18, 86]). However, while comparisons of traditional and behavioral policy instruments exist in the literature ([53]; [36]), little is known about how treatment effects would carry forward outside the specific area they were designed for [32]. Moreover, most existing studies compare only few policy instruments within the same design, and, to our knowledge, no assessment of the relative effectiveness of a broader set of behavioral and traditional policies has yet been conducted.

We aim to address these gaps in the literature. In this paper we measure the direct and spillover effects of policy on pro-social behavior, examine whether they depend on the type of instrument used - considering four of the most common behavioral and traditional policy tools- and evaluate whether they extend beyond the context directly targeted by the intervention. We tackle these questions through a large scale experiment conducted online on Amazon Mechanical Turk[72], using a pool of US workers. The design is articulated in 2 stages. In the first stage, subjects participate in a Dictator Game (DG). We randomly assign participants to one of four different policy interventions, in addition to the baseline case: *defaults, social norms, economic incentives, and regulation.* Subjects subsequently play a second game, randomly chosen between another DG and a continuous Prisoner's Dilemma (PD).The design thus allows us to examine how behavioral norms induced in the first stage spill over to the second stage, both when the relevant norm remains the same (DG) and when it changes from fairness to efficiency (PD), and whether such spillover effects depend on first stage treatment.

2

Our main result is that, at least in the abstract setting of an anonymous experiment, traditional economic interventions (incentives and regulation) perform better than psychological interventions, both in terms of their direct effectiveness and in terms of spillovers. We provide evidence that the effect is mostly attributable to the higher capacity of traditional instruments to change initial behavior. Somewhat surprisingly, we also find that these results are not correctly anticipated by observers. Specifically, we measure third parties' beliefs and we find that they systematically overestimate the effect of behaviorally informed interventions, both at the level of initial donations and in terms of their spillovers.

In the next section we review the literature on spillovers and provide some background into economic and psychological findings on the effectiveness of each policy lever examined within our design. Section 3 describes our experimental design. Section 4 presents experimental results, and Section 5 concludes.

2 Background

We first review the literature on spillovers, then discuss evidence on each policy lever examined within our design and review the literature evaluating the relative effectiveness of different policies.

2.1 Spillovers

Following ([90]), we define spillover as an effect of an intervention on subsequent behaviors not targeted by the intervention. Therefore, we refer to 'spillovers' both when the initial and subsequent behavior belong to the same context and when they do not. Recent reviews of the experimental literature have highlighted a variety of different possible mechanisms that can explain the presence of spillovers, in both directions[32, 90, 13]. Moral licensing, moral cleansing and ego depletion are examples of psychological mechanisms that can explain the presence of negative spillovers: actions or primes that improve one's self-image or deplete one's self-control may be conducive to less ethical behavior. Cognitive dissonance and foot-in-the-door effects are instead examples of mechanisms explaining positive spillovers, which rely on individuals' desire for self-consistency.[1] In [12], the authors have introduced a theoretical economic framework which can reconcile both positive and negative spillovers. In this model, moral values are inferred by individuals from their behavior: identity-enhancing behaviors are affected by small differences in the decision context, and depend on prior confidence in being a moral person. This theory thus introduces history-dependence in behavior and reconciles many divergent experimental findings on behavioral spillovers.

Consistent with the theoretical literature, empirical studies offer evidence of both negative and positive spillovers. In the realm of negative spillovers, psychologists have documented the moral licensing effect by priming the moral self-concept positively (negatively) and observing a subsequent

---

[1]We refer to [90] for references on the different mechanisms behind spillovers.

3

decrease (increase) in pro-sociality [68, 70, 80, 56, 43]. On the other hand, recent studies show that costly investments in social image generate positive spillovers in terms of generosity later on [42], that cooperative behavior learned in a particularly cooperative setting spills over to a number of other settings involving pro-sociality [75] and that different pro-social behaviors are all positively correlated [22, 21, 74]. Experimental evidence from sequential and simultaneous games shows that behavior is influenced by play in other contexts, and attribute such spillover effects to learning and to the use of heuristics ([26]).

The present study contributes to the literature on spillovers by investigating the implications of different policy interventions on subsequent behavior, particularly focusing on the different impact of behavioral and traditional policy instruments. In principle, behavioral interventions aimed at promoting intrinsic motivation should have better chances of inducing positive spillovers than traditional policy instruments, such as extrinsic economic incentives that have been shown to lead to "crowding out" effects (see[16] for a review of the large experimental literature). However, to our knowledge no previous work has systematically tested this hypothesis experimentally.

2.2 Impact of traditional and behavioral interventions

The present study compares the relative effectiveness of four common policy instruments: defaults, norms, monetary incentives, and regulations. We begin by providing an overview of the literature on these four tools and their effectiveness, with the aim of putting our experiment into context.

Defaults refer to the option, which individuals receive if they fail to make an active choice. Defaults have become a very popular policy tool after numerous studies have shown their effectiveness in influencing behavior in a wide range of settings, from organ donation ([54]), to retirement savings ([27]; [65]), to insurance choice ([55]) and choice of green energy provider ([76]). In spite of the fact that default effects are stronger when applied to situations that are complex or unfamiliar to the decision maker ([88]), laboratory evidence from simple decision environments provides evidence of the power of defaults ([14]; [49]), although these studies generally use more prominent forms of defaults than we do in the present study. This literature offers contrasting evidence on the spillover effects of defaults: while [31] find that defaults set in the first round affect directly round one contributions in a DG and persist in the following rounds through status quo bias, or anchoring, [20] show no spillover effects of defaults in a public good game setting.

Norm-based messages and social comparisons are other non-monetary strategies that are increasingly being used to influence individual decision making. A growing literature supports this policy trend, by demonstrating the effect of pro-social messages and social comparison on charitable giving ([19]; [78]; [40]; [66]; [82]) and pro-environmental behavior ([52]; [64]; [92]; [84]; [85]; [81]; [47]; [71]). An influential study within this literature is by [2], which shows how providing utility customers with information on their water use relative to their neighbors leads to energy savings among

high consumption users. Coupling such comparative information with normative judgments and making relative performance public increase the effect of comparative information ([47];[29]). Laboratory evidence also supports the effectiveness of social norms: contributions to public goods, dictator and ultimatum games are influenced by information on other subjects' individual or average decisions ([10]; [28]; [38]; [58]; [91]; [25]; [15]).

The literature on monetary incentives is too large to be covered exhaustively, so here we focus only on studies concerned with the effect of incentives on pro-social behavior. A common view is that economic incentives may backfire in the presence of intrinsic motivation to perform a task ([39]). Field experiments on blood donations ([69]; [61]; [46]), pro-environmental behavior ([79]; [41]), pro-social work tasks ([8]), charitable giving ([5]; [63, 57]), voluntary work ([45]) and norm violation ([44]) yield mixed evidence on the motivation crowding effects of incentives. Such sensitivity to the context is confirmed by laboratory studies on incentives and pro-social behavior, which offer examples of crowding out, as well as crowding in.[2]. Rebates, similar to the ones we use in our design, are a widely used policy tool to foster charitable giving and pro-environmental behavior, and have been shown to be effective in inducing higher transfers in dictator games in the lab and to charity in the field ([4]; [63]).

Regulations that limit individuals' choice sets produce results similar to incentives. Regulations, fines and monitoring are found to backfire when intrinsic motivation is relevant to the task ([9]). Laboratory experiments where subjects can impose minimum contribution rules or fine low contributors also demonstrate that control is costly, as those who exert it earn lower payoffs ([24]; [35]; [34]). The counterproductive effect of regulations is attributed to the fact that they signal mistrust on the part of the agents imposing them, and reduce the autonomy of control averse individuals ([16]). An implication of the signaling theory is that endogenously adopted regulations should have stronger negative effects than exogenously imposed ones. However, [23] finds similar effects of exogenous and endogenous fines. Our design implements regulation as minimum contribution rules, similar in spirit to [34], exogenously imposed by the experimenter.

Only a relatively small number of studies compare different policy tools, focusing instead on the effect of behavioral and traditional versions of the same intervention, or of combining different tools. An example of the former set of studies is [7], which examines the relative effect of monetary and symbolic rewards on social workers' motivation. [36], which compares the effect of normative messages with and without a social comparison feature and finds the former to be more effective, provides an example of the latter set of studies. Comparison of different policy instruments generally occurs between pairs of alternatives. A recent example of comparison of different policy tools is [53], which tests the relative impact of moral suasion, a type of normative message, and monetary incentives on energy use during peak times, finding the latter to be more effective in generating sustained behavioral change. Research on spillover effects of policy within economics focuses on measuring unintended consequence of a policy intervention ([37]; [89]), but does not normally compare the relative effect of

---

[2]For a review of the evidence and for an explanation of such contrasting results, see [16]

different policy treatments in inducing spillovers. The psychology literature offers an example of study that compares the spillover effect on pro-environmental behaviors of inducing green purchases through monetary incentives or moral nudges, and finds the former to be more effective[62]. To the best of our knowledge, no study exists that compares the direct and spillover effects of a wide range of policy instruments.

3   Design and implementation

As mentioned above, we implemented a large scale experiment articulated in two stages in Amazon Mechanical Turk (AMT) [72], using a pool of US workers. AMT allows to have larger and more representative samples than laboratory experiments, at the expenses of less control of the experimenter over participants. In this paper we use two standard economic games -a dictator and a prisoner dilemma game. Several papers have replicated the main findings and general results of these experiments using online platforms such as AMT ([73];[51];[3];[48];[59]). The consistency between data gathered on AMT and those collected in physical labs goes beyond incentivised economic games: it also regards self-reported measures, even though providing random responses to unincentivised questions minimizes the time needed to complete the survey without having any consequence on earnings ([77]; [67]; [83]; [50]).

Overall, we recruited 3267 participants: 564 participated in the pilot study and 2703 in the main experiment, 1322 as decision makers and 1381 as recipients. From the sample of recipients we also elicited beliefs on behavior of decision makers in the main experiment.

The pilot study consisted in a single DG with two experimental treatments, differing in the size of the participation fee (50 and 70 cents). This test was conducted in order to check for the presence of income effects in DG contributions, so as to rule them out as potential causes of treatment effects in stage 2 of the main study.

All participants in the main experiment received \$0.50 as participation fee, in addition to the earnings from the tasks. Subjects in the role of decision makers participated in a first stage Dictator Game (DG), and subsequently played a second game, randomly chosen between another DG and a continuous Prisoner's Dilemma (PD). Participants who played as decision makers in Stage 1, played as decision makers also in Stage 2.

In the first stage, dictators had an endowment of \$0.20 and were asked to decide how much of it (in fractions of 2 cents), to give to the recipient, if any. We randomly assigned participants to one of four different policy interventions, in addition to the baseline case: *defaults, social norms, economic incentives, and regulation.* For each of them, we created a focal donation point. In the *defaults* case, donations were pre-set at 50% of the endowment by simply pre-marking the button corresponding to a 50% contribution. This is a mild intervention, since the player could choose a different contribution with very limited effort. Under the *norms* treatment, we provided subjects with

a social reference by informing them of the behavior of others subjects participating in the pilot: in particular, subjects were told that roughly half of the previous donations were equal or above 50% (this was the actual outcome, no deception was used). *Incentives* rewarded subjects for giving at least half of their endowment, by returning half of the donation to the donors when they gave 50% or more. Finally, *regulation* established a minimum donation rule, set at 50% of the endowment. Participants were not allowed to proceed if their donations were strictly below 50%. Roughly half of the subjects assigned to the role of decision-makers in the experiment faced these treatments, which we refer to as the *high focal point* treatments. In order to disentangle the effect of the policy instrument from the effect of the level of donation induced in the first stage, we simultaneously conducted an additional set of four treatments, where defaults, norms, incentives and regulation were used to foster a lower level of donations, with a *low focal point* at 10% of the endowment.

The second stage game was either a standard DG or a PD, without any policy intervention. We clarified that the second stage was independent from the first one and played with a different partner. The second stage DG was identical to the baseline first stage DG for all subjects. Participants playing the PD in the second stage had an endowment of \$0.10, and were asked how much of it, if any, they wanted to give to the other person (in steps of 1 cent): the amount given would be multiplied by 2 and earned by the other person. The payoff in the continuous PD for subject i, matched with subject j, was thus given by the following formula $p_i = (10 - d_i) + 2 * d_j$, where d is the amount donated by a subject and 10 is the size of the initial endowment for the task in cents.

Finally, subjects assigned to the role of receivers in the main experiment faced two belief elicitation tasks. In each treatment, participants were grouped with two other participants, person A (a dictator) and person B (another recipient), were shown the screenshots of the instructions received by person A, and were asked to guess person A's donation to Person B. This procedure was used to elicit beliefs both on the first and on the second stage DG. Stage 1 treatments mirrored the ones facing the dictators. Correct guesses were incentivised with a \$0.20 reward. This design allows us to observe if beliefs on DG donation in stage 1 and 2 varied between subjects depending on how giving was encouraged in the first stage game.

All subjects participating in the main experiment completed a survey at the end of the second stage. The survey contained questions on basic demographic and socioeconomic characteristics, as well as questions on social preferences. In particular, we used a set of experimentally validated survey questions designed by [11] to measure individual altruism, negative and positive reciprocity.

Table 1 summarizes our treatments, and displays the number of donors for each of them. We adopted a series of checks on IP addresses and Turk IDs to ensure that no subject took the experiment more than once. Participants had to answer correctly a series of comprehension questions to be admitted to the decision stage of the experiment. In the first stage, participants were asked two questions about the donation maximizing their own and the other participant payoff in the DG. Roughly 73% of participants responded correctly to both questions. In the second stage, if randomized

		Control	50% Focal Point				10% Focal Point			
			Default	Norms	Incentives	Regulation	Default	Norms	Incentives	Regulation
Stage 1	DG	130	162	153	138	155	153	144	140	147
Stage 2	DG	80	96	92	81	93	91	84	82	96
	PD	50	66	61	57	62	62	60	58	51

Table 1: List of treatments and number of contributors per treatment. For DG games, we recruited an equal number of recipients.

to the PD, subjects were asked four additional comprehension questions about the rules of the game, with an average success rate of 70%. These different success rates explain why the sample size in the second stage PD is smaller than in the second stage DG. However, this differential attrition has not led to statistically significant differences between the two sub-samples along observable characteristics, as shown in Table 8 in the Appendix. Table 10 reports the success rate in the comprehension questions by treatment. Full experimental instructions are reported in the Appendix.

4 Results

We now describe the results of the experiment. We begin with the direct effects of treatments, then discuss the spillover effects on DG contribution and PD cooperation. We focus on the treatments with a focal point of 50%. The low focal point (10%) treatments are discussed in the robustness section subsubsection 4.2.1. We conclude the discussion of the results with the guesses of treatment effects.

4.1 Direct effects

Figure 1 and Table 2 report average donations (normalized to 10) in Stage 1. Average giving in the first stage DG is 26.7% in the baseline treatment, in line with the results from the many dictator games published so far[33]. Looking at the treatments, donations are 28.2 % in the default one, 32.3% when norms of fairness are made salient, 46.1% under economic incentives, and 55.3% for the regulation treatment. Therefore, in terms of direct effectiveness of policy instruments, traditional interventions -incentives and regulation- yield the higher impact on average donation, by almost doubling it (2-sided t-test; all p's $< .0001$). On the other hand, we only find a marginally significant positive effect of the norm treatment ($p = 0.0711$) and no effect of the default treatment ($p = 0.6160$).

To ensure comparability between the incentive treatment and the other treatments, throughout the analysis we report DG donations as the amount received by the receiver, when not otherwise stated. When relevant, we discuss treatment effects on net donations (donations minus rebate) for the incentive treatments in the main text and Appendix. For instance, in terms of direct effects, we find that the difference in gross donations between the incentive and baseline treatments is driven, and actually paid, by the institution: subtracting the amount reimbursed from the actual donation, i.e.

looking at net donations, we find that donors, on average, give only 23.5% of their endowment, which is not statistically different from the amount given in the baseline ($p = 0.2490$).

Table 2 provides additional details on the distribution of individual donations. Contributions in the first stage DG tend to cluster around two modes, at 0% and 50% of the endowment, reproducing results obtained in the laboratory using the same game [33]. The only exception is of course the regulation treatment, with a uni-modal distribution due to the minimum threshold. Donations above 50% of the endowment are more frequent in the regulation than in the baseline treatment, suggesting that the minimum contribution rule did not crowd out intrinsic motivation to give[3]. Other treatment effects appear to be the shift in the main mode from 0 in the baseline to 50% in the other treatments, and that in intermediate donations from below 50% in the baseline, default and norms treatments, to above 50% in the incentive one. Overall, we observe a much stronger effect of traditional policy instruments, such as incentives and regulation, than of behaviorally-informed ones, like defaults and norms.[4] These results are not particularly surprising, given the rigidity of aggregate behavior in the simple setting of the Dictator game, polarized around either payoff maximizing or equal split allocation strategies, and the mildness of the behavioral treatments relative to the incentive and regulation ones.

---

[3]In the analysis, we measure crowding out through between subjects comparison of stage 1 behavior under each experimental treatment, relative to baseline. Spillover effects are instead measured through within subjects comparison of stage 1 and stage 2 behavior.

[4]Further details are provided Appendix A, whereFigure 5 shows the distribution of donations in the first stage, by treatment.
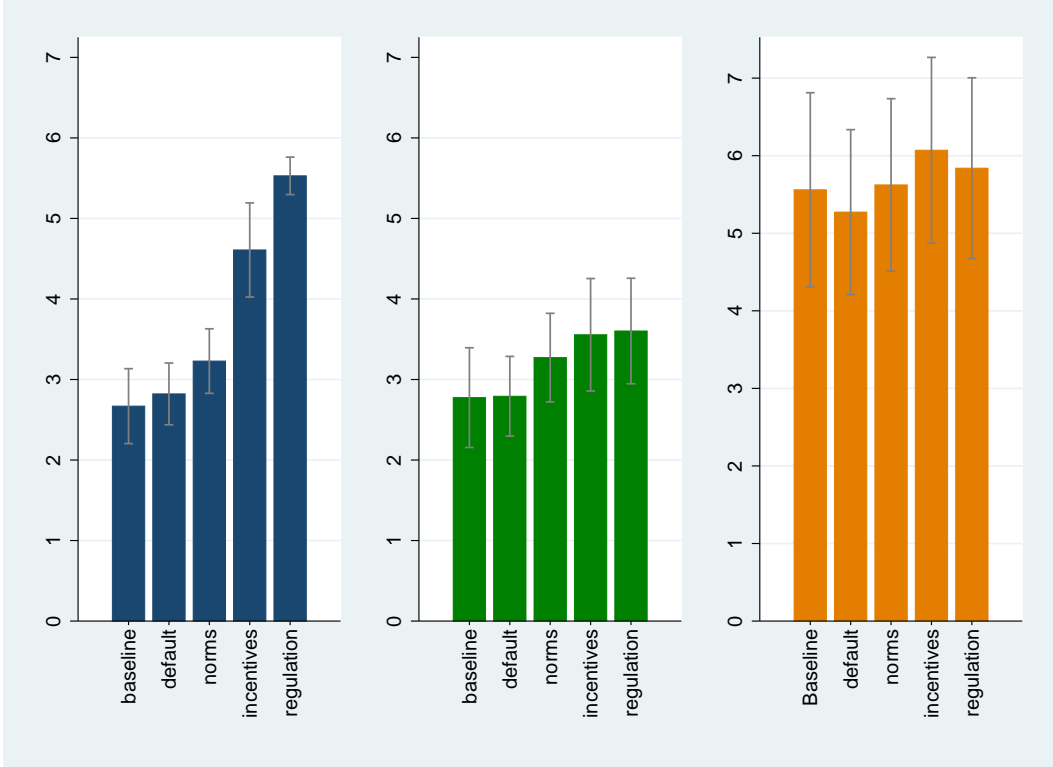
Figure 1: Average contributions (normalized to 10) in stage 1 DG (left panel), stage 2 DG (central panel) and stage 2 PD (right panel), by first stage treatments. Error bars represent 95% confidence intervals.

We use regression analysis to confirm the robustness of treatment effects to the inclusion of individual characteristics, and examine other determinants of contributions in the first stage DG. In order to do this, in Table 3 we pool the high focal point DG-DG and DG-PD treatments (442 and 296 subjects respectively, for a total of 738 observations), and run an OLS regression of stage 1 contributions on treatment dummies (Column 1) and demographic controls, such as age, gender, altruism, education and experience with similar tasks on AMT (Column 2). Regression results confirm the findings of Figure 1 in terms of treatment effects, which are statistically significant only for the traditional policy instruments. Consistent with the discussion above, the coefficient on the incentive treatment dummy is insignificant, once we consider net donations. As for individual characteristics, older subjects tend to give significantly more. Women appear to be more generous, but the effect is not significant, as not significant is the effect of having participated to similar studies on AMT before. The index of altruism, constructed from the experimentally validated survey questions, is significantly and positively linked to giving, while higher education levels are associated with lower altruism.

Regression analysis reported in Table 3 also helps us investigate the presence of crowding out of

10

	Average	**Zero**	**>Zero & < Half**	**Half**	**>Half & < All**	**All**
Stage 1 DG						
Baseline	2.67	43.1	13.8	39.2	1.5	2.3
Default	2.82	37	17.3	43.2	1.2	1.2
Norms	3.23	31.4	13.7	50.3	2.6	2
Incentives	4.61	28.3	2.2	40.6	10.1	18.8
Regulation	5.53	0	0	86.5	4.5	9
Stage 2 DG						
Baseline	2.77	43.7	11.3	38.7	3.8	2.5
Default	2.79	37.5	16.7	43.8	1	1
Norms	3.27	31.5	15.2	46.6	4.3	3.3
Incentives	3.55	35.8	9.9	39.5	4.9	9.9
Regulation	3.6	35.5	6.4	44.1	4.3	9.7
Stage 2 PD						
Baseline	5.56	32	2	20	4	42
Default	5.27	31.8	7.6	16.7	4.5	39.4
Norms	5.62	31.2	1.6	22.9	0	44.3
Incentives	6.07	31.6	1.7	12.3	1.8	52.6
Regulation	5.84	32.3	4.8	9.7	1.6	51.6

Table 2: Average contributions (normalized to 10) and distribution of contributions (in %) in stage 1 DG (top panel), stage 2 DG (middle panel) and stage 2 PD (bottom panel), by first stage treatments.

intrinsic motivation to contribute, generated by the treatments. Namely, if we consider the four treatments as embodying higher levels of control, from minimum in the baseline treatment to maximum in the regulation one, in the presence of crowding out we would expect a negative correlation between the level of control and the contribution. To test this hypothesis, we construct a variable (*Control)* equal to 0 for subjects in the baseline treatment, 1 for subjects in the default treatment, and so on, up to 4 for subjects in the regulation treatment. In the presence of crowding out, the coefficient on the control variable should be negative. On the contrary, Column 3 of Table 3 shows a positive and significant coefficient for the *Control* variable, an effect which is not mechanically driven by the regulation treatment alone, as shown in the same table in Column 4, where the same regression is run excluding it. However, if we consider net donation in the incentive treatment, the coefficient on the *Control* variable becomes statistically insignificant. Table 9 in the Appendix reproduces Table 3 using net instead of gross donations for the incentive treatment.

Overall, the regression results confirm that the differences in direct treatment effects on gross donations between behavioral and traditional policy instruments are not consistent with the presence of crowding out of intrinsic motivation due to the control exerted on individuals by the different treatments.

Dependent variable	Stage 1 contribution			
	All	**All**	**All**	**Excl. Regulation**
	(1)	**(2)**	**(3)**	**(4)**
Default	0.152	0.204		
	(0.305)	(0.290)		
Norm	0.560*	0.636**		
	(0.310)	(0.294)		
Incentive	1.939***	1.957***		
	(0.378)	(0.358)		
Regulation	2.860***	2.974***		
	(0.263)	(0.249)		
Control			0.785***	0.638***
			(0.058)	(0.110)
Age		0.030***	0.030***	0.039***
		(0.010)	(0.010)	(0.012)
Female		0.304	0.308	0.396*
		(0.186)	(0.188)	(0.228)
Altruism		0.005***	0.005***	0.006***
		(0.001)	(0.001)	(0.001)
High School		-1.695**	-1.795**	-1.744**
		(0.777)	(0.863)	(0.885)
Vocational training		-2.094**	-2.197**	-2.382**
		(0.854)	(0.930)	(0.976)
Attended college		-2.218***	-2.347***	-2.399***
		(0.746)	(0.832)	(0.845)
Bachelor degree		-2.111***	-2.188***	-2.274***
		(0.739)	(0.823)	(0.836)
Graduate degree		-1.838**	-1.932**	-1.982**
		(0.756)	(0.843)	(0.862)
Experience with AMT		-0.056	-0.043	0.028
		(0.084)	(0.084)	(0.099)
Constant	2.669***	3.149***	1.997**	1.793*
	(0.235)	(0.829)	(0.906)	(0.967)
Number of Obs	738	736	736	581
R-Squared	0.160	0.243	0.230	0.161

Table 3: Determinants of stage 1 DG contributions. Significance levels: *: < 0.1, **: < 0.05; ***: < 0.01.

4.2 Spillover effects

Moving to the analysis of spillovers from stage 1 to stage 2, the central and right panels of Figure 1 provide visual evidence that, by simple average comparison with baseline stage 2 behavior, the incentives and regulation treatment effects persist, even though attenuated, in the second DG. As Table 2 reports, average DG contributions are increasing from 27.7 % in the baseline to 32.7% and 35.5 % in

the incentives and regulation treatments, mirroring qualitatively the effects of stage 1. On the other hand, no clear pattern emerges from the PD, where cooperation is lowest, at 52.7 %, in the default treatment and highest, at 60.7 %, in the incentive one. When looking at the distribution of contributions, Table 2 shows that, as in stage 1, all treatments shift the main modal contribution from 0 to 50% of the endowment in the second stage DG. Moreover, the incentive and regulation treatments increase the share of subjects who give more than half of the endowment, by correspondingly reducing the frequency of contributions between 0 and 50%. The economic and regulation treatments also increase the share of high contributions in the PD, but they decrease the middle ones too in this game, with a net balancing effect. [5]

The distributional data also offer an overview of the differences in behavior generated by the two different strategic environments of the DG and PD. As already mentioned above, donations in the baseline DG appear supported on two modes (donating 0 or donating half), while contributions in the baseline PD appear supported on three modes (contributing 0, contributing half, or contributing everything, see Figure 7). Moreover, average donation in the baseline PD is about twice as large as that in the baseline DG. Consistent with this, the main modal contribution in the baseline DGs is 0, while in the PD is 10 cents. These differences are consistent with previous literature [33, 21] and support our claim that our design allows us to explore spillover effects within and across contexts.

Regression analysis helps us explore treatment effects on stage 2 behavior more precisely. Table 4 shows results from random-effects OLS regressions of game behavior on treatment, an indicator variable for stage 2 (*Post*), and their interaction. Columns 1 and 2 report results for the DG-DG treatments, while Columns 3 and 4 for the DG-PD ones. The coefficients on the treatment dummies capture their effect on donations in stage 1, confirming the statistically significant impact of the incentive and regulation treatments. The effect of the norms treatment, which appeared to be marginally statistically significant when considering the pooled sample, does not retain statistical significance when we consider the sub-samples of subjects in the DG and PD treatments separately.

Focusing on contribution dynamics, DG giving in the control treatment doesn't differ between stages, as the coefficient on the post dummy shows for the DG. The positive and significant coefficient of the *Post* dummy in the PD treatments is simply due to the fact that the PD and the DG are different games with different contribution profiles, as discussed above. The coefficients on the interaction terms capture the difference in donations between stages for each treatment: while there is no clear trend in donations under the default and norms treatments, the negative and significant coefficients of the incentive and regulation interaction terms indicate that the levels of donation achieved in the first stage are not sustained in the second. The overall spillover effect of treatments is captured by the marginal effects, that is by the sum of the main effects of treatment, *Post*, and their interaction, which confirm the small and not statistically significant effect of the behavioral treatments, and the persistence of the traditional ones. Marginal effects of treatments (Figure 2, left panel) indicate that 78% and 82% of

---

[5]Appendix A shows the distribution of stage 2 DG and PD contributions, by treatment, in Figure 6 and Figure 7.

the increase in contribution generated by the incentive and regulation treatments, respectively, persist in the second stage DG, and that these effects are statistically significant at the 10% and 5% level, respectively. Marginal effects are 51% and 28%, and not statistically significant, when we look at the impact of the two treatments on PD cooperation (Figure 2, right panel), thus indicating weak spillovers across different games. These results are robust to the inclusion of individual controls in the regressions, as shown in Columns 2 and 4 of Table 4.

Dependent variable	**All contributions**			
	DG-DG		**DG-PD**	
	(1)	**(2)**	**(3)**	**(4)**
Default	0.029	0.094	0.362	0.423
	(0.416)	(0.394)	(0.673)	(0.675)
Norm	0.428	0.464	0.778	1.048
	(0.420)	(0.399)	(0.685)	(0.677)
Incentive	1.977***	1.985***	1.926***	2.121***
	(0.433)	(0.411)	(0.695)	(0.703)
Regulation	2.811***	2.971***	2.954***	3.117***
	(0.419)	(0.397)	(0.682)	(0.673)
Post	-0.112	-0.113	3.240***	3.408***
	(0.204)	(0.204)	(0.583)	(0.586)
Default*Post	-0.013	-0.012	-0.649	-0.817
	(0.276)	(0.277)	(0.773)	(0.773)
Norm*Post	0.069	0.069	-0.715	-0.884
	(0.279)	(0.280)	(0.786)	(0.787)
Incentive*Post	-1.196***	-1.196***	-1.415*	-1.584**
	(0.288)	(0.288)	(0.799)	(0.799)
Regulation*Post	-1.984***	-1.984***	-2.675***	-2.844***
	(0.278)	(0.279)	(0.784)	(0.784)
Age		0.050***		0.000
		(0.012)		(0.016)
Female		0.096		0.672*
		(0.238)		(0.348)
Altruism		0.006***		0.005***
		(0.001)		(0.001)
High School		-3.516		-2.751
		(2.465)		(2.078)
Vocational training		-3.152		-3.762*
		(2.514)		(2.180)
Attended college		-3.731		-3.632*
		(2.450)		(2.043)
Bachelor degree		-3.718		-2.944
		(2.447)		(2.051)
Graduate degree		-3.382		-2.247
		(2.463)		(2.062)
Experience with AMT		-0.017		0.118
		(0.108)		(0.162)
Constant	2.887***	4.117*	2.320***	4.155*
	(0.307)	(2.438)	(0.508)	(2.245)
Number of Obs	884	882	592	590
R-Squared	0.103	0.218	0.121	0.183

Table 4: Determinants of contributions. Random effects OLS. Significance levels: *: < 0.1, **: < 0.05; ***: < 0.01.
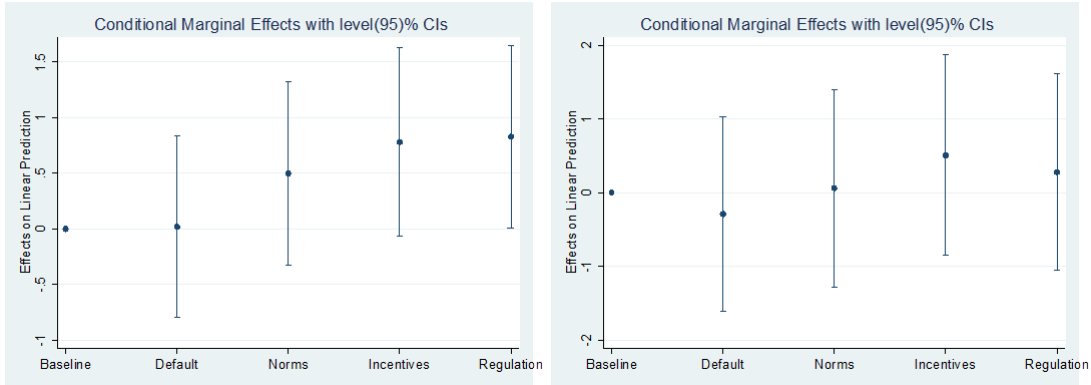
Figure 2: Marginal effects of treatments on stage 2 DG (left) and PD (right). Bars show 95% confidence intervals

The positive impact of the incentive and regulation treatments appears to be driven by their ability to influence the first stage game contribution. In order to isolate the effect of treatment working through the increase in stage 1 contributions, we run 2SLS regressions, instrumenting DG giving in stage 1 with treatment: the results (see Table 5) confirm a positive and significant role of increased donation in stage 1 on stage 2 pro-sociality, but only for the DG. It thus appears that, when the strategic context is the same, subjects anchor their second stage contribution to the one made in the first stage. These results are remarkable, especially in light of the fact that, consistent with the findings on direct effects, the spillover effect of the incentive treatment on net donations is not statistically significant. The anchoring effect of gross contributions thus appears to work, even if the incentive treatment does not affect the net donation level with respect to the baseline treatment, and even if the regulation treatment exogenously truncates the range of donation. Such strength of anchoring is consistent with existing evidence on how even arbitrary anchors can affect preferences ([6]).

Dependent variable	Stage 2 contribution			
	DG		PD	
	(1)	(2)	(3)	(4)
DG1 contribution	0.282***	0.306***	0.185	0.202
	(0.105)	(0.094)	(0.233)	(0.220)
Age		0.043***		-0.014
		(0.012)		(0.023)
Female		0.004		0.653
		(0.207)		(0.505)
Altruism		0.005***		0.005**
		(0.001)		(0.002)
High School		-2.547***		-4.248***
		(0.442)		(0.854)
Vocational training		-2.149***		-4.778***
		(0.559)		(1.220)
Attended college		-2.825***		-4.713***
		(0.371)		(0.746)
Bachelor degree		-2.665***		-3.953***
		(0.350)		(0.657)
Graduate degree		-2.357***		-2.882***
		(0.415)		(0.727)
Experience with AMT		-0.050		0.378
		(0.093)		(0.241)
Constant	2.089***	2.728***	5.008***	7.744***
	(0.395)	(0.468)	(0.848)	(1.561)
Number of Obs	442	441	296	295
R-Squared	0.347	0.458	0.074	0.135

Table 5: Determinants of second stage DG and PD contributions. 2SLS. Significance levels: *: < 0.1, **: < 0.05; ***: < 0.01.

4.2.1 Robustness with a low focal point

The set of additional treatments with a low donation focal point -that is, the treatments where defaults, norms, incentives and regulation were used to induce giving 10% as opposed to 50% of one's endowment- allows us to disentangle the spillover effect of treatments net of their impacts on stage 1 contribution. Consistent with our goals, contribution levels in these treatments are not significantly different from that in the baseline one (Figure 3). In the same figure, we can also observe no spillover of any treatments on either DG or PD.
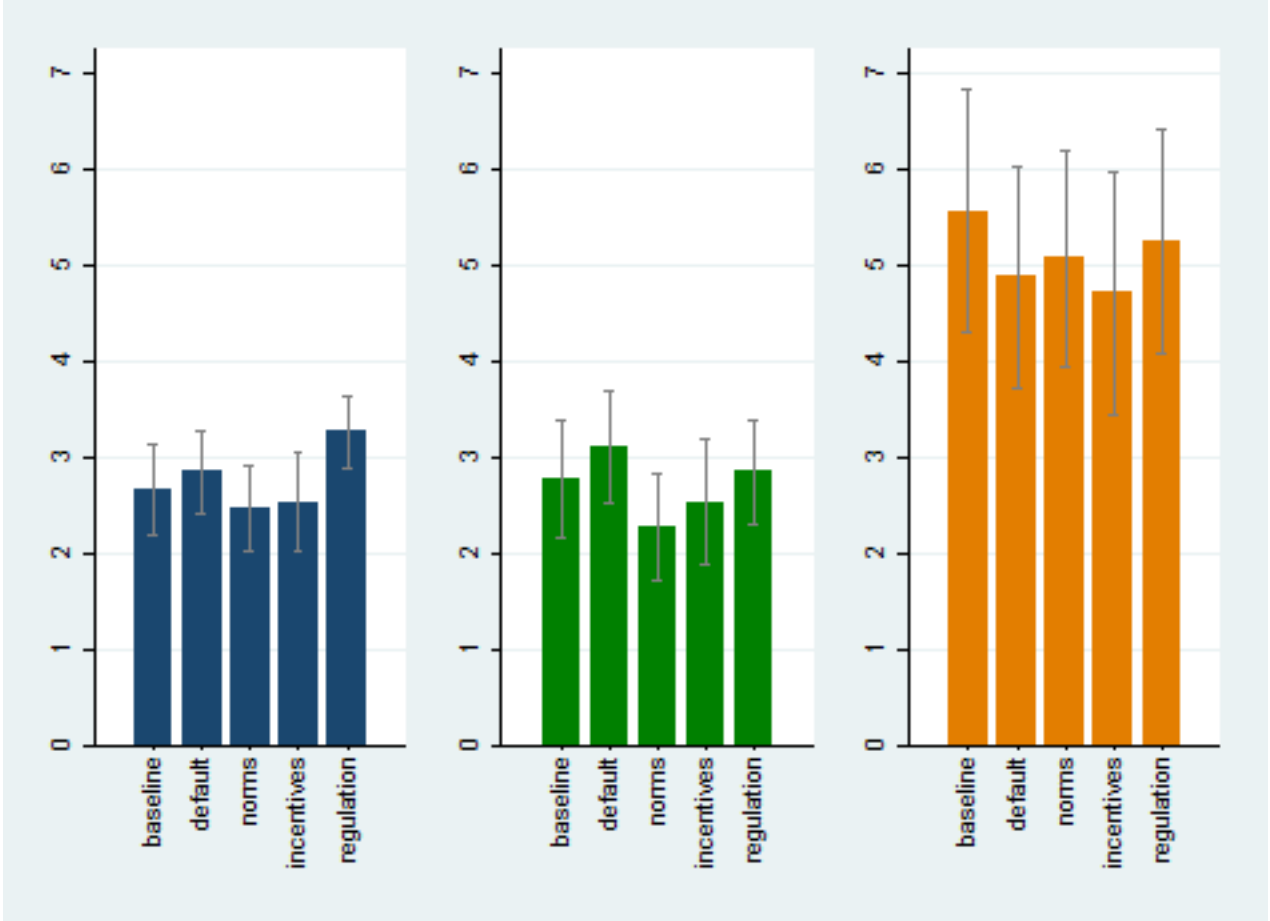
Figure 3: Average contributions (normalized to 10) in stage 1 DG (left panel), stage 2 DG (central panel) and stage 2 PD (right panel), by first stage treatments, for the 10% focal donation point. Error bars represent 95% confidence intervals.

Random effect OLS regressions, shown in Table 6, confirm the non statistically significant effect of treatments and spillovers in these conditions: in spite of the significance of individual coefficients, none of the marginal effects of treatments are statistically significant (Figure 8 in Appendix A). All marginal effects are negative, especially in the case of the PD.

The low focal point conditions also help us further assess the effect of the behavioral and traditional treatments, by comparing average contributions by treatment between high and low focal point conditions. While we find no difference in stage 1 giving between the default treatment with high and low focal point, in all the other treatments the high focal point induces significantly higher contributions than the low one (p=0.0111 for norms, p<0.0000 for incentives and regulation). In terms of spillover effects, we find statistically significant differences in stage 2 DG contributions between

high and low focal point conditions in the norms (p=0.0130), incentive (p=0.0364) and regulation treatments (p=0.0752), but no effect in the default one. No difference between high and low focal point conditions is significant when we consider PD contributions. These results are consistent with our previous findings and provide further support to our assessment of the relative effectiveness of the different policy instruments.

| Dependent variable | Stage 2 game decision: contribution | | | | | |
| | DG-DG | | | DG-PD | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Default	0.200	0.263		0.196	0.053	
	(0.419)	(0.392)		(0.697)	(0.689)	
Norm	-0.614	-0.394		0.413	0.646	
	(0.427)	(0.400)		(0.703)	(0.704)	
Incentive	-0.717*	-0.589		0.732	0.858	
	(0.429)	(0.403)		(0.708)	(0.698)	
Regulation	0.279	0.416		1.131	1.152	
	(0.414)	(0.389)		(0.730)	(0.720)	
Post	-0.113	-0.113	0.036	3.240***	3.408***	3.443***
	(0.146)	(0.146)	(0.151)	(0.542)	(0.544)	(0.548)
Default*Post	0.134	0.134		-0.885	-1.053	
	(0.201)	(0.201)		(0.728)	(0.727)	
Norm*Post	0.113	0.113		-0.907	-1.075	
	(0.205)	(0.205)		(0.733)	(0.733)	
Incentive*Post	0.478**	0.478**		-1.585**	-1.753**	
	(0.206)	(0.206)		(0.739)	(0.738)	
Regulation*Post	-0.210	-0.210		-1.436*	-1.604**	
	(0.198)	(0.198)		(0.762)	(0.761)	
Control			0.009			0.314**
			(0.087)			(0.157)
Control*Post			-0.018			-0.385**
			(0.045)			(0.166)
Age		0.024**	0.025**		0.012	0.012
		(0.012)	(0.012)		(0.018)	(0.018)
Female		-0.020	-0.017		0.794**	0.774**
		(0.246)	(0.247)		(0.370)	(0.367)
Altruism		0.007***	0.007***		0.007***	0.007***
		(0.001)	(0.001)		(0.001)	(0.001)
High School		-2.459	-2.454		-0.505	-0.168
		(2.503)	(2.506)		(2.261)	(2.231)
Vocational training		-2.370	-2.285		-0.660	-0.276
		(2.562)	(2.565)		(2.414)	(2.380)
Attended college		-2.985	-2.982		-0.494	-0.171
		(2.484)	(2.485)		(2.205)	(2.178)
Bachelor degree		-3.607	-3.637		-0.658	-0.348
		(2.482)	(2.485)		(2.192)	(2.164)
Graduate degree		-3.610	-3.587		-0.824	-0.482
		(2.501)	(2.504)		(2.259)	(2.225)
Experience with AMT		-0.057	-0.055		-0.199	-0.201
		(0.106)	(0.107)		(0.172)	(0.171)
Constant	2.887***	4.652*	4.565*	2.320***	1.949	1.254
	(0.306)	(2.470)	(2.482)	(0.519)	(2.335)	(2.275)
Number of Obs	866	866	866	562	560	560
R-Squared	0.017	0.168	0.154	0.096	0.165	0.161

Table 6: Determinants of DG2 and PD giving, low focal point. Random effects OLS. Significance levels: *: < 0.1, **: < 0.05; ***: < 0.01.

An alternative explanation of our results relies on income effects. Treatment effects in stage 1 determine differences in subjects' income as they make their stage 2 decisions. Spillover effects could thus be caused indirectly by treatments through income effects. We rule out this alternative hypothesis thanks to the data from our pilot study, where treatments differed in the size of the endowment. We find no differences in contributions across treatments. Specifically, in the treatment with a participation fee of 50 cents the average donation was 4.84 cents, while in the treatment with a participation fee of 70 cents the average donation was 4.82 cents (2-tailed t test, p = 0.9785).

4.3   Subjects' guess of treatment effects

An interesting question is whether individuals correctly predict the observed effects. Recipients in the DG games in both stage 1 and 2 were asked to guess the contribution of donors. Figure 4 compares people's beliefs with actual donations for each of the treatments. Pairwise comparison (between baseline beliefs and treated beliefs) shows that people highly overestimate both the direct effectiveness of behaviorally informed interventions (all p's $< .0001$), and the presence of spillover effects (all p's $< .001$). On the other hand, observers correctly guessed the presence of direct effects of traditional interventions and their capacity to spill over. Specifically, treated beliefs are significantly higher than untreated beliefs both in Stage 1 (all p's $< .0001$) and in Stage 2 (incentive: $p = 0.0030$; regulation: $p = 0.0337$). The effect of behavioral policy instruments on beliefs is reassuring, as it demonstrates that the design of these treatments was not so mild as to make them undetectable.

The data on beliefs, while giving us a sense of the sign of the effects, does not allow us to say whether subjects correctly guessed also the magnitude of these effects. The main reason behind the limited informativeness of beliefs is that our experiment is based on a between-subjects design and we do not have a measure of untreated beliefs and donations for treated subjects. In other words, quantifying the biases in beliefs, relative to the observed treatment effects on donations, would require a within-subjects design, where data on both donations and beliefs were collected for each individual. Moreover, the data on beliefs contains some noise: namely, 11% of observers in the regulation treatment believed that donors would give less than half of their endowment. Since giving less than half was not allowed, it is likely that these observers did not understand the decision problem at hand. Indeed, eliminating these subjects from the analysis, we find that average beliefs in the regulation treatment increase to 5.22 cents and, consequently, the difference between average beliefs in the regulation treatment and average untreated beliefs in stage 1 baseline decreases to 5.22 - 2.31 = 2.91 cents, which is essentially the same as the difference between average donations in the regulation treatment and average untreated donations in stage 1 baseline (5.53 - 2.67 = 2.86). Thus, controlling for potential noise, it seems that subjects correctly predict the magnitude of the spillover effect of the regulation treatment. However, as already mentioned, comparing the differences between donations and beliefs would require a within-subjects design. Future research may help clarify whether there is a true bias in the estimation of the

magnitude of direct and spillover effects of traditional policy interventions.
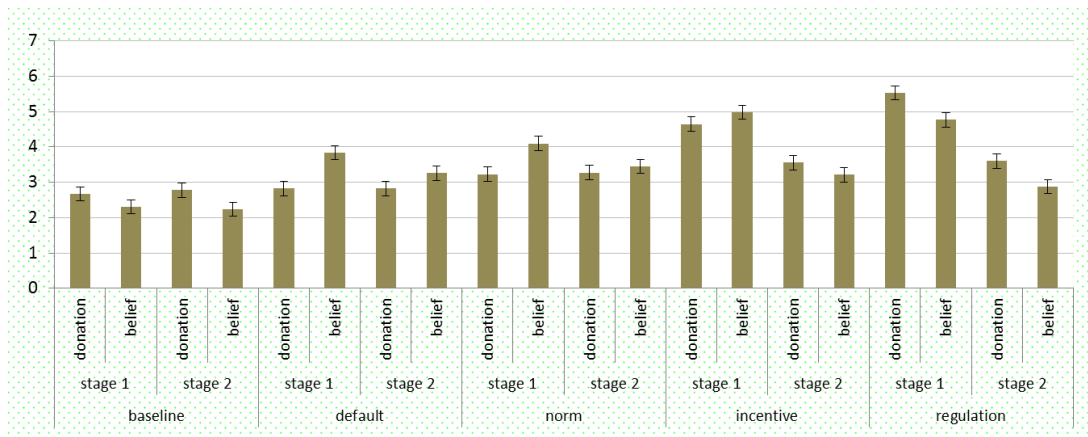
Figure 4: Difference between DG donations and guesses, by treatment and stage

Random effects OLS regressions in Table 7 confirm that individuals predict a positive impact of all interventions -including the behavioral ones- on both first stage contributions and spillovers to second stage DG (the guessing was only done by DG receivers). When contrasted with the actual results of Table 4, it appears that individuals overestimate the positive direct and spillover effect of the behavioral treatments. As for the traditional treatments, regression results confirm that subjects underestimate both direct and spillover effects.

Dependent variable	Belief (guessing) of stage 1 and stage 2 DG contributions		
	(1)	(2)	(3)
Default	0.826***	0.828***	
	(0.279)	(0.277)	
Norm	0.977***	0.953***	
	(0.282)	(0.281)	
Incentive	1.812***	1.800***	
	(0.281)	(0.280)	
Regulation	1.460***	1.461***	
	(0.279)	(0.277)	
Post	-0.045	-0.045	0.293**
	(0.131)	(0.131)	(0.135)
Default*Post	-0.309*	-0.309*	
	(0.181)	(0.181)	
Norm*Post	-0.351*	-0.353*	
	(0.183)	(0.183)	
Incentive*Post	-1.115***	-1.115***	
	(0.182)	(0.182)	
Regulation*Post	-1.130***	-1.130***	
	(0.181)	(0.181)	
Control			0.384***
			(0.062)
Control*Post			-0.306***
			(0.040)
Age		0.002	0.002
		(0.008)	(0.008)
Female		0.461***	0.458***
		(0.170)	(0.171)
Altruism		0.002**	0.002**
		(0.001)	(0.001)
High School		-1.971*	-1.878*
		(1.102)	(1.103)
Vocational training		-1.639	-1.605
		(1.144)	(1.145)
Attended college		-2.052*	-2.012*
		(1.084)	(1.085)
Bachelor degree		-2.212**	-2.170**
		(1.082)	(1.083)
Graduate degree		-2.668**	-2.628**
		(1.103)	(1.103)
Experience with AMT		-0.188**	-0.189**
		(0.075)	(0.075)
Constant	2.125***	4.329***	4.164***
	(0.202)	(1.116)	(1.121)
Number of Obs	2378	2376	2376
R-Squared	0.034	0.059	0.052

Table 7: Beliefs on DG giving. Random effects OLS. Significance levels: *: < 0.1, **: < 0.05; ***: < 0.01.

5 Discussion

We provide one of the first experimental assessments of whether and how spillovers of pro-social behavior depend on the policy instruments used to motivate behavioral changes in the first place, examining within the same experimental design four among the most common behavioral (defaults and social norms) and traditional (economic incentives and regulation) policy levers. Results show that traditional policy instruments appear to be more effective than behaviorally-informed ones to foster giving directly, and to generate positive spillovers in terms of higher giving in a subsequent game, but only in the same context. In our study, traditional policy instruments have the greatest influence on behavior within the same context, followed by behavioral policy instruments based on social comparison. The mechanism driving the positive spillovers appears to be that of anchoring on the initial donations. The results are confirmed by the non significant spillover effects of the set of treatments focused on a low donation focal point, which did not induce a change in first stage contributions. Summing up, subjects appear to anchor their second stage contribution to the one they made in the first stage, provided the context is the same.

Our results have direct policy implications: we are able to show that, although traditional policy interventions based on economic incentives and regulation might crowd-out some intrinsic motivation, their high effectiveness in raising initial altruistic donation has overall positive consequences on subsequent pro-social behavior, thanks to an anchoring effect on initial contributions. Almost all of the policy induced contribution in the first stage carries forward in the second one. For economic incentives, taking into account the spillover effects of the treatment greatly influences the evaluation of its cost-effectiveness: while incentives do not significantly increase net donations if we only consider stage 1, they have a positive and significant effect on overall donations (including stage 2), since the anchoring appears to be on gross contributions, without factoring in the government rebate.

Although our main results appear to be robust to several specifications, and have been generated by a large sample of more than 3000 subjects, additional research is required on comparing behavioral and traditional policy interventions, especially since behavioral interventions are receiving increased attention from policy makers. For example, other behavioral interventions (e.g. priming, identity manipulation, framing etc.) or other ways to implement defaults or norms from the way we have done exist, and require further scrutiny. Some of these alternative designs could be able to induce large initial changes, comparable to those of strong traditional policy interventions, though unlikely to be as determinant as our regulation policy lever. Whether in such a case the anchoring would also hold for the behavioral policy instruments remains an open question. Furthermore, an assessment of behavioral spillovers in the field, especially in multidimensional policy areas such as pro-environmental behavior, would allow for a more robust evaluation of contexts of added complexity and less clear preferences than those represented by economic games, and for a true comparison of the relative cost-effectiveness of the different policies, once taking into account their implementation, monitoring and enforcement

costs.

Finally, our research provides suggestive evidence on the welfare implications of different policy instruments. Behavioral policies are often preferred to traditional ones on the grounds of their minimal invasiveness on individual decision making autonomy and thus impact on welfare. Our results do not confirm this common view. Within the stylized setting of our experiment, we capture welfare effects through a happyness question included in the endline survey. We don't observe any statistically significant differences in subjects' self-reported happyness levels across treatments.[6] Further research in the field, possibly making use of revealed preferences approached, should explore whether these welfare effects are confirmed.

Acknowledgments

References

[1] Bridging the gap. the role of monitoring and evaluation in evidence-based policy making. 00023. 1

[2] Hunt Allcott. Social norms and energy conservation. 95(9):1082–1095. 00344. 1, 2.2

[3] Ofra Amir, David G Rand, and others. Economic games on the internet: The effect of \$1 stakes. 00089. 3

[4] James Andreoni and John Miller. Giving according to GARP: An experimental test of the consistency of preferences for altruism. pages 737–753. 01332. 2.2

[5] James Andreoni and A. Abigail Payne. Is crowding out due entirely to fundraising? evidence from a panel of charities. 95(5):334–343. 00144. 2.2

[6] Dan Ariely, George Loewenstein, and Drazen Prelec. "coherent arbitrariness": Stable demand curves without stable preferences. 118(1):73–106. 00000. 4.2

[7] Nava Ashraf, Oriana Bandiera, and B. Kelsey Jack. No margin, no mission? a field experiment on incentives for public service delivery. 120:1–17. 00026. 2.2

---

[6]Pairwise t-tests between self-reported happyness in the baseline and each of the high focal point treatments yield p-values $>= 0.5967$.

[8] Nava Ashraf, Oriana Bandiera, Scott S. Lee, and others. Do-gooders and go-getters: career incentives, selection, and performance in public service delivery. 00009. 2.2

[9] Oriana Bandiera, Iwan Barankay, and Imran Rasul. Social preferences and the response to incentives: Evidence from personnel data. pages 917–962. 00477. 2.2

[10] Nicholas Bardsley and Rupert Sausgruber. Conformity and reciprocity in public good provision. 26(5):664–681. 00115. 2.2

[11] Anke Becker, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse. The relationship between economic preferences and psychological personality measures. 4(1):453–478. 00084. 3

[12] R. Benabou and J. Tirole. Identity, morals, and taboos: Beliefs as assets. 126(2):805–855. 00179. 2.1

[13] Irene Blanken, Niels van de Ven, and Marcel Zeelenberg. A meta-analytic review of moral licensing. 41(4):540–558. 00006. 2.1

[14] Iris Bohnet and Stephan Meier. Deciding to distrust. 00012. 2.2

[15] Iris Bohnet and Richard Zeckhauser. Social comparisons in ultimatum bargaining. 106(3):495–510. 00117. 2.2

[16] Samuel Bowles and Sandra Polanía-Reyes. Economic incentives and social preferences: Substitutes or complements? 50(2):368–425. 00116. 2.1, 2.2, 2

[17] Jordi Brandts and David J. Cooper. A change would do you good.... an experimental study on how to overcome coordination failure in organizations. pages 669–693. 00174. 1

[18] Gharad Bryan, Dean Karlan, and Scott Nelson. Commitment devices. 2(1):671–698. 00110. 1

[19] James H. Bryan and Mary A. Test. Models and helping: naturalistic studies in aiding behavior. 6(4):400. 00411. 2.2

[20] Dominique Cappelletti, Luigi Mittone, and Matteo Ploner. Are default contributions sticky? an experimental analysis of defaults in public goods provision. 108:331–342. 2.2

[21] Valerio Capraro, Jillian J. Jordan, and David G. Rand. Heuristics guide the implementation of social preferences in one-shot prisoner's dilemma experiments. 4. 00014. 2.1, 4.2

[22] Valerio Capraro, Conor Smyth, Kalliopi Mylona, and Graham A. Niblo. Benevolent characteristics promote cooperative behaviour among humans. 9(8):e102881. 00007. 2.1

[23] Juan-Camilo Cardenas. Norms from outside and from inside: an experimental analysis on the governance of local ecosystems. 6(3):229–241. 00083. 2.2

[24] Jeffrey P. Carpenter and David Dolifka. Exploitation aversion: When financial incentives fail to motivate agents. 00002. 2.2

[25] Timothy N. Cason and Vai-Lam Mui. Social influence in the sequential dictator game. 42(2):248–265. 00266. 2.2

[26] Timothy N. Cason, Anya C. Savikhin, and Roman M. Sheremeta. Behavioral spillovers in coordination games. 56(2):233–245. 00042. 2.1

[27] James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. In *Tax Policy and the Economy, Volume 16*, pages 67–114. MIT Press. 00513. 2.2

[28] Rachel TA Croson. Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. 45(2):199–216. 00295. 2.2

[29] Magali A. Delmas and Neil Lessem. Saving power to conserve your reputation? the effectiveness of private versus public information. 67(3):353–370. 00004. 2.2

[30] Giovanna Devetag. Precedent transfer in coordination games: An experiment. 89(2):227–232. 00038. 1

[31] Nikhil Dhingra, Zach Gorn, Andrew Kener, and Jason Dana. The default pull: An experimental demonstration of subtle default effects on preferences. 7(1):69–76. 00008. 2.2

[32] Paul Dolan and Matteo M. Galizzi. Like ripples on a pond: Behavioral spillovers and their implications for research and policy. 47:1–16. 00000. 1, 2.1

[33] Christoph Engel. Dictator games: a meta study. 14(4):583–610. 00323. 4.1, 4.2

[34] Armin Falk and Michael Kosfeld. The hidden costs of control. 96(5):1611–1630. 00584. 2.2

[35] Ernst Fehr and Bettina Rockenbach. Detrimental effects of sanctions on human altruism. 422(6928):137–140. 00645. 2.2

[36] Paul J. Ferraro and Michael K. Price. Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. 95(1):64–73. 00052. 1, 2.2

[37] Erica Field. Entitled to work: Urban property rights and labor supply in peru. 122(4):1561–1602. 00442. 2.2

[38] Urs Fischbacher, Simon Gächter, and Ernst Fehr. Are people conditionally cooperative? evidence from a public goods experiment. 71(3):397–404. 01653. 2.2

[39] Bruno S. Frey and Reto Jegen. Motivational interactions: Effects on behaviour. pages 131–153. 00020. 2.2

[40] Bruno S Frey and Stephan Meier. Social comparisons and pro-social behavior: Testing "conditional cooperation" in a field experiment. 94(5):1717–1722. 2.2

[41] Bruno S. Frey and Felix Oberholzer-Gee. The cost of price incentives: An empirical analysis of motivation crowding-out. pages 746–755. 01229. 2.2

[42] Ayelet Gneezy, Alex Imas, Amber Brown, Leif D. Nelson, and Michael I. Norton. Paying to be nice: Consistency and costly prosocial behavior. 58(1):179–187. 00030. 2.1

[43] Uri Gneezy, Alex Imas, and Kristóf Madarász. Conscience accounting: Emotion dynamics and social behavior. 60(11):2645–2658. 00008. 2.1

[44] Uri Gneezy and Aldo Rustichini. Fine is a price, a. 29:1. 01411. 2.2

[45] Uri Gneezy and Aldo Rustichini. Pay enough or don't pay at all. 115(3):791–810. 01168. 2.2

[46] Lorenz Goette and Alois Stutzer. Blood donations and incentives: Evidence from a field experiment. 00063. 2.2

[47] N. J. Goldstein, R. B. Cialdini, and V. Griskevicius. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. 35(3):472–482. 00600. 2.2

[48] Joseph K Goodman, Cynthia E Cryder, and Amar Cheema. Data collection in a flat world: The strengths and weaknesses of mechanical turk samples. 26(3):213–224. 00276. 3

[49] Zachary Grossman. Strategic ignorance and the robustness of social preferences. 60(11):2659–2665. 00030. 2.2

[50] Christopher J Holden, Trevor Dennie, and Adam D Hicks. Assessing the reliability of the m5-120 on amazon?s mechanical turk. 29(4):1749–1754. 00000. 3

[51] John J. Horton, David G. Rand, and Richard J. Zeckhauser. The online laboratory: Conducting experiments in a real labor market. 14(3):399–425. 00351. 3

[52] R. Bruce Hutton and Dennis L. McNeill. The value of incentives in stimulating energy conservation. pages 291–298. 00040. 2.2

[53] Koichiro Ito, Takanori Ida, and Makoto Tanaka. The persistence of moral suasion and economic incentives: Field experimental evidence from energy demand. 00003. 1, 2.2

[54] Eric J. Johnson and Daniel G. Goldstein. Do defaults save lives? 302:1338–1339. 00749. 2.2

[55] Eric J. Johnson, John Hershey, Jacqueline Meszaros, and Howard Kunreuther. *Framing, probability distortions, and insurance decisions.* Springer. 00718. 2.2

[56] Jennifer Jordan, Elizabeth Mullen, and J. Keith Murnighan. Striving for the moral self: The effects of recalling past moral actions on future moral behavior. 37(5):701–713. 00087. 2.1

[57] Dean Karlan and John A. List. Does price matter in charitable giving? evidence from a large-scale natural field experiment. 97(5):1774–1793. 00000. 2.2

[58] Claudia Keser and Frans Van Winden. Conditional cooperation and voluntary contributions to public goods. 102(1):23–39. 00570. 2.2

[59] Richard A Klein, Kate A Ratliff, Michelangelo Vianello, Reginald B Adams Jr, Štěpán Bahník, Michael J Bernstein, Konrad Bocian, Mark J Brandt, Beach Brooks, Claudia Chloe Brumbaugh, and others. Investigating variation in replicability. 00071. 3

[60] Marc Knez and Colin Camerer. Increasing cooperation in prisoner's dilemmas by establishing a precedent of efficiency in coordination games. 82(2):194–216. 00120. 1

[61] Nicola Lacetera and Mario Macis. Do all material incentives for pro-social activities backfire? the response to cash and non-cash incentives for blood donations. 31(4):738–748. 00078. 2.2

[62] Pietro Lanzini and John Thøgersen. Behavioural spillover in the environmental domain: An intervention study. 40:381–390. 00001. 2.2

[63] John A. List. The market for charitable giving. pages 157–180. 00093. 2.2

[64] Paul D. Luyben. Prompting thermostat setting behavior public response to a presidential appeal for conservation. 14(1):113–128. 00049. 2.2

[65] Brigitte C. Madrian and Dennis F. Shea. The power of suggestion: Inertia in 401 (k) participation and savings behavior. 116(4):1149–87. 01646. 2.2

[66] Richard Martin and John Randal. How is donation behaviour affected by the donations of others? 67(1):228–238. 00080. 2.2

[67] Winter Mason and Siddharth Suri. Conducting behavioral research on amazon's mechanical turk. 44(1):1–23. 00468. 3

[68] Nina Mazar and Chen-Bo Zhong. Do green products make us better people? 21(4):494–498. 00236. 2.1

[69] Carl Mellström and Magnus Johannesson. Crowding out in blood donation: was titmuss right? 6(4):845–863. 00314. 2.2

[70] Benoît Monin and Dale T. Miller. Moral credentials and the expression of prejudice. 81(1):33–43. 00445. 2.1

[71] Jessica M. Nolan, P. Wesley Schultz, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius. Normative social influence is underdetected. 34(7):913–923. 2.2

[72] Gabriele Paolacci and Jesse Chandler. Inside the turk understanding mechanical turk as a participant pool. 23(3):184–188. 00014. 1, 3

[73] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. Running experiments on amazon mechanical turk. 5(5):411–419. 01336. 3

[74] Alexander Peysakhovich, Martin A. Nowak, and David G. Rand. Humans display a 'cooperative phenotype' that is domain general and temporally stable. 5. 00017. 2.1

[75] Alexander Peysakhovich and David G. Rand. Habits of virtue: Creating norms of cooperation and defection in the laboratory. 00000. 2.1

[76] Daniel Pichert and Konstantinos V. Katsikopoulos. Green defaults: Information presentation and pro-environmental behaviour. 28(1):63–73. 00141. 2.2

[77] David G. Rand. The promise of mechanical turk: How online labor markets can help theorists run behavioral experiments. 299:172–179. 00159. 3

[78] Peter H. Reingen. Inducing compliance with a donation request. 106(2):281–282. 00010. 2.2

[79] Jens Rommel, Vera Buttmann, Georg Liebig, Stephanie Schönwetter, and Valeria Svart-Gröger. Motivation crowding theory and pro-environmental behavior: Experimental evidence. 129:42–44. 00002. 2.2

[80] Sonya Sachdeva, Rumen Iliev, and Douglas L. Medin. Sinning saints and saintly sinners the paradox of moral self-regulation. 20(4):523–528. 00198. 2.1

[81] P. W. Schultz, J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius. The constructive, destructive, and reconstructive power of social norms. 18(5):429–434. 00749. 2.2

[82] Jen Shang and Rachel Croson. A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. 119(540):1422–1439. 00291. 2.2

[83] Danielle N Shapiro, Jesse Chandler, and Pam A Mueller. Using mechanical turk to study clinical populations. page 2167702612469015. 00122. 3

[84] Frans W. Siero, Arnold B. Bakker, Gerda B. Dekker, and Marcel TC Van Den Burg. Changing organizational energy consumption behaviour through comparative feedback. 16(3):235–246. 00202. 2.2

[85] Henk Staats, Paul Harland, and Henk AM Wilke. Effecting durable change a team approach to improve environmental behavior in the household. 36(3):341–367. 00261. 2.2

[86] Cass R. Sunstein and Lucia A. Reisch. Green by default. 66(3):398–402. 00003. 1

[87] Sophie Sutcliff and Julius Court. Evidence-based policymaking: What is it? how does it work? what relevance for developing countries? 00000. 1

[88] Richard H. Thaler and Cass R. Sunstein. Libertarian paternalism. 93(2):175–179. 00955. 2.2

[89] Verena Tiefenbeck, Thorsten Staake, Kurt Roth, and Olga Sachs. For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. 57:160–171. 00024. 2.2

[90] Heather Barnes Truelove, Amanda R. Carrico, Elke U. Weber, Kaitlin Toner Raimi, and Michael P. Vandenbergh. Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. 29:127–138. 00007. 1, 2.1, 1

[91] Gabriel Weimann. *The influentials: People who influence people.* SUNY Press. 00016. 2.2

[92] Richard A. Winett, Ingrid N. Leckliter, Donna E. Chinn, Brian Stahl, and Susie Q. Love. Effects of television modeling on residential energy conservation. 18(1):33. 00101. 2.2

Appendix

Appendix A: Additional Results
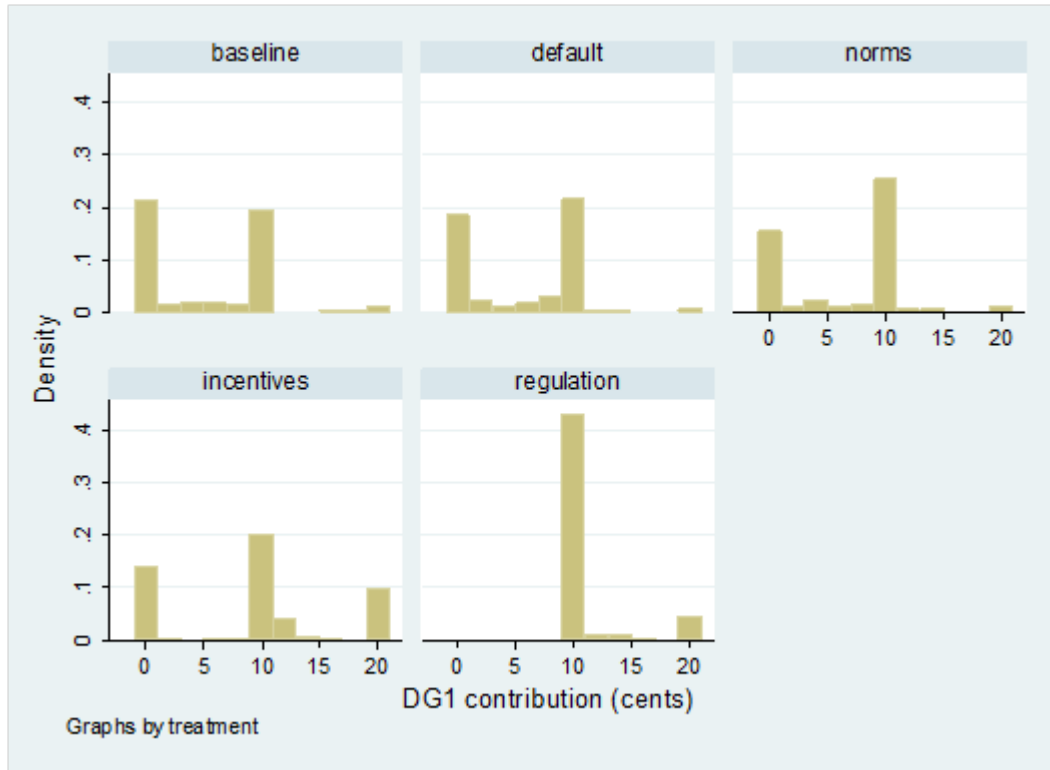
Figure 5: Distribution of stage 1 DG contribution, by treatment
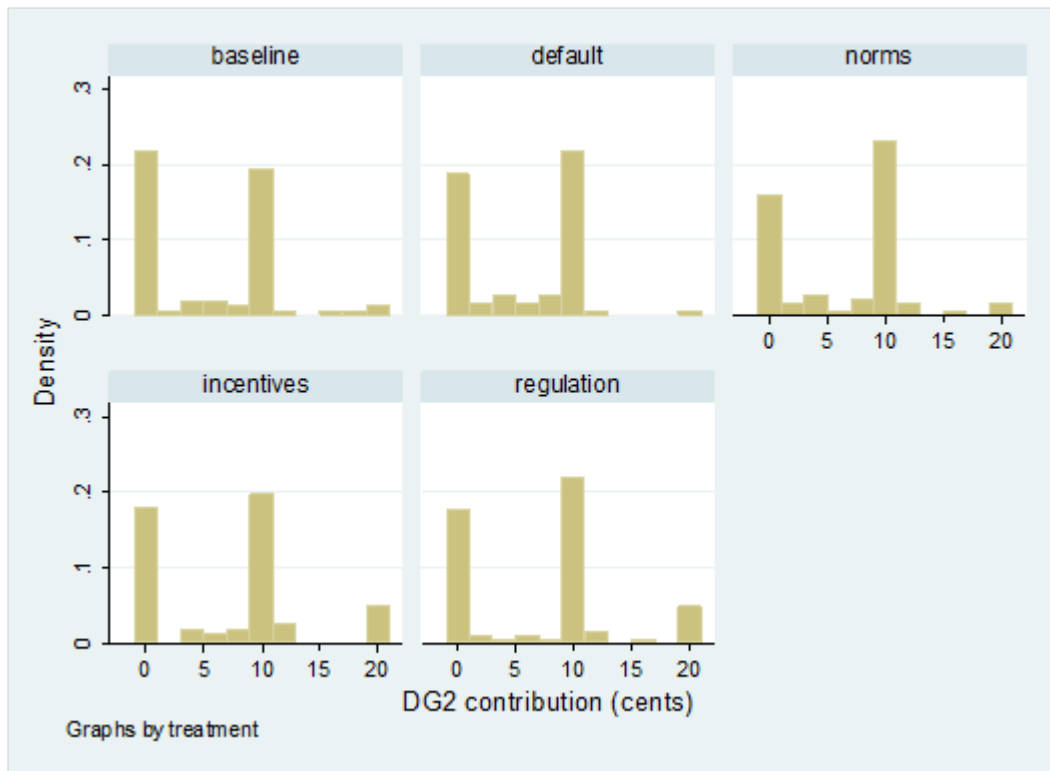
Figure 6: Distribution of stage 2 DG contribution, by stage 1 treatment
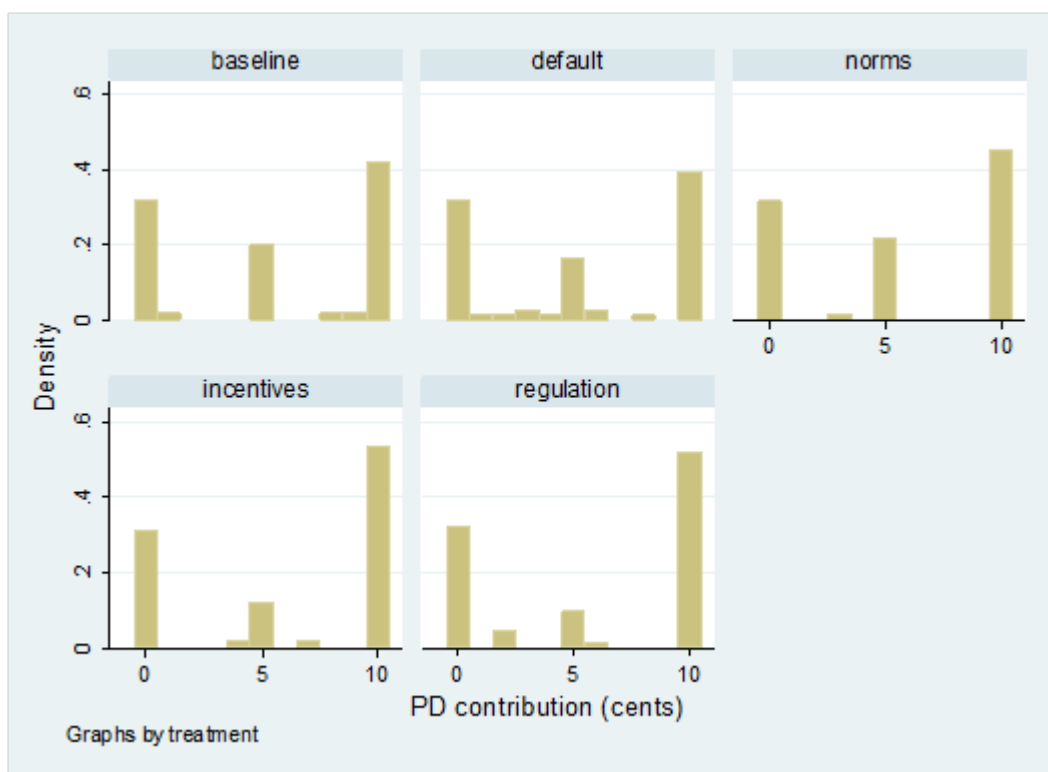
Figure 7: Distribution of stage 2 PD contribution, by stage 1 treatment

	DG-PD			DG-DG			Mean equality (p)	Distr. equality (p)
	N	Mean	Std.Dev	N	Mean	Std.Dev		
Age	527	31.930	10.426	794	32.340	10.395	**0.483**	.
Female (%)	527	0.408	0.492	795	0.401	0.490	**0.808**	**1.000**
Happyness (1 sad-9 happy)	527	6.059	1.463	795	6.169	1.494	**0.188**	.
Experience with AMT (1 never-5 often)	527	2.831	1.050	795	2.730	1.105	**0.095***	.
Altruism index	527	102.349	123.228	795	108.387	124.782	**0.387**	.
Married	527	0.298	0.458	795	0.317	0.466	**0.463**	**1.000**
High school	527	0.101	0.301	795	0.122	0.328	**0.229**	.
Vocational training	527	0.034	0.182	795	0.039	0.194	**0.649**	**1.000**
Attended college	527	0.343	0.475	795	0.322	0.468	**0.418**	.
Bachelor degree	527	0.391	0.488	795	0.389	0.488	**0.936**	**1.000**
Graduate degree	527	0.121	0.327	795	0.127	0.333	**0.763**	**1.000**
Income bracket	527	4.932	2.526	795	5.117	2.430	**0.182**	.
Political affiliation (1 right-7 left)	527	4.761	1.589	795	4.713	1.558	**0.589**	.
Religious faith (1 no-7 yes)	527	3.205	2.417	795	3.589	2.507	**0.006***	.

'**Mean equality** (p)' reports the p-value from a two-sample t test with equal variances.

'**Distr. equality** (p)' reports the p-value from a two-sample Kolmogorov-Smirnov test (calculated exactly, rather than approximately).

Confidence: '*': 90%; '**': 95%; '***': 99%.

Table 8: Balance between participants in DG-DG and in DG-PD treatments

Dependent variable	Stage 1 contribution			
	All	**All**	**All**	**Excl. Regulation**
	(1)	**(2)**	**(3)**	**(4)**
Default	0.152	0.205		
	(0.305)	(0.290)		
Norm	0.560*	0.623**		
	(0.310)	(0.295)		
Incentive	-0.336	-0.313		
	(0.278)	(0.263)		
Regulation	2.860***	2.965***		
	(0.263)	(0.249)		
Control			0.576***	-0.044
			(0.058)	(0.084)
Age		0.023***	0.021**	0.029***
		(0.008)	(0.009)	(0.010)
Female		0.222	0.270	0.307
		(0.162)	(0.174)	(0.195)
Altruism		0.005***	0.005***	0.005***
		(0.001)	(0.001)	(0.001)
High School		-1.641**	-1.569	-1.526*
		(0.757)	(1.262)	(0.902)
Vocational training		-1.806**	-1.218	-1.558
		(0.832)	(1.318)	(0.999)
Attended college		-2.045***	-1.725	-1.840**
		(0.736)	(1.245)	(0.882)
Bachelor degree		-2.011***	-1.530	-1.831**
		(0.729)	(1.240)	(0.874)
Graduate degree		-1.796**	-1.425	-1.656*
		(0.743)	(1.252)	(0.888)
Experience with AMT		-0.056	-0.037	0.007
		(0.072)	(0.075)	(0.085)
Constant	2.669***	3.353***	1.999	3.014***
	(0.235)	(0.793)	(1.269)	(0.941)
Number of Obs	738	736	736	581
R-Squared	0.214	0.292	0.171	0.105

Table 9: Determinants of stage 1 DG contributions. Significance levels: *: < 0.1, **: < 0.05; ***: < 0.01.

			50% Focal Point				10% Focal Point			
		Control	**Default**	**Norms**	**Incentives**	**Regulation**	**Default**	**Norms**	**Incentives**	**Regulation**
Stage 1	DG	72%	75%	79%	70%	74%	66%	76%	77%	73%
Stage 2	DG	-	-	-	-	-	-	-	-	-
	PD	65%	73%	65%	77%	76%	80%	70%	56%	73%

Table 10: Percentage of participants who passed the comprehension questions
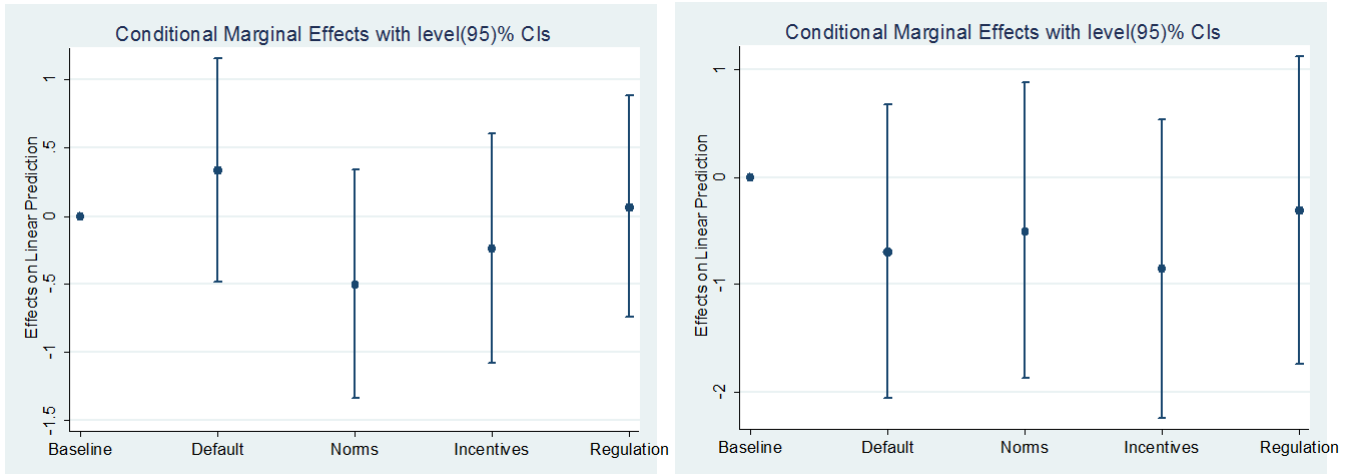
Figure 8: Marginal effects of treatments on stage 2 DG (left) and PD (right) in the low focal point conditions

Appendix B: Experimental Instructions

DG donors, Stage 1.

The first four screens were the same across treatments.

Screen 1. This is the first part of the HIT. You have been paired with another participant. The amount of money you can earn depends only on your choice. You are given 20c and the other participant is given nothing. You have to decide how much, if any, to give to the other participant. The other participant has no choice, is REAL, and will really receive your donation. No deception is used. You will really get the amount of money you decide to keep.

Screen 2. Here are some questions to ascertain that you understand the rules. Remember that you have to answer all of these questions correctly in order to get the completion code. If you fail any of them, the survey will automatically end and you will not get any payment.

Screen 3. What is the donation by you that maximizes your bonus? (Here participants could select any amount of money between 0c and 20c, with a 2c increase)

Screen 4. What is the donation by you that maximizes the other participant's bonus? (Here participants could select any even amount of money between 0c and 20c)

Screen 5 (baseline). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. What amount will you give to the other person? (Here participants could select any even amount of money between 0c and 20c)

Screen 5 (low default). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. What amount will you give to the other person? (Here participants could select any amount of money between 0c and 20c, with a 2c increase. The only

difference with the baseline treatment is that the option 2c was pre-selected).

Screen 5 (high default). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. What amount will you give to the other person? (Here participants could select any amount of money between 0c and 20c, with a 2c increase. The only difference with the baseline treatment is that the option 10c was pre-selected).

Screen 5 (low norm). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. IN AN IDENTICAL EXPERIMENT CARRIED OUT TWO WEEKS AGO ON MTURK, ABOUT 50% OF THE PARTICIPANTS GAVE 2c OR LESS. What amount will you give to the other person?

Screen 5 (high norm). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. IN AN IDENTICAL EXPERIMENT CARRIED OUT TWO WEEKS AGO ON MTURK, ABOUT 50% OF THE PARTICIPANTS GAVE 10c OR MORE. What amount will you give to the other person?

Screen 5 (low incentive). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. FOR EVERY CENT THAT YOU GIVE TO THE OTHER PERSON, WE WILL GIVE YOU BACK 0.5c, PROVIDED THAT YOU GIVE 2c OR LESS. What amount will you give to the other person?

Screen 5 (high incentive). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. FOR EVERY CENT THAT YOU GIVE TO THE OTHER PERSON, WE WILL GIVE YOU BACK 0.5c, PROVIDED THAT YOU GIVE 10c OR MORE. What amount will you give to the other person?

Screen 5 (low regulation). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. YOU WILL NOT BE ALLOWED TO GIVE LESS THAN 2c. IF YOU DECIDE TO GIVE LESS THAN 2c, THE SURVEY WILL STOP AND ASK YOU TO REVISE YOUR DECISION. What amount will you give to the other person?

Screen 5 (high regulation). Congratulations, you have answered both comprehension questions correctly! It is now time to make your choice. YOU WILL NOT BE ALLOWED TO GIVE LESS THAN 10c. IF YOU DECIDE TO GIVE LESS THAN 10c, THE SURVEY WILL STOP AND ASK YOU TO REVISE YOUR DECISION. What amount will you give to the other person?

DG recipients, Stage 1.

The first, the second, and the fourth screens were the same across conditions:

Screen 1. This is the first part of the HIT. Your role is to guess how much a participant (PERSON A) will behave towards another participant (PERSON B).You will win 20c if you make the right guess. The screenshots below show the instructions presented to Person A.

Screen 2. This contained a screenshot of the instructions seen by the donors in their Screen 1.

Screen 3. This contained a screenshot of the instructions seen by the donors in their Screen 5.

Screen 4. What amount do you believe Person A will give to Person B? (Here participants could select any even amount of money between 0c and 20c).

Stage 2,

Dictator Game: These instructions were exactly the same as those of the baseline condition in Stage 2. The only slight difference was in Screen 1, which started with the sentence: This is the second part of the HIT. You have been paired with another participant, different from the one you were paired with before. The amount of money you can earn depends only on your choice.

Prisoner's Dilemma: \underline{Screen 1}. This is the second part of the HIT. You have been paired with another participant, different from the one you were paired with before. The amount of money you can earn depends on your choice and also on the choice of the other participant. You are both given 10c and you have to decide how much, if any, to transfer to the other person. Any cent you decide to transfer will be multiplied by 2 and earned by the other participant. Any cent you decide not to transfer will be earned by yourself, but without being multiplied by any factor. Thus: If both you and the other participant decide to transfer all of your money to each other, then you both end this part of the HIT with 20c If you decide to keep all of your money and the other participant decides to transfer all of his money, then you end this part of the HIT with 30c and the other participant ends this part of the HIT with 0c. If you decide to transfer all of your money and the other participant decides to keep all of his money, then you end this part of the HIT with 0c and the other participant ends this part of the HIT with 30c. If both you and the other participant decide to keep all of your money, then you both end this part of the HIT with 10c. The other participant is REAL, and will really make a choice. No deception is used. You will really get the amount of money you decide to keep plus twice the amount that the other participant decides to transfer to you.

\underline{Screen 2}. Here are some questions to ascertain that you understand the rules. Remember that you have to answer all of these questions correctly in order to get the completion code. If you fail any of them, the survey will automatically end and you will not get any payment.

\underline{Screen 3}. How much should YOU transfer in order to maximize YOUR bonus? (Here participants could select any amount of money between 0c and 10c).

\underline{Screen 4}. How much should YOU transfer in order to maximise THE TOTAL BONUS bonus (i.e. the sum of your and the other participant's bonus)? (Here participants could select any amount of money between 0c and 10c).

\underline{Screen 5}. How much should THE OTHER PARTICIPANT transfer in order to maximize THE TOTAL bonus (i.e., the sum of your and the other participant's bonus)? (Here participants could select any amount of money between 0c and 10c).

\underline{Screen 6}. How much should THE OTHER PARTICIPANT transfer in order to maximize THEIR bonus? (Here participants could select any amount of money between 0c and 10c).

\underline{Screen 7}. Congratulations! You have passed all comprehension questions. It is now time to make your choice. How much do you want to transfer to the other participant? (Here participants could select any amount of money between 0c and 10c).

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei Working Paper Series

Our Note di Lavoro are available on the Internet at the following addresses:

http://www.feem.it/getpage.aspx?id=73&sez=Publications&padre=20&tab=1
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=266659
http://ideas.repec.org/s/fem/femwpa.html
http://www.econis.eu/LNG=EN/FAM?PPN=505954494
http://ageconsearch.umn.edu/handle/35978
http://www.bepress.com/feem/

NOTE DI LAVORO PUBLISHED IN 2015