Reluctant Donors and their Reactions to Social Information

David Klinowski*

*Santiago Centre for Experimental Social Sciences, Nuffield College, University of Oxford, and Department of Economics, Universidad de Santiago de Chile. Email: dklinowski@gmail.com.

July 2019

Abstract

Recent work shows that people donate when asked and at the same time avoid being asked, suggesting that separate mechanisms may influence the decisions to sort into an opportunity to give and of how much to give. In a laboratory experiment with donations to a charity, we examine how social information affects these two decisions. We vary information about the size of a previous donation, and also the timing of the information. We show that both the extensive- and intensive-margin decisions are highly responsive to the information, sometimes in opposite directions. The observed pattern is largely predicted by an individual's tendency to share money in a separate game while preferring not having been asked to share. Our results have implications for whether sorting is a good indicator of welfare in fundraising, and for the use of social information in the design of effective fundraising mechanisms.

Keywords: charitable giving, reluctant altruism, social information
JEL codes: C91, D64, D80, D91, H41, L31

Acknowledgements: The author gratefully acknowledges funding for this project from the Science of Philanthropy Initiative.

1. Introduction

Fundraisers often announce past contributions in order to encourage further donations. We might see, for instance, crowdfunding platforms post previous gifts on a project's website, or charities advertise seed money at the start of a fundraising campaign. The effectiveness of announcing past donations at raising further funding depends on the motivations of future donors. In the pure altruism model, donors care only about the recipient's welfare, and thus their contributions are perfect substitutes for each other (Becker, 1974). The prediction in this case for a charity with continuous production is a negative relationship between announced donations and subsequent donations (Bergstrom, Blume, and Varian, 1986; Varian, 1994). In contrast, if donors derive prestige or status from giving more than others (Romano and Yildirim, 2001), or if they seek to conform to social norms (Bernheim, 1994) or to reciprocate others' behavior (Sugden, 1984; Rabin, 1993), or if donors infer the charity's quality based on how much others give (Vesterlund, 2003), then announced donations and future donations are predicted to be positively correlated.

Several studies directly test these predictions by manipulating the provision of social information (i.e., information about the behavior of others) to potential future donors. Frey and Meier (2004) solicit donations to a university fund and vary information about the percentage of peers that donate. Martin and Randal (2008) vary the monetary contents of a transparent box that is used for collecting donations to an art gallery. Alpizar, Carlsson, and Johansson-Stenman (2008) and Shang and Croson (2009) solicit donations to a charity (a national park and a public radio station, respectively) and vary the size of a previous donation that is announced to the solicitee. In general, these studies find that giving by others has a positive effect on subsequent giving. More recently, Kessler (2017) shows that even nonbinding, non-verifiable announcements of support by others

can increase subsequent donations. Thus, these studies highlight the important role that conformity, reciprocity, and learning have in determining how social information affects giving.[1]

Some recent work sheds further light on the motivations for giving by studying not how donors respond to social information, but rather how they "interact with the ask." This work finds that people avoid the social pressure of engaging with a solicitor by opting out of impending or future solicitations (DellaVigna, List, and Malmendier, 2012; Kamdar et al., 2015), or resist the temptation of giving by avoiding proximity with a solicitor (Andreoni, Rao, and Trachtman, 2017; Trachtman et al., 2015), or use time to quickly think of excuses to decline to donate, and thus a surprise ask increases compliance (Exley and Petrie, 2018).[2] While the underlying mechanisms may vary, these findings all suggest that the decision to participate in a prosocial act—whether to give, or whether to seek or avoid the ask—might involve a different decision process from the decision of how much to give.[3] If so, the question arises of whether social information interacts with this two-stage process. That is, does the individual use social information differently when making decisions about sorting and decisions about how much to give?

[1] Related studies investigate the effect of social information on behavior in dictator, ultimatum, and public goods games in the laboratory, and in general also find a positive effect (Cason and Mui, 1998; Keser and van Winden, 2000; Fischbacher, Gächter, and Fehr, 2001; Bohnet and Zeckhouser, 2004). Other studies use observational data to examine the relationship between donations by others and self (for instance, Andreoni and Scholz, 1998; Smith, Windmeijer, and Wright, 2015). For reviews of motivations for charitable giving more generally, see Vesterlund (2006, 2016) and Andreoni and Payne (2013).

[2] These experiments demonstrate that individuals give when asked, but at the same time avoid the ask. A related finding is that individuals share money when asked, but then retract their giving if they get an opportunity to do so (see for instance Dana, Cain, and Dawes, 2006; Broberg, Ellingsen, and Johannesson, 2007; and Lazear, Malmendier, and Weber, 2012).

[3] Krupka and Croson (2016) study the effect of normative cues on charitable giving and find behavior consistent with a two-stage decision process.

This paper provides evidence that this is the case. We solicit donations to a local charity in a laboratory experiment. The solicitation occurs in two stages: first, participants decide whether to donate; then, only if they accepted, they indicate the size of their contributions. We vary information about the size of a previous donation (a relatively small or large amount), and also the timing of the information (before or after sorting). We demonstrate that both the extensive- and intensive-margin decisions are highly sensitive to the information. But the decisions can respond in opposite directions, suggesting that the individual engages with the information differently at different stages of the solicitation. For example, being informed of a relatively large donation induces substantial sorting out; however, if the same information is disclosed only after the sorting decision, individuals respond by increasing their gift amounts, and no individual gives 0.

This behavior is difficult to rationalize with a theory of conformity, reciprocity, or learning. To try to shed light on the mechanism, a second part of the experiment elicits the participants' intensity of preferences over an opportunity to exit a dictator game without compromising their moral image. In this part of the experiment, participants play a $10 dictator game with an anonymous recipient in the room, and then face an unexpected choice between following through with their dictator decision, or receiving $9 and leaving the recipient with $0. A computer makes this choice for them, but the participants must indicate the probability distribution over the two options that the computer will use to make the choice. The probability distribution a participant indicates is thus a measure of their "reluctant altruism"—their willingness to quietly retract their giving in the dictator game.[4] We find that this measure is highly predictive of how a participant responded to

[4] The mechanism we implement allows us to offer dictators the opportunity to exit the dictator game and leave the recipient uniformed of their exit, while keeping all participants in the same room. This contrasts with previous work that implements the quiet exit by keeping dictators and

the social information during the solicitation for a donation to the charity. The stronger a participant's willingness to exit the dictator game, the greater the likelihood that the participant sorted out of the solicitation in response to a relatively large donation by others, or sorted into the solicitation in response to a relatively small donation by others. Moreover, reluctant altruists significantly increased their giving at the intensive margin in response to all information treatments, while other participants (those that either shared or not shared in the dictator game but did not indicate a willingness to exit) tended not to change the amounts they donated in response to any information treatment.

We interpret the observed responses to the social information, and their correlation with the reluctant altruism measure, as suggestive of *social pressure avoidance* (DellaVigna, List, and Malmendier, 2012; Kamdar et al., 2015), with a plausible avoidance mechanism being *motivated reasoning* (Kunda, 1990; Gino, Norton, and Weber, 2016). According to this interpretation, individuals do not want to give their money to the charity, but they also do not want to think of themselves as selfish. They draw from features of the decision context in order to come up with excuses to avoid giving that do not compromise their moral image; and if they are unable to find such excuses, or if sorting is not possible, they give. Social information can influence giving by affecting both channels—it can increase the pressure to give a large-enough amount so that one

recipients in separate rooms (Dana, Cain, and Dawes, 2006; Broberg, Ellingsen, and Johannesson, 2007; and Lazear, Malmendier, and Weber, 2012). As holding participants in separate rooms may complicate experiment logistics and decrease the number of usable observations, a mechanism such as ours could serve as a feasible way to elicit reluctant altruism in general settings, and thus provide a more nuanced picture of an individual's altruistic preferences than what is obtained with the dictator game alone. This is especially relevant in light of concerns that giving in the dictator game does not necessarily reflect altruism (Bardsley, 2008; List, 2007), and recent interest in eliciting moral preferences (Benabou, Falk, and Tirole, 2018a). Klinowski (2018) shows that responses to this mechanism can help to explain gender differences in giving.

does not appear selfish, and it can facilitate (or invalidate) justifications to sort out, if sorting is possible. For example, learning that others have donated a large amount adds social pressure on the individual, but at the same time can help the individual to come up with justifications to sort out such as "I cannot afford to give that much," or "the charity must have raised enough money already." Information about a relatively small donation, on the other hand, makes such excuses less plausible, inducing the individual to sort in.[5]

Although the experiment cannot test this interpretation directly, recent work demonstrates the important role that motivated reasoning can play in shaping giving behavior.[6] An alternative mechanism that has recently been shown to explain a tendency of individuals to donate when asked while at the same time avoiding the ask, is *empathy avoidance* (Shaw, Batson, and Todd, 1994; Andreoni, Rao, and Trachtman, 2017). In this view, individuals get joy from giving, and being asked to give triggers in them empathy and an impulse to give that is hard to resist. Individuals that are aware of their self-control problem, or "empathic vulnerability", avoid the empathic stimulus by avoiding being asked to give (Andreoni, Rao, and Trachtman, 2017). We suggest that a way to distinguish between social pressure avoidance and empathy avoidance in our experiment is to examine the participants' personality characteristics, which we collect in an exit

[5] This is consistent with Cialdini and Schroeder (1976), who propose that solicitors can invalidate excuses for noncompliance by legitimizing small contributions. They write: "When the most minimal of monetary donations is said to be acceptable, excuses for failing to help that might ordinarily be offered (e.g., `I can't afford to give to all the various charities'; `We're too low on money this week,' etc.) become inapplicable" (Cialdini and Schroeder, 1976). At the same time, solicitees are unlikely to respond with a small donation because doing so would "jeopardize personal image."

[6] Exley (2016) shows that people evaluate risk self-servingly in order to avoid contributing to a charity, and Exley and Petrie (2018) show that people can find ways to decline a solicitation simply by being given time to decide. For the role of motivated reasoning in other economic domains, see Haisley and Weber (2010), Gino, Norton, and Weber (2016) and Exley and Kessler (2018).

questionnaire. We hypothesize that if reluctant altruism is due to social pressure, reluctant altruists will score relatively high on Neuroticism, a trait associated with increased anxiety and rumination (Muris et al., 2005), with experiencing the world as threatening (Baumeister and Vohs, 2007), and with negative effects on decision-making from situations involving peer pressure and social evaluation (Byrne, Silasi-Mansat, and Worthy, 2015). On the other hand, if reluctant altruism is driven by empathy avoidance, reluctant altruists will score relatively high on Agreeableness, a trait associated with dispositional empathy and greater tendency to offer help, even to strangers (Baumeister and Vohs, 2007). Following the work by Ottoni-Wilhelm and Bekkers (2010) and Bekkers and Ottoni-Wilhelm (2016), we also examine the role of the internalization of a moral value that one should help those in need in predicting reluctant altruism and in mediating the empathy-giving link. We find weak support for Agreeableness as a predictor of reluctant altruism, and strong support for Neuroticism and a principle that one should help as predictors. At the same time, Agreeableness and an internalization of a moral value that one should help (but not Neuroticism) are associated with giving and subsequently *not* exiting the game. Thus, the results are consistent with the view that reluctant altruists in the experiment are individuals that feel social pressure from being asked to give and look for justifications to sort out, with these motivations potentially affecting how they responded to social information in the solicitation. On the other hand, "willing altruists" are more likely motivated by empathy and adherence to the imperative that one should help, and thus less susceptible to social information in the solicitation.[7]

[7] This is consistent with the framework in Benabou, Falk, and Tirole (2018b) that separates motivations for moral behavior into *narratives* and *imperatives*.

Together our results provide evidence that social information can affect giving in a two-stage decision process. This has practical implications for the use of social information in the design of effective fundraising mechanisms. In particular, it points to the potential value to fundraisers from carefully adjusting not just the content of the information, but also its timing, which has received little attention in the literature. As we discuss at the end of the paper, it also raises questions about whether we can conclude that a solicitation is welfare-enhancing or welfare-decreasing based on observed sorting, given that entry in our experiment was highly sensitive to an intervention that did not change the cost of giving.[8]

In the remaining of the paper, Section 2 details the experimental design, Section 3 presents the results, and Section 4 closes with a discussion. The Appendix presents a theoretical model that formalizes reluctant giving and illustrates that this behavior can result in equilibrium if donors care about their moral image and have a tendency to seize on the decision context to justify selfishness.[9]

[8] Thus, our paper enters the literature on the demand side of giving, that examines fundraising strategies and their effect on giving. See for instance List and Lucking-Reiley (2002) on seed money, Eckel and Grossman (2003) and Karlan and List (2007) on rebates and matches, and Argo et al. (2019) on goal setting. Our paper also fits in the literature that studies the effect of social influence more generally on giving behavior (see for instance Meer, 2011; Linardi and McConnell, 2011; Samek and Sheremeta, 2014; Castillo, Petrie, and Wardell, 2014, 2017), and a large literature that examines the effect of social information on prosocial behavior more generally, including energy consumption (Allcott, 2011), tax compliance (Hallsworth et al., 2017), and prosocial job choice (Coffman, Featherstone, and Kessler, 2017).

[9] Our model is only illustrative, and other models have more detailed treatments of related concepts such as image concerns and information avoidance. Andreoni and Bernheim (2009), in their online appendix, explain the decision to exit a dictator game in a model where dictators care about their moral image, by assuming an exogenous utility value to the exit option. Relative to their work, our model derives the utility from declining to give endogenously, from beliefs about behavior in equilibrium. Grossman and Van der Weele (2017) model the binary decisions to acquire information about the social benefits of an action, and to subsequently engage in that action. Our equilibrium is similar to their "ignorance equilibrium." Relative to their work, we model the intensive margin of giving as a continuous variable, and formalize reluctant altruism by comparing equilibrium behavior in a game with and without the possibility of sorting. Related models of

2. Experimental design

Each experiment session consisted of four parts. First, participants generated earnings by solving two real-effort tasks. Then they were solicited for a donation from their earnings to a local charity. After responding to the solicitation, participants played a $10 dictator game with an anonymous participant in the session, and then had an opportunity to quietly exit the game and obtain a relatively selfish outcome. Finally, participants completed a questionnaire that elicited demographic and personality information. Participants knew at the start of the session that the experiment consisted of four parts, but received instructions for each part only at the beginning of the corresponding part (instructions were shown on the screen and also read aloud by the experimenter). In the online Appendix, we include the experiment instructions and screenshots. Below we describe each part of the experiment in more detail.

Part 1: real-effort tasks. Participants solved two computerized real-effort tasks. The first was a modified version of the slider task (Gill and Prowse, 2012) in which participants had 90 seconds to slide seven scroll bars to their center positions. Participants earned $1.50 for successfully sliding all seven bars in the allotted time, and $0 otherwise. The second task involved clicking a button located at the center of the screen exactly when a timer shown next to the button displayed the number 15. The timer counted seconds elapsed since the start of the task; thus, participants had a one-second window, fifteen seconds after the start of the task, to click the button. They earned $1.50 for doing so successfully, and $0 otherwise. The two tasks appeared one after the other a total of five rounds; thus, participants could earn at most $15 in this part of the experiment. The

social image and dual-self behavior are Bodner and Prelec (2003), Bénabou and Tirole (2006), and Ellingsen and Johanesson (2008).

tasks were designed to be low in difficulty, in order to maximize the number of participants that earned the $15 while at the same time providing them with earned (rather than windfall) money to be used later in the solicitation. We also included these effort tasks to mask the experiment's main research question (studying giving behavior), in order to minimize experimenter demand effects (De Quidt, Vesterlund, and Wilson, 2019).

Part 2: solicitation. After completing the real-effort tasks and being informed of their earnings, participants were solicited for a donation from their earnings to *Pittsburgh Cares.*[10] The solicitation occurred via the computer in two stages. In the first stage, all participants received a brief description of Pittsburgh Cares and of the protocol that would be used to ensure confidentiality of their decisions.[11] Participants were also told that their donations would be matched one-to-one by the sponsor of the study (we matched contributions in order to encourage donations). At this stage, participants had to decide whether they wished to donate (a binary yes-no decision), and were told that if they accepted, they would then have to indicate the amount they wished to donate (anywhere from $0 to their earnings in Part 1). If the participant decided not to donate, the solicitation ended there. Only the participants that decided to donate advanced to the second stage of the solicitation. In the second stage, participants had to indicate how much they wished to give (where, again, donating $0 was explicitly allowed).

[10] Pittsburgh Cares is a local nonprofit organization in Pittsburgh that promotes volunteerism by connecting other local nonprofits in need of volunteers with individuals seeking to volunteer their time. Pittsburgh Cares does not focus on a particular population or issue, which helps to lessen concerns that donations in the experiment are driven by unobserved characteristics of the participants (such as political views).

[11] Participants received an envelope and a receipt form that they needed to fill if they wanted to get a receipt from the charity. To prevent participants from identifying who donated and who did not based on their filling the envelope, all participants were asked to put their receipt form (even if left blank) in their envelope, seal the envelope, and leave the envelope on their desks.

Table 1. Treatments and sample sizes

Treatment	Information in stage 1	Information in stage 2	Observations
No-information	None	None	44
50¢-stage-1	50¢	50¢	64
50¢-stage-2	None	50¢	60
$5-stage-1	$5	$5	74
$5-stage-2	None	$5	66

The experiment consisted of five treatments, administered in a between-subjects design that varied the content and timing of the social information participants received during the solicitation. Participants in the *no-information* treatment received no social information. Participants in the *50¢-stage-1* treatment were informed in stage 1 that another participant in a previous session had donated 50¢. Participants in the *50¢-stage-2* treatment received no social information in stage 1 and were informed in stage 2 (if they advanced there) that another participant in a previous session had donated 50¢. Participants in the *$5-stage-1* treatment were informed in stage 1 that another participant in a previous session had donated $5. Participants in the *$5-stage-2* treatment received no social information in stage 1 and were informed in stage 2 (if they advanced there) that another participant in a previous session had donated $5. In all information treatments, the information was shown by adding the following sentence to the section on the screen where participants had to enter their decisions: "For your information, a participant in a previous session donated [X] dollars." We conducted the *no-information* sessions first, and observed non-zero donations ranging from 50¢ to $5. We then referred to these two values in the information treatments, to provide information about relatively small and large donations without deception. Table 1 summarizes the experiment treatments. All participants in any given session received the same treatment.

Part 3: dictator game and exit. Following the solicitation, participants in the session were paired anonymously. Each participant received $10 and had to decide how to split that money between themselves and their partner, in multiples of $1 (namely, 10–0, 9–1, …, 0–10). They were informed that only one decision among the pair would be randomly selected for payment at the end of the experiment. Once all participants decided how to split the $10, they faced a previously unannounced second choice to make. For each participant, the computer would determine the allocation that would become the participant's decision in this part of the experiment. The computer could either confirm the $10 split the participant had just indicated, or replace it with an allocation of $9 for the participant and $0 for the partner. The computer would choose one of these two options at random, but the participant had to select the probability distribution over the two options. Specifically, the participant had to indicate a value between 10 and 90 percent (inclusive) for the probability of choosing the $9–$0 allocation, and the complement value was automatically assigned as the probability of confirming the original $10 split. While making this decision, participants were informed that, if they and this part of the experiment were randomly selected for payment, the partner would be informed only of the final allocation chosen by the computer, and never informed of their original $10 split (unless it was confirmed by the computer) or of the probability distribution they indicated.

This part of the experiment was designed to elicit a measure of reluctant altruism. The decision to split the $10 measures the participant's willingness to share money with a stranger, while the decision over the probability distribution measures the willingness to quietly retract one's giving in the dictator game in favor of obtaining $9 and leaving the partner with $0. The retraction is quiet, in the sense that if the computer chooses $9–$0 and this is selected for payment, the partner

never learns the participant' original allocation decision or the selected probability distribution. The partner cannot even infer that the participant moved the odds at all in favor of $9–$0, because there is always at least a 10-percent probability that the computer chooses this option regardless of what probability weights the participant indicated.[12]

Part 4: questionnaire. Participants completed an unincentivized questionnaire before receiving feedback on their payments for part 3 and their total earnings. The questionnaire included a 44-item Big Five personality inventory (John, Donahue, and Kentle, 1991) that produces a score for each of the five personality dimensions: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness. The questionnaire also included a 3-item instrument employed by Ottoni-Wilhelm and Bekkers (2010) to elicit an individual's internalization of a moral value that one should help those in need—a value known as the *principle of care* (Hoffman, 2000; Batson, 2011). For this, participants indicate their agreement (from 1 to 5) with the following statements: "*People should be willing to help others who are less fortunate*" (principle of care 1), "*These days people need to look after themselves and not overly worry about others*" (principle of care 2), and "*Personally assisting people in trouble is very important to me*" (principle of care 3). Note that stronger agreement with the second statement implies weaker endorsement of a value that one should help. Finally, participants provided demographic information such as age, gender, and race.

[12] This mechanism was inspired by Dana, Cain, and Dawes (2006). In their study, dictators and recipients are seated in separate rooms, and the exit decision is a binary choice between confirming one's dictator decision or implementing a (9,0) allocation that also leaves the recipient unaware that the game was played. By eliciting probabilities between 10 to 90 percent rather than a binary choice, we are able to implement the quiet exit with all participants in the same room. This also generates a continuous measure of preferences over exit, with one observation per participant, and maintains common knowledge of the instructions at all times of the game.

After completing the questionnaire, participants received feedback on part 3 and their individual earnings for the entire experiment. Total earnings were the sum of earnings in part 1, net of donations, and part 3 (no show-up fee was added). The experiment was conducted at the Pittsburgh Experimental Economics Laboratory (PEEL) from April to September 2014. A total of 308 undergraduate students participated in 14 sessions. Each session involved 16 to 30 participants and lasted approximately one hour. The experiment was programmed in z-Tree (Fischbacher, 2007).

3. Results

Not all participants earned the maximum possible ($15) in the effort tasks: 82 percent of all participants earned $15, and 98 percent earned either $13.50 or $15 (the two largest possible amounts). As giving may have depended on earnings, throughout the analysis we control for the participant's pre-donation earnings by including an indicator of whether the participant earned $15 in the effort tasks. We also control for the participant's age and gender.

3.1. Intent to donate

Figure 1 shows the fraction of participants that accepted to donate in stage 1, given the information received in this stage. With no information, 35 percent accepted to donate. This fraction dropped to 22 percent with information about a $5 donation, and grew to 58 percent with information about a 50¢ donation. Note that the no-information condition in stage 1 combines observations from the *50¢-stage-2*, *$5-stage-2*, and *no-information* treatments, as in these three treatments participants received no social information in stage 1. As expected, intent to donate is statistically similar across these three treatments: 35 percent in *no-information*, 30 percent in *$5-stage-2*, and 43 percent in the *50¢-stage-2*. In a probit regression that estimates intent to donate and restricts the sample to

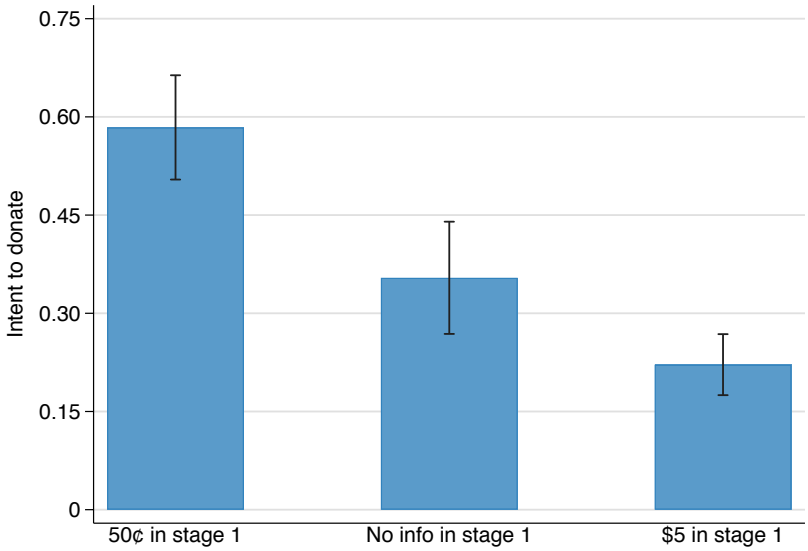

Figure 1. Intent to donate to the charity by information in stage 1
Notes: Black lines show 95-percent confidence intervals.

these three treatments (controlling for age, gender, and pre-donation earnings), the marginal effect of each of the latter two treatments relative to the *no-information* treatment is insignificant (p=0.422 and p=0.346, respectively). And a test of joint significance of the treatments in the same regression is insignificant (p=0.409).

Table 2 explores the effect of the social information on the intent to donate with regression analysis. For now we focus on the *Pooled* column, which considers the entire participant sample. The column shows results from a probit regression that predicts the decision to donate in stage 1 with the information seen in stage 1, controlling for the participant's age, gender, and pre-donation earnings. The first two rows show marginal effects (in percentage points) of the information, relative to receiving no information. The last row shows the mean intent to donate when no information is received. We see that information had a large impact on intent to donate. Information about a 50¢ donation increased intent to donate by 23 percentage points (a 65-percent

Table 2. Effects of social information on the intent to donate to the charity

	Pooled	By type		
		Selfish	Reluctant	Generous
50¢	0.230***	0.150	0.212*	0.203
	(0.064)	(0.108)	(0.111)	(0.167)
$5	-0.133***	-0.075	-0.127	-0.334
	(0.051)	(0.063)	(0.110)	(0.209)
No-info mean	0.354	0.231	0.460	0.588

Notes: Marginal effects (in percentage points) of the information seen in stage 1, relative to receiving no information in that stage. Estimates from a probit regression predicting the decision to donate in stage 1, controlling for the participant's age, gender, and pre-donation earnings (estimates *by type* come from a single regression that includes in addition the participant's type and its interaction with the information received in stage 1). For both regressions, N=308. Standard errors clustered at the experimental session in parentheses. *p<0.1, **p<0.05, ***p<0.01.

increase), while information about a $5 donation depressed intent to donate by 13 percentage points (a 38-percent drop), both changes statistically significant.

In principle, intent to donate need not equal actual donation rates, since participants could have accepted to donate in stage 1 and donate zero in stage 2. However, this occurred very rarely: once in the *no-information* treatment and twice in the *$5-stage-1* treatment. Over all treatments, the actual donation rate conditional on intending to donate is 97 percent. Thus, intent to donate and donation rates were virtually equivalent.

3.2. Donation size

Table 3 estimates treatment effects on the amount donated. To give a fuller picture of the effects on the distribution rather than just at the mean, Table 3 presents three separate analyses. *Donation size* shows marginal effects of each information treatment at the mean, relative to the no-information treatment, from a single OLS regression. *Prob. of donating more than $1* shows marginal effects on the probability of donating more than $1, relative to the no-information

treatment, from a single probit regression. And *Distribution shift* shows the test statistic and corresponding p-value of pairwise Mann-Whitney U tests that compare the full distribution of donations under the corresponding information treatment to the distribution under no information (a positive statistic indicates a shift toward larger amounts). Columns under *All participants* present results before sorting (that is, not conditioning on intent to donate). Donation size in this case corresponds to the expected revenue to the charity per experiment participant. This value is equal to $0.44 in the *no-information* treatment, and increased with all information treatments, by an amount that ranged from $0.07 (a 16-percent increase) in the *$5-stage-1* treatment to $0.55 (a 124-percent increase) in the *50¢-stage-1* treatment. Only the latter is a statistically significant increase, and also economically large, since it brought well over twice more revenue to the charity. Thus, providing information about a 50¢ donation in stage 1 turned out to be the most profitable fundraising strategy in the experiment.

Turning to the likelihood of making a donation larger than $1, 8 percent of all participants in the *no-information* treatment made a donation larger than $1. All information treatments led to a large increase in the probability of donating more than $1, by an amount that ranged from 6 percentage points (a 78-percent increase) in the *$5-stage-2* treatment to 12 percentage points (a 151-percent increase) in the *50¢-stage-1* treatment. This is notable because one might have predicted a drop in the likelihood of receiving relatively large donations for treatments informing of a 50¢ donation, either because such information could have signaled low charity quality or signaled that relatively small donations were socially acceptable. We come back to this point later when we interpret the results. Finally, looking at the shifts in the distribution, these reflect mainly the changes in sorting

Table 3. Effects of social information on the amount donated to the charity

	All participants			Conditional on intent to donate		
	Donation size	Prob. of donating more than $1	Distribution shift	Donation size	Prob. of donating more than $1	Distribution shift
$5 stage 1	0.069	0.077***	-1.352	1.073**	0.528***	2.128
	(0.119)	(0.024)	p=0.176	(0.465)	(0.082)	p=0.033
$5 stage 2	0.123	0.062**	-0.004	0.606***	0.260***	2.141
	(0.143)	(0.031)	p=0.997	(0.190)	(0.070)	p=0.032
50¢ stage 1	0.551**	0.119***	2.592	0.391	0.115*	1.359
	(0.190)	(0.028)	p=0.010	(0.239)	(0.062)	p=0.174
50¢ stage 2	0.214	0.111*	1.082	0.266	0.256***	1.563
	(0.214)	(0.063)	p=0.279	(0.221)	(0.092)	p=0.118
No-info mean	0.444	0.079	-	1.295	0.210	-
N	308	308	-	114	114	-

Notes: Marginal effect of each information treatment, relative to the no-information treatment. *All participants* include observations from the full sample, and *Conditional on intent to donate* restricts the sample to participants that accepted to donate in stage 1. Estimates for *Donation size* from an OLS regression that predicts the donation amount with the treatment, and estimates for *Prob. of donating more than $1* from a probit regression that predicts the probability of donating more than $1. Both regressions control for the participant's age, gender, and pre-donation earnings. *Distribution shift* shows the Mann-Whitney U statistic of a test that compares the distributions of the donation amount for the corresponding information treatment and the no-information treatment. A positive statistic indicates an upward shift in the distribution. Standard errors clustered at the experimental session in parentheses. *p<0.1, **p<0.05, ***p<0.01.

caused by information in stage 1: a shift downward in the *$5-stage-1* treatment, and a shift upward in the *50¢-stage-1* treatment.

Looking at donations conditional on intent to donate allows us to examine treatment effects at the intensive margin of giving. Figure 2 plots the distribution of donation amounts over all treatments, conditional on intent to donate. While the average donation is $1.74, the majority of donations are concentrated over smaller amounts. Approximately 58 percent of all donations are no larger than $1, and 42 percent of all donations are exactly equal to $1.

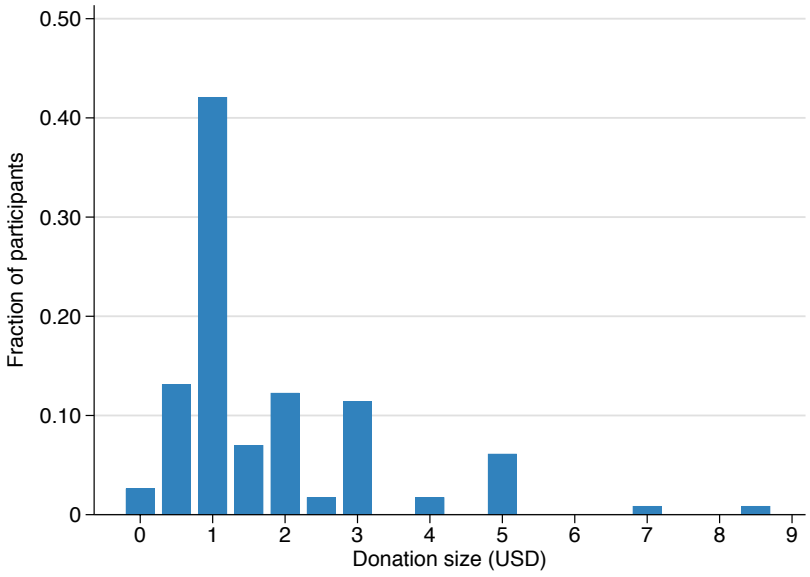

Figure 2. Amount donated to the charity conditional on intent to donate

In Table 3, columns *Conditional on intent to donate*, we repeat the analysis but now restrict the sample to participants that sorted in. The mean donation in the no-information treatment is $1.30. All information treatments led to an increase in the mean donation, by $0.27 (a 21-percent increase) in the *50¢-stage-2* treatment, $0.39 (a 30-percent increase) in the *50¢-stage-1* treatment, $0.61 (a 47-percent increase) in the *$5-stage-2* treatment, and $1.07 (an 83-percent increase) in the *$5-stage-1* treatment. The latter two are statistically significant. Similarly, all information treatments led to a substantial increase in the likelihood that the donation was larger than $1, relative to the *no-information* treatment. The probability of making a donation larger than $1, conditional on intent to donate, was 21 percent in the *no-information* treatment, and grew by 12 percentage points (a 55-percent increase) in the *50¢-stage-1* treatment, by 26 percentage points (a 123-percent increase) in the *50¢-stage-1* and *$5-stage-2* treatments, and by 53 percentage points (a 251-percent increase) in the *$5-stage-1* treatment. This increase in the intensive margin due to the information treatments is also illustrated by the upward shift in the distribution of donations in each information treatment relative to the *no-information* treatment (although the shift is statistically significant for *$5-stage-1* and *$5-stage-2* only).

3.3. Possible mechanisms

What drives the observed responses to the social information at the extensive and intensive margins of giving? It is difficult to reconcile the observed behavior with a model of altruism. The classical model of pure altruism with continuous production (Bergstrom, Blume, and Varian, 1986; Varian, 1994) assumes that agents care about aggregate donations to the charity, and thus predicts that participants give more when informed of a 50¢ donation than when informed of a $5 donation. This crowd-out prediction matches the observed effects at the extensive margin, where information about a $5 donation depressed intent to donate relative to information about a 50¢ donation. But it does not match the effects at the intensive margin, where, if anything, the *$5-stage-2* treatment elicited larger donations than the *50¢-stage-2* treatment. Neither do participants appear to be acting on the information they received by making inferences about the quality of the charity, as they might have done had they held asymmetric information about the charity's quality (Vesterlund, 2003; Potters, Sefton, and Vesterlund, 2007). If participants inferred quality based on others' donations, this would explain that information in stage 2 about a $5 donation elicited larger donations, but not that the same information, when provided in stage 1, reduced intent to donate. A similar argument seems to apply against a view that participants sought to conform to social norms (Bernstein, 1984) or reciprocate others' donations (Sugden, 1984; Rabin, 1993). If norm-compliance or reciprocity led participants to increase donations in response to information in stage 2 about a $5 donation, why did the same information in stage 1 reduce intent to donate?

A satisfactory explanation, then, needs to account for the opposing responses to the information at the extensive and intensive margins of giving. The findings appear more in line with recent work that pays particular attention to how individuals "interact with the ask." This work finds that

individuals give unless they can seize on opportunities or justifications to decline to give, or to avoid being asked to give (DellaVigna, List, and Malmendier, 2012; Kamdar et al., 2015; Andreoni, Rao, and Trachtman, 2017; Trachtman et al., 2015; Exley and Petrie, 2018). In our experiment, a tendency to exploit the context to find plausible justifications to sort out may have led participants to respond to the social information differently depending on whether the information was provided before or after sorting. For instance, learning that another participant donated a relative large amount may have increased the social pressure on the individual to make a large-enough donation, but at the same time, if the information was provided before sorting, it may have allowed the individual to come up with justifications to sort out, such as "I cannot afford to give that much," or "the charity must have raised enough money already." Information about a small donation of others may have countered such justifications, resulting in greater sorting in (and not in smaller donations at the intensive margin if participants considered that donating 50¢ was too small to be morally acceptable). According to this description, the individual engages with the social information through a combination of *social pressure avoidance* and *motivated reasoning*. Individuals prefer not to give, and look for justifications to say no that do not seem selfish. Social information can affect sorting by making it more or less difficult for the individual to find such justifications. When sorting is not possible, there is increased social pressure on the individual to give more as others give more.

In the next section, we provide further support for this interpretation by showing that participants that shared money in a dictator game and subsequently took advantage of an opportunity to exit quietly—a behavior consistent with giving due to pressure from being asked and then seizing on an opportunity to opt out—were more likely to have responded to the social information in the

previous part of the experiment in the manner described above. That is, we observe a correlation in behavior across parts of the experiment that is consistent with our interpretation and is difficult to rationalize with other motivations.

3.4. Identifying reluctant altruists

Following the solicitation part of the experiment, participants played a $10 dictator game with a subsequent possibility to exit (by indicating the probability of implementing a $9–$0 outcome instead). Figure 3 shows the fraction of participants that transferred a given amount in the dictator game. The color shading indicates the observed exit probability, divided in three cases: 10 percent (dark shade), 11 to 89 percent (medium shade), and 90 percent (light shade).

Ignoring the decision to exit, a total of 47 percent of participants transferred a positive amount, and 21 percent split the endowment evenly; the mean transfer over all participants was $1.80. This bimodal distribution with the largest mass at 0 is in line with typical behavior in the dictator game (Camerer, 2003). The amount transferred did not vary significantly across treatments: when regressing the amount transferred on the treatment alone, an F-test fails to reject joint non-significance of the treatments ($p=0.357$), and a chi-squared test fails to reject that the observations across treatments come from the same distribution ($p=0.338$).

Focusing on the decision to exit, 41 percent of participants selected a probability larger than 10 percent. Of those, 24 percent selected 50 and 33 percent selected 90. No payoff-based preference can explain why a participant would choose an exit probability other than 10 percent, since $9–$0 is strictly less efficient than any dictator game allocation, and Pareto inferior (for the dictator and

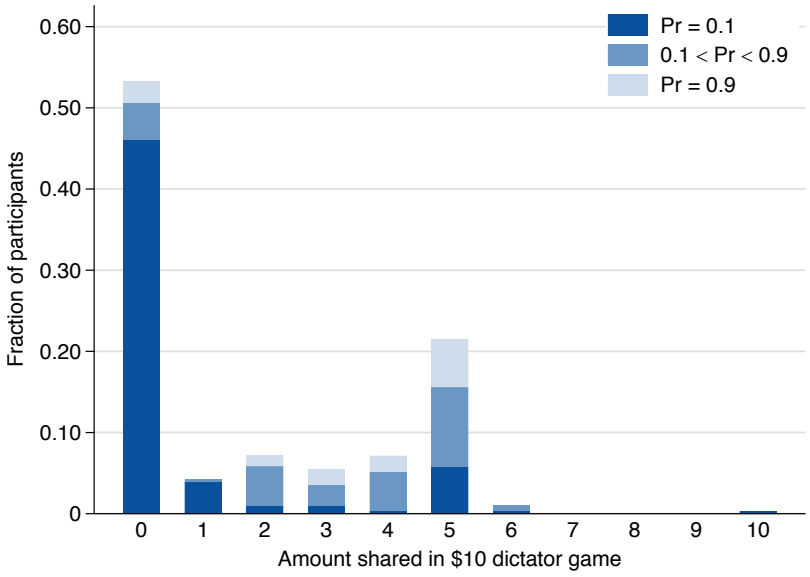

Figure 3. Distribution of giving and exit in the dictator game

the recipient) to a dictator game allocation. However, such choice is consistent with giving in response to the pressure from being asked to, while preferring a selfish option that does not compromise one's moral image. In addition, the observed exit pattern suggests participants did not exit at random or by mistake, but rather did so in order to increase their own payoff: among participants that gave $2 or more to the recipient, 79 percent chose an exit probability larger than 10 percent, while only 20 percent did so among participants that gave $1 or $0.

To explore how exiting the dictator game correlates with the response to social information in the solicitation part of the experiment, we classify participants into three types based on their behavior in the dictator game: (i) *Selfish* denotes participants that transferred $0 or $1 and chose an exit probability equal to 10 percent; (ii) *Reluctant* denotes participants that chose an exit probability

larger than 10 percent, independently of the amount they transferred;[13] and (iii) *Generous* denotes participants that transferred $2 or more and chose an exit probability equal to 10 percent.

The distribution of types in the sample is 50 percent *selfish*, 41 percent *reluctant*, and 9 percent *generous*. The distribution did not vary significantly across treatments: a chi-squared test fails to reject that the observations across treatments come from the same distribution ($p=0.133$), and, when predicting the likelihood of being classified as reluctant on the treatment alone in an OLS regression, an F-test fails to reject joint non-significance of the treatments ($p=0.347$). We also find no evidence of moral compensation in either direction—either moral licensing, where individuals display selfish behavior after initially acting morally (Zhong and Liljenquist, 2006; Merritt, Effron, and Monin, 2010), or moral cleansing, where individuals restore moral standing lost to initial selfish behavior by subsequently acting generously (Sachdeva, Iliev, and Medin, 2009). That is, we do not observe that participants transferred smaller (larger) amounts in the dictator game after making larger (smaller) donations in the solicitation. In fact, participants were consistent in their prosocial behavior across parts: transferring an additional $1 in the dictator game is associated with donating an additional $0.40 in the solicitation ($p=0.012$), and having transferred a positive amount in the dictator game is associated with an increase of 27 percentage points in the intent to donate in the solicitation ($p<0.01$).[14] Thus, types appear stable across treatments, and there seems to be no "moral wealth effects" across parts of the experiment.

[13] 23 *reluctant* participants shared $0 or $1 in the dictator game. This behavior is consistent with a concern for not appearing selfish, as it might indicate that the participant was willing to pay a cost of $1 (or impose a cost of $1 on the recipient) in order to hide from the recipient the fact that they chose a relatively selfish split in the dictator game.

[14] This is in line with other findings of individual consistency in prosocial behavior. For instance, Dariel and Nikiforakis (2014) find that cooperation in a public-good game predicts reciprocation in a subsequent gift-exchange game.

3.5. Reluctant altruists and their response to social information

To examine the relationship between behavior in the dictator game and the response to social information in the solicitation, columns *By type* in Table 2 replicate the analysis of the effect of information on intent to donate, but now adding to the regression an indicator of the participant's type and the interaction between type and the information seen in stage 1. We find a natural progression in intent to donate in which, with no information in stage 1, *selfish* participants are least likely to donate, followed by *reluctant* and then *generous* types. This is a first indication that the types capture something informative about giving behavior. All types increased their intent to donate with information about a 50¢ donation and decreased it with information about a $5 donation, though only statistically significantly so for *reluctant* participants in the 50¢ condition. However, there is substantial heterogeneity in the change in intent to donate within the *reluctant* type not reflected in this result. The intensity of preferences over exiting the dictator game—that is, the selected exit probability—was highly predictive of the effect of social information on the extensive margin of giving, in a way consistent with the mechanism described above.

To illustrate this, Figure 4 plots the predicted intent to donate conditional on the information seen in stage 1 for *reluctant* participants only, as a function of their selected exit probability. A larger probability is associated with greater likelihood of sorting in in the 50¢ condition and of sorting out in the $5 condition, particularly at the upper range of the exit probability. Thus, while *reluctant* participants that selected an exit probability no larger than 50 were equally likely to sort in in the 50¢ and $5 conditions (intent to donate equal to 0.44 and 0.45, respectively), *reluctant* participants that selected an exit probability equal to 90 were almost five times more likely to sort in with information about a 50¢ donation than with information about a $5 donation (intent to donate equal

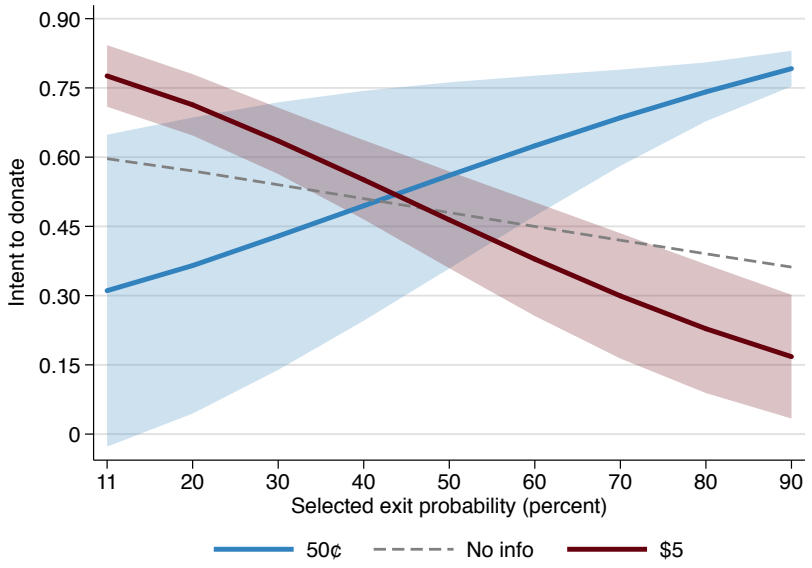

Figure 4. Effect of social information on intent to donate to the charity as function of exit in the dictator game

Notes: Shade areas show 90-percent confidence intervals

to 0.84 and 0.18, respectively). This correlation between changing one's intent to donate in response to social information and showing a willingness to exit the dictator game is consistent with the idea that the individual donates unless they can find an excuse not to, and that social information can make it more or less difficult for the individual to find a justification to sort out.

To investigate the effects on the amount donated, Table 4 replicates the analysis of the treatment effects on the donation size and the probability of donating more than $1 in Table 3, but now the predictors are the information treatment, the participant's type, and the interaction of the two. We combine *selfish* and *generous* types into a single category (labelled *other* in Table 4) because dividing the sample into three types and five treatments leaves some cells with too few observations for the analysis to be feasible. For the same reason, we omit the pairwise comparisons of the distributions of amounts donated, as there are cases with too few observations for a pair-wise comparison to be feasible. The results show that *reluctant* participants were in general more

Table 4: Effects of social information on the amount donated to the charity, by participant type

	All participants				Conditional on intent to donate			
	Donation size		Prob. of donating more than $1		Donation size		Prob. of donating more than $1	
	Reluctant	Other	Reluctant	Other	Reluctant	Other	Reluctant	Other
$5 stage 1	0.428	-0.160	0.217**	-0.017	1.446**	0.911	0.735***	0.339*
	(0.291)	(0.332)	(0.094)	(0.061)	(0.626)	(1.083)	(0.164)	(0.205)
$5 stage 2	0.630**	-0.268	0.185*	-0.043	1.337**	-0.234	0.457***	-0.028
	(0.290)	(0.291)	(0.103)	(0.053)	(0.559)	(0.704)	(0.165)	(0.198)
50¢ stage 1	0.481**	0.639	0.118**	0.131*	0.313	0.786	0.149*	0.151*
	(0.211)	(0.453)	(0.055)	(0.076)	(0.378)	(0.858)	(0.082)	(0.085)
50¢ stage 2	0.796**	-0.154	0.368**	-0.052	1.031**	-0.346	0.614***	-0.126
	(0.317)	(0.305)	(0.183)	(0.032)	(0.516)	(0.718)	(0.106)	(0.107)
No-info mean	0.282	0.575	0.029	0.128	0.778	1.673	0.053	0.367

Notes: Marginal effect of each information treatment, relative to the no-information treatment. *All participants* include observations from the full sample, and *Conditional on intent to donate* restricts the sample to participants that accepted to donate in stage 1. Estimates for *Donation size* from a single OLS regression that predicts the donation amount, and estimates for *Prob. of donating more than $1* from a single probit regression that predicts the probability of donating more than $1. In both regressions, the regressors are the information treatment, an indicator of whether the participant is classified as *reluctant* or other (which combines *selfish* and *generous* types), and the interaction of the two, with additional controls for the participant's age, gender, and pre-donation earnings. In parentheses are bootstrap standard errors from 1000 repetitions for the OLS regressions on *Donation size*, and standard errors clustered at the experimental session for the probit regressions on *Prob. of donating more than $1*. *p<0.1, **p<0.05, ***p<0.01.

sensitive to information than other participants, as the changes in giving tended to be larger in magnitude, and more often statistically significant, for this group than for other participants. Before sorting (namely, not conditioning on intent to donate), the expected revenue to the charity and the probability of receiving a donation larger than $1 increased for *reluctant* participants in all information treatments, relative to the *no-information* treatment (and significantly so in all but one case). On the other hand, other participants tended not to respond to any information treatment.

We get a clearer picture of the effects at the intensive margin of giving by looking at the results conditional on intent to donate. There we see, again, that it is the *reluctant* participants that mostly drove the treatment effects. Their giving increased significantly with all information treatments,

while giving intensity by other participants tended not to change. It is particularly noteworthy that, relative to the *no-information* treatment, *reluctant* participants increased their donation size and their likelihood of donating more than $1 when they received information about a $5 donation in stage 2—by 172 percent and 962 percent, respectively—even though the same information induced them to sort out when the information was shown in stage 1. Thus, even though *reluctant* participants responded positively to the social information at the intensive margin, and ended up making fairly large donations, it appears they would have preferred to sort out in the first place. Other participant types did not respond in this way. This correlation suggests that the mechanism driving the response to social information in our experiment may be related to the mechanism driving exit in the dictator game.

3.6. Social pressure avoidance and empathy avoidance

We interpret the results as suggestive of a combination of social pressure avoidance and motivated reasoning: Individuals dislike giving, and seek excuses to opt out of the solicitation without compromising their moral image. Social information increases the pressure to give, but at the same time can provide (or invalidate) excuses to opt out, depending on the content of the information.[15] Although consistent with this mechanism, our results do not demonstrate it unambiguously. It may be that reluctant giving in our experiment is associated with *empathy avoidance* (Shaw, Batson, and Todd, 1994), or what Andreoni, Rao, and Trachtman (2016) call *avoidance of empathic stimulation* to describe why individuals give when asked, but at the same time avoid the solicitor. According to this view, individuals get joy from giving, and the solicitation triggers in them a

[15] In a related decision-theoretic model, Yagasaki (2018) proposes a combination of pride and shame to explain these results.

feeling of empathy and an impulse to give that is hard to resist. Individuals that are aware of their "empathic vulnerability" avoid the ask as a way to self-regulate (Andreoni, Rao, and Trachtman, 2016). Interestingly, the two models described above take almost opposite views of the underlying psychology. In one, givers dislike giving and feel pressured to do so, but engage in self-deceit by rationalizing a way out. In the other, givers enjoy giving and feel overstimulated by the ask, but their acute self-awareness allows them to engage in avoidance as a self-control mechanism. Both models predict that individuals give when asked and at the same avoid being asked.[16] One strategy that might help to distinguish between these motivations may be to investigate what personality characteristics are associated with reluctant giving. Of the Big Five personality traits (John, Donahue, and Kentle, 1991), research shows that Neuroticism is associated with increased anxiety and rumination (Muris et al., 2005), with experiencing the world as threatening (Baumeister and Vohs, 2007), and with negative effects on decision-making from situations involving peer pressure and social evaluation (Byrne, Silasi-Mansat, and Worthy, 2015). On the other hand, Agreeableness is associated with dispositional empathy and greater tendency to offer help, even to strangers (Baumeister and Vohs, 2007). Thus, based on this research, we hypothesize that participants in our experiment will score relatively higher than average on Neuroticism if their behavior is driven by social pressure avoidance and rationalization, while they will score relatively higher than average on Agreeableness if their behavior is driven by empathy avoidance.

[16] It is unclear, however, what prediction empathy avoidance makes with respect to the effect of social information on giving. Is the feeling of empathy and the tendency to avoid the empathic stimulus dependent on the giving of others? Also, while empathy avoidance can account for giving and at the same time avoiding the ask, it is less clear how it relates to exiting *after* a decision to share money (as in the dictator game in our experiment). If giving is utility-enhancing ex post being asked to give, then, if one already failed to avert the empathic stimulus of the ask in the dictator game, why retract one's gift (since doing so leads to a decrease in utility and to guilt)?

Table 5: Personality and principle of care as predictors of giving type

	Reluctant		Generous	
	Big 5	Big 5 + Princ. care	Big 5	Big 5 + Princ. care
Neuroticism	0.012**	0.010**	0.001	0.001
	(0.005)	(0.005)	(0.003)	(0.003)
Agreeableness	0.011**	0.003	0.006**	0.005*
	(0.006)	(0.005)	(0.003)	(0.003)
Conscientiousness	-0.002	-0.003	-0.001	-0.001
	(0.006)	(0.005)	(0.004)	(0.004)
Extraversion	-0.005	-0.005	-0.004***	-0.004***
	(0.005)	(0.005)	(0.001)	(0.001)
Openness	0.007	0.003	0.001	0.001
	(0.006)	(0.007)	(0.002)	(0.002)
Principle care 1	-	0.076***	-	-0.011
		(0.026)		(0.018)
Principle care 2	-	-0.038	-	-0.035*
		(0.023)		(0.020)
Principle care 3	-	0.041	-	-0.008
		(0.038)		(0.021)
Female	0.133***	0.123***	-0.048	-0.048
	(0.037)	(0.042)	(0.036)	(0.035)
Age	0.012	0.013	-0.012	-0.011
	(0.011)	(0.010)	(0.011)	(0.011)
Earnings	-0.015	0.000	-0.082	-0.079
	(0.079)	(0.074)	(0.062)	(0.062)
Pseudo R squared	0.048	0.085	0.055	0.080
N	308	308	308	308

Notes: Marginal effects (in percentage points) from probit regressions on the probability of being classified as *reluctant* vs. all other types (for Reluctant columns), and classified as *generous* versus all other types (for Generous columns). The Big 5 specification includes as regressors the participant's score in each of the big five personality dimensions, and the Big 5 + Princ. care specification includes in addition intensity of agreement with each of the three principle of care statements. Earnings refer to pre-donation earnings. Standard errors clustered at the experimental session in parentheses. *p<0.1, **p<0.05, ***p<0.01.

Table 5 examines this hypothesis. It presents results from probit regressions that estimate the probability that the participant is classified as *reluctant* versus all other types (in the Reluctant columns), and classified as *generous* versus all other types (in the Generous columns). The "Big 5" specification includes as predictors the participant's score for each of the Big Five personality

dimensions, based from their responses to the questionnaire at the end of the experiment (we control in addition for gender, age, and pre-donation earnings). We find that *reluctant* participants scored significantly higher than other types on both Neuroticism and Agreeableness, while *generous* participants scored significantly higher than other types on Agreeableness (and also lower on Extraversion) but not Neuroticism. These results give preliminary support for both social pressure and empathy as potentially important determinants of reluctant giving in our experiment. Giving that is not retracted, on the other hand, appears to be motivated by empathy but not social pressure.

However, we probe this result further by looking at the potential mediating role of the *principle of care* in the observed relation between empathy and reluctant giving. Ottoni-Wilhelm and Bekkers (2010) and Bekkers and Ottoni-Wilhelm (2016) show in a series of studies that helping in response to needs of an abstract recipient (such as a charity or an anonymous participant) is more strongly associated with an internalization of a moral value that one should help those in need—a value known as the *principle of care* (Hoffman, 2000; Batson, 2011)—than with an empathic reaction toward the recipient. Indeed, Ottoni-Wilhelm and Bekkers (2010) and Bekkers and Ottoni-Wilhelm (2016) show that agreement with the principle of care mediates the empathy–giving relationship in cases of abstract helping to people in need, and thus empathy loses its explanatory power when agreement with the principle of care is accounted for. Building on this work, in our experiment we elicited agreement with the principle of care using the 3-question instrument in Ottoni-Wilhelm and Bekkers (2010). Table 5, columns "Big 5 + Proc. care" examine this mediating hypothesis by replicating the analysis in the "Big 5" specification, but now including in addition the participant's intensity of agreement with each of the three principle of care items (note

that agreement with principle 2 implies weaker internalization). We find that the inclusion of these variables has a significant explanatory power over the probability of being classified as *reluctant*. Moreover, in this augmented specification, the coefficient on Neuroticism is hardly changed and continues to be significant, while the coefficient on Agreebleness shrinks by 75 percent (from 0.011 to 0.003) and is no longer significant. On the other hand, inclusion of the principle of care variables also contributes significantly to explaining the probability of being classified as *generous*, but in this augmented specification, the coefficient on Agreeableness barely changes and continues to be significant. Thus, the potential role of Neuroticism in explaining reluctant giving is robust to the inclusion of the principle of care, while the role of Agreeableness is not. This gives stronger support to social pressure avoidance as a mechanism driving reluctant giving in our experiment. On the other hand, the role of Agreeableness in explaining giving by *generous* participants is robust to the inclusion of the principle of care. This suggests that empathy might be an important motivator of giving that is then not retracted in our experiment.[17]

4. Discussion

We investigate the effect of social information on the decisions to sort into an opportunity to donate to a charity and of how much to give. Our results demonstrate that altruism, conformity, and learning have difficulty explaining the observed patterns. We find that exit in a dictator game is

[17] These results replicate the findings in Ottoni-Wilhelm and Bekkers (2010) and Bekkers and Ottoni-Wilhelm (2016), while adding a potentially important distinction. We find that the principle of care mediates the empathy–giving link for givers that later retracted their giving (the majority of givers in our experiment). However, for the subset of givers that did not retract their giving, the empathy–giving link was not mediated by the principle of care, suggesting that some individuals might be moved to help due to empathy even in contexts of help to abstract recipients. Our results are also consistent with Lotz et al.'s (2013) finding that individuals that score low on *justice sensitivity*—a measure of the importance an individual places on justice in everyday life—tend to avoid giving and become selfish when circumstances give them an excuse to do so.

highly predictive of the responses to social information, both at the extensive and intensive margins of giving. In combination with an analysis of personality traits and values, the results are consistent with a model in which social information increases social pressure, and at the same time affects the individual's ability to find plausible justifications for sorting out of the solicitation.

The results reveal that social information can be used to influence sorting substantially, in both directions. By simply mentioning a relatively small or large previous donation, we were able to increase sorting by 65 percent, or decrease it by 38 percent, respectively (relative to providing no information), without changing the cost of giving. If information has this effect because it makes it more or less difficult for the individual to find an excuse not to give, one implication is that observed entry may not be taken as evidence that the solicitation is welfare-enhancing for the donor, because the donor might have felt pressured to sort in, just as we have come to recognize that observed giving may not be taken as evidence that the solicitation is welfare-enhancing for the donor, because the donor might have felt pressured to give. Thus, to assess the welfare effects of fundraising we may have to consider variables beyond sorting. And policies that seek to respect donor welfare, such as providing individuals an opportunity to sort or to opt out, may have to pay careful attention to how this opportunity is framed, as subtle factors (such as social information) could have important effects in the direction opposite to what is intended, as we demonstrate. Our findings also have practical implications to fundraisers, for the design of mechanisms in order to increase revenue. In particular, when thinking about providing social information, fundraisers may find it profitable to view content and timing of the information as two separate dimensions to be calibrated. The magnitude and direction of the effect of the information along these two dimensions may depend on whether solicitees are likely to feel social pressure from the fundraiser.

APPENDIX A

A model of reluctant giving and motivated reasoning

We present a model to describe how reluctant giving can arise as equilibrium behavior if donors are concerned about maintaining a moral image, and at the same time have a tendency to exploit the decision context in order to find excuses to decline the ask without appearing selfish.

A.1. The solicitation game

Consider a game between a donor, indexed by i, and an observer, denoted by s. The donor decides whether and how much to contribute in response to a solicitation, and the observer has no actions in the game. The donor's level of generosity, v_i, is a uniform random draw from the support $[0,2\bar{v}]$. The distribution over v_i is common knowledge to the players, but the value v_i is private information of the donor.

The solicitation occurs in two stages. In the first stage (or *invitation* stage), the donor decides whether to sort into the second stage. Denote this decision by $d_i \in \{0,1\}$, where $d_i = 0$ if the donor sorts out, and $d_i = 1$ if the donor sorts in. If $d_i = 0$, the solicitation ends there, with no contribution from the donor. If $d_i = 1$, the donor moves to the second stage (or *payment* stage). In the second stage, the donor decides an amount $g_i \geq 0$ to contribute. Note that giving zero is allowed. Modeling the decision to give zero separately from the decision to sort out will allow us to capture in the equilibrium a tendency of the donor to exploit flexibility in the decision context in order to find justifications to sort out, and to give a positive amount if such justifications are not available. We model flexibility in the decision context with a single parameter as follows. Let $p \in (0,1)$ be the probability that the donor is forced to stay out of the payment stage for reasons beyond her

control, conditional on having sorted in. This captures situations in which an individual fails to give despite her best intentions. For example, the individual decides to attend a charity banquet but fails to arrive on time because her car breaks down on the way to the event. Or the individual waits at home to meet a door-to-door solicitor but honestly fails to hear the doorbell. Crucially, the failure to participate in the payment stage is truly unintentional, and unrelated to the donor's level of generosity. Thus, we assume that p is common knowledge to the players and is a value independent of v_i. It bears stressing that with this we are not assuming that less generous donors are more likely to experience an event that forces them out of the payment stage. Rather, all donors are equally likely to experience such an event.

As the donor may decide to sort in but be forced out, we distinguish sorting and actual participation. Denote participation by $a_i \in \{0,1\}$, where $a_i = 0$ if the donor fails to participate in the payment stage, and $a_i = 1$ if the donor participates. Obviously, the donor never participates if she sorts out; thus, $d_i = 0$ implies $a_i = 0$. Similarly, the donor participates only if she sorts in, thus $a_i = 1$ implies $d_i = 1$. But the converse of neither statement is necessarily true, because of the probability p that $a_i = 0$ when $d_i = 1$.

Finally, the role of the observer in the game is simply to observe the outcome of the solicitation and make an inference about the donor's level of generosity v_i based on that outcome. We assume the observer sees the outcome of the game but not the donor's intentions. That is, he observes whether the donor participates and how much she donates (a_i and g_i), but does not observe the donor's sorting decision (d_i). Importantly, if the donor fails to participate in the payment stage, the observer does not observe whether that happened intentionally or unintentionally. We assume

the donor cares about the observer's inference, such that the observer's posterior belief about the donor's level of generosity v_i will enter directly into the donor's utility function. In this sense, the solicitation game is a psychological game between the donor and the observer (Geanakoplos, Pearce, and Stacchetti, 1989; Battigalli and Dufwenberg, 2009).

A.2. Preferences

Following Benabou and Tirole (2006), we assume that when the donor contributes an amount g_i, she obtains "intrinsic" utility equal to $g_i v_i$, and incurs in cost equal to $\frac{k{g_i}^2}{2}$, where k is fixed and common knowledge. The donor also obtains "image" utility resulting from the observer's posterior belief about her level of generosity. If the donor fails to participate in the payment stage, she obtains image utility equal to $\gamma E_s(v_i|a_i = 0)$, where $\gamma > 0$ is the donor's intensity of her image concern, and is fixed and common knowledge. $E_S(v_i|a_i = 0)$ is the observer's posterior belief about v_i conditional on observing that the donor did not participate in the payment stage. On the other hand, if the donor participates in the payment stage and makes a gift equal to g_i, the donor obtains image utility equal to $\gamma E_s(v_i|g_i)$ (short for $\gamma E_s(v_i|g_i, a_i = 1)$).[18] Note that if the donor contributes nothing—either because she did not participate or because she made a gift of $g_i = 0$—she obtains no intrinsic utility and incurs in no cost, but still obtains image utility from the observer's inference equal to $\gamma E_S(v_i|a_i = 0)$ or $\gamma E_S(v_i|g_i = 0)$. These values result in equilibrium. For completeness, the observer obtains constant utility c in the game.

[18] More strictly, the donor's image utility is not a function of $E_s(\cdot)$, but of her beliefs about $E_s(\cdot)$, since the observer's beliefs are not known to the donor. But in equilibrium beliefs are assumed to be correct, and so formulating the donor's preferences in terms of the observer's beliefs is without problem and avoids the more cumbersome notation of second-order beliefs (Geanakoplos, Pearce, and Stacchetti, 1989).

A.3. Equilibrium

Our goal with the model is to illustrate that reluctant giving can be supported as equilibrium behavior with the game and preferences described above. In this equilibrium, donors intentionally sort out, and use the fact that any donor might be prevented from participating by chance as a plausible justification for their lack of participation. Unable to access the donor's intentions, the observer will be convinced (to some extent) by this justification. However, when we remove the availability of this plausible justification from the game, a donor with the same preferences now donates a positive amount rather than 0. We solve for a Perfect Bayesian equilibrium with these characteristics.

Given the preferences, the donor's problem is

$$\max_{d_i\in\{0,1\},g_i\geq 0} d_i\left\{p\gamma E_s(v_i|a_i=0)+(1-p)\left[v_i g_i-\frac{k{g_i}^2}{2}+\gamma E_s(v_i|g_i)\right]\right\} \\ +(1-d_i)[\gamma E_s(v_i|a_i=0)] \tag{1}$$

Letting $R_s(g_i)\equiv E_s(v_i|g_i)$ and assuming that $R_s(g_i)$ is differentiable, it follows that in the payment stage, the donor chooses the amount ${g_i}^*(v_i)$ given by the first-order condition

$$v_i-k{g_i}^*(v_i)+\gamma\frac{dR_s\left({g_i}^*(v_i)\right)}{dg_i}=0 \tag{2}$$

The beliefs $R_s(g_i)$ will result endogenously in equilibrium, but are taken as given by the donor when solving her problem. If $R_s(g_i)$ is weakly increasing in g_i, equation (2) implies that the donor's level of generosity v_i is perfectly revealed from the size of her gift, since there is a unique optimum ${g_i}^*$ different for each v_i. It follows that, under rational expectations, beliefs in equilibrium pin v_i perfectly when ${g_i}^*(v_i)$ is observed; namely

$$R_s\left({g_i}^*(v_i)\right)={g_i}^{*-1}\left({g_i}^*(v_i)\right)=v_i \tag{3}$$

Substituting (3) into (2) results in the differential equation

$$R_s\left(g_i{}^*(v_i)\right) - k g_i{}^*(v_i) + \gamma \frac{dR_s\left(g_i{}^*(v_i)\right)}{dg_i} = 0 \tag{4}$$

The solution to this equation is the belief function that maps the observed donation size g_i to the posterior expected level of generosity v_i that the observer infers in equilibrium. This function is

$$R_s(g_i) = k\left[g_i - \gamma\left(1 - e^{-\frac{g_i}{\gamma}}\right)\right] \tag{5}$$

where the initial condition has been set to $R_s(0) = 0$. This initial condition states that the observer identifies the minimum possible donation amount ($g_i = 0$) as coming from a donor with the lowest possible level of generosity ($v_i = 0$). This condition also guarantees that $R_s(g_i)$ is increasing in g_i for any $k > 0$ and $\gamma > 0$.

The donor's optimal gift size in the payment stage, as a function of her level of generosity and given the belief function $R_s(g_i)$, is given by $g_i{}^*(v_i)$, which is found by taking the derivative of the belief function and substituting it into equation (2), obtaining

$$g_i{}^*(v_i) = \frac{v_i}{k} + \gamma\left[1 + \mathcal{W}_0\left(-e^{-1-\frac{v_i}{k\gamma}}\right)\right] \tag{6}$$

where $\mathcal{W}_0$ is the principal branch of the Lambert W function.[19] Note that $g_i{}^*(v_i)$ is increasing in v_i, and $g_i{}^*(0) = 0$. Figure A1 (left panel) illustrates the shape of the optimal donation size.

[19] The Lambert W function is defined as the solution to $x = \mathcal{W}(x)e^{\mathcal{W}(x)}$. Its principal branch, $\mathcal{W}_0(x)$, is real-valued and increasing for $x \geq -\frac{1}{e}$, with $\mathcal{W}_0\left(-\frac{1}{e}\right) = -1$, which implies for our problem that $g_i{}^*(v_i)$ is real-valued and increasing in v_i for any $v_i \geq 0$, given that $k > 0$ and $\gamma > 0$. For a discussion on the Lambert W function, see Corless et al. (1996).

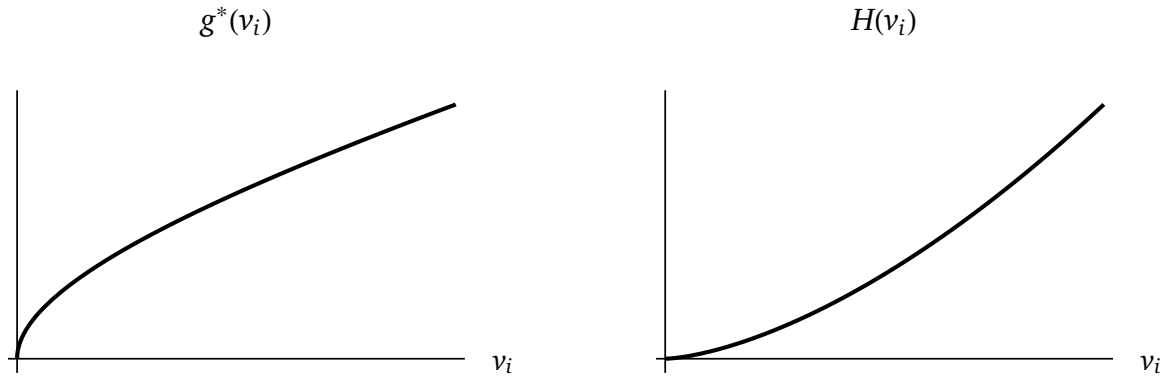

Figure A1. Shape of $g_i{}^*(v_i)$ and $H(v_i)$ for arbitrary values of k and γ

Given $R_s(g_i)$ and $g_i{}^*(v_i)$, the utility the donor expects to obtain if she decides to participate in the payment stage is given by the value function

$$U^*(v_i) \equiv p\gamma E_s(v_i|a_i = 0) + (1-p)H(v_i) \tag{7}$$

where $H(v_i) \equiv \gamma v_i + \frac{1}{2k}(v_i{}^2 - k^2\gamma^2\Psi^2(v_i))$ and $\Psi(v_i) \equiv 1 + \mathcal{W}_0\left(-e^{-1-\frac{v_i}{k\gamma}}\right)$. Note that $H(v_i)$ is increasing in v_i for any $k > 0$ and $\gamma > 0$, is continuous, and $H(0) = 0$. Figure A1 (right panel) illustrates the shape of $H(v_i)$.

The value of $E_s(v_i|a_i = 0)$, which remains to be solved for, depends on the sorting behavior in equilibrium. First, note that a situation in which all donors sort in cannot be an equilibrium. If all donors sorted in, failure to participate would happen only unintentionally, and therefore the observer's beliefs conditional on observing non-participation would be uniform over $[0,2\bar{v}]$, resulting in $E_s(v_i|a_i = 0) = \bar{v}$. But then a donor with low-enough level of generosity would prefer to sort out and be mistaken for the average type than to sort in and potentially reveal her low level

of generosity.[20] Thus, in equilibrium, the observer's beliefs $E_s(v_i|a_i = 0)$ have to incorporate the fact that some donors intentionally sort out, and sorting behavior has to correspond to these beliefs.

Consider, then, an equilibrium in which there exists a level of generosity v_c, $0 < v_c < 2\bar{v}$, such that donors with $v_i < v_c$ sort out, and donors with $v_i > v_c$ sort in and make a gift of ${g_i}^*(v_i)$. In this case, any donation amount $g_i \in [{g_i}^*(v_c), {g_i}^*(2\bar{v})]$ is observed in the payment stage in equilibrium with probability $\frac{1-p}{2\bar{v}-v_c}$. The observer's beliefs are such that, if the observer sees $g_i \in [{g_i}^*(v_c), {g_i}^*(2\bar{v})]$, he places full weight on the donor being of generosity level equal to $R_s(g_i)$. Failure to participate in the payment stage happens either unintentionally for any donor, or intentionally for donors with $v_i < v_c$. Thus, if the observer sees that the donor fails to participate in the payment stage, his beliefs about the donor's level of generosity are uniform over $[0,2\bar{v}]$ with probability p, and uniform over $[0, v_c]$ with probability $1 - p$. The expected value of this mixture is $E_S(v_i|a_i = 0) = p\bar{v} + (1-p)\frac{v_c}{2}$. Gift amounts $\widetilde{g_i} \in [0, {g_i}^*(v_c))$ are never observed in equilibrium, and therefore beliefs upon observing such amounts must be specified outside of Bayes' Rule. We assume that if the observer sees $\widetilde{g_i} \in [0, {g_i}^*(v_c))$, he places full weight on the donor being of generosity level $R_S(\widetilde{g_i})$.[21] In the Appendix B, we derive conditions that guarantee existence of this equilibrium (Proposition B.1) and that it is "reasonable" in the sense that it survives Cho and Kreps' (1987) equilibrium domination test (Proposition B.2).

[20] To see this, take for instance the least generous donor (i.e., $v_i = 0$). If she decides to sort in she would give ${g_i}^*(0) = 0$, and her expected payoff is $p\gamma\bar{v} + (1-p)0 = p\gamma\bar{v}$. If she instead decides to sort out, her expected payoff is $\gamma\bar{v}$, which is strictly greater than $p\gamma\bar{v}$ for any valid p. Therefore, she prefers to sort out.

[21] The cutoff type v_c is a function of p. For ease of exposition, at the moment we avoid the notation $v_c(p)$. But later we take this notation explicitly when we discuss how changes in the "amount" of justifications for failing to participate available in the game (captured by p) affects sorting behavior.

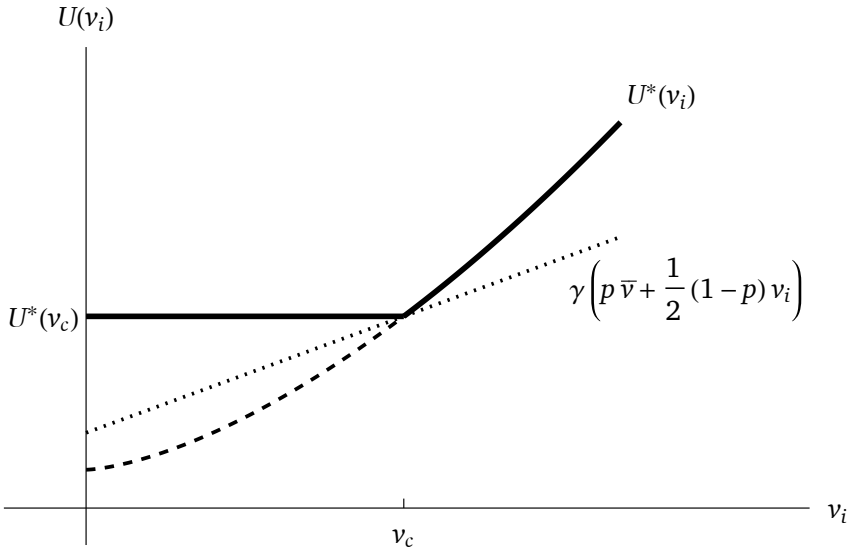

Figure A2. Semi-separating equilibrium in the two-stage game

Notes: The solid line traces the expected payoff to the donor in equilibrium as a function of the donor type. Donor types below v_c pool at sorting out intentionally, and donor types above v_c sort in and given their optimum gifts.

Figure A2 illustrates the behavior of donors in equilibrium. Donor types below v_c sort out intentionally, while donor types above v_c sort in and give their optimal amounts, unless they are unintentionally forced out of the payment stage. Upon observing a gift, the observer correctly identifies the type. Upon observing lack of participation, the observer correctly accounts for intentional and unintentional lack of participation, and thus he believes that a donor that does not participate is in expectation of type $p\bar{v} + (1-p)\frac{v_c}{2}$. Donor types below v_c strictly prefer to sort out and be mistaken by this type than to sort in and potentially reveal their type. Donor types above v_c strictly prefer to sort in and make their optimal gift than to sort out. The cutoff type is indifferent between the two options.

A.4. Identifying reluctant giving

To describe reluctant giving within our framework, we consider now a game in which donors are unable to sort out of the payment stage. We continue to assume that preferences are as described initially, but now suppose the game consists only of the payment stage. In this case, there is no sorting decision to be made and no event that prevents participation. However, the donor is still allowed to donate zero, and therefore the new game is no more restrictive in terms of the monetary options available to the donor. The donor must choose an amount $g_i \geq 0$ to contribute, and the observer infers the donor's level of generosity v_i based on g_i. The donor's problem is

$$\max_{g_i \geq 0} \left[v_i g_i - \frac{k {g_i}^2}{2} + \gamma E_S(v_i | g_i) \right] \tag{9}$$

One can show that there is now a fully-separating equilibrium in which each donor contributes her unique and distinct optimal amount ${g_i}^*(v_i)$ as defined previously. More interestingly, there is also a semi-separating equilibrium with similar characteristics to the equilibrium in the two-stage game. Donor types below a cutoff v_c' pool at $g_i = 0$, and donor types above v_c' give their optimal gift ${g_i}^*(v_i)$, where v_c' is such that $H(v_c') = \gamma \frac{v_c'}{2}$. The observer cannot distinguish the donors that pool, and perfectly identifies the donors that contribute. His beliefs are $E_S(v_i | g_i) = \frac{v_c'}{2}$ upon observing $g_i = 0$, and $E_S(v_i | g_i) = R_S(g_i)$ upon observing $g_i \in ({g_i}^*(v_c'), {g_i}^*(2\bar{v})]$. Gift amounts $\widetilde{g_i} \in \left(0, {g_i}^*(v_c')\right)$ are never observed in equilibrium. We assume that upon observing such an amount, the observer places full weight on the donor being of type $R_S(\widetilde{g_i})$.

This equilibrium is almost identical to the one in the two-stage game, with the difference that some donors now pool at $g_i = 0$ rather than at sorting out intentionally. Notice that a situation in which

all donors sort in and then partially pool at $g_i = 0$ is not an equilibrium in the two-stage game.[22] What leads to this difference between the two games, and gives bite to the availability of sorting, is the existence of the probability p. The fact that it is plausible that donors unintentionally missed the payment stage forces the observer to interpret non-participation more generously than he interprets giving $g_i = 0$ in the payment stage. As long as $p > 0$, donors have "an excuse" that makes sorting out preferable to giving $g_i = 0$, because it yields greater image utility. Thus, with $p > 0$, the set of donors that pool at sorting out in the two-stage game contains all donors that pool at $g_i = 0$ in the payment-stage-only game, plus an additional set of donors.[23] This additional group of donors might be thought of as the reluctant givers. They sort out if sorting is possible, but donate a strictly positive amount (and do not pool at $g_i = 0$) when the payment stage cannot be avoided.

Figure A3 illustrates this result. When sorting is not possible, donors below v_c' pool at $g_i = 0$ and obtain a payoff of $\gamma E_s(v_i|g_i = 0)$, whereas when sorting is possible, a larger set of donors (those below v_c) sort out and obtain a larger payoff of $\gamma E_s(v_i|a_i = 0)$. Donors between v_c' and v_c are the reluctant givers. They sort out if possible, but otherwise give. The least generous donors (those below v_c') give in neither case, and the most generous donors (those above v_c) give in both cases. Only the reluctant donors change their behavior in response to the availability of sorting.

[22] When all donors sort in, failure to participate happens only unintentionally, thus the observer's beliefs upon observing non-participation are $E_s(v_i|a_i = 0) = \bar{v}$. Donors that sort in and pool at $g_i = 0$ obtain expected payoff equal to $p\gamma\bar{v} + (1-p)\gamma\frac{v_c'}{2}$, but sorting out would yield a larger payoff of $\gamma\bar{v}$.

[23] Proposition B.3 in the Appendix B proves this by finding conditions that guarantee that $v_c(p)$ is increasing in p and that can always be satisfied while satisfying conditions for Propositions B.1 and B.2.

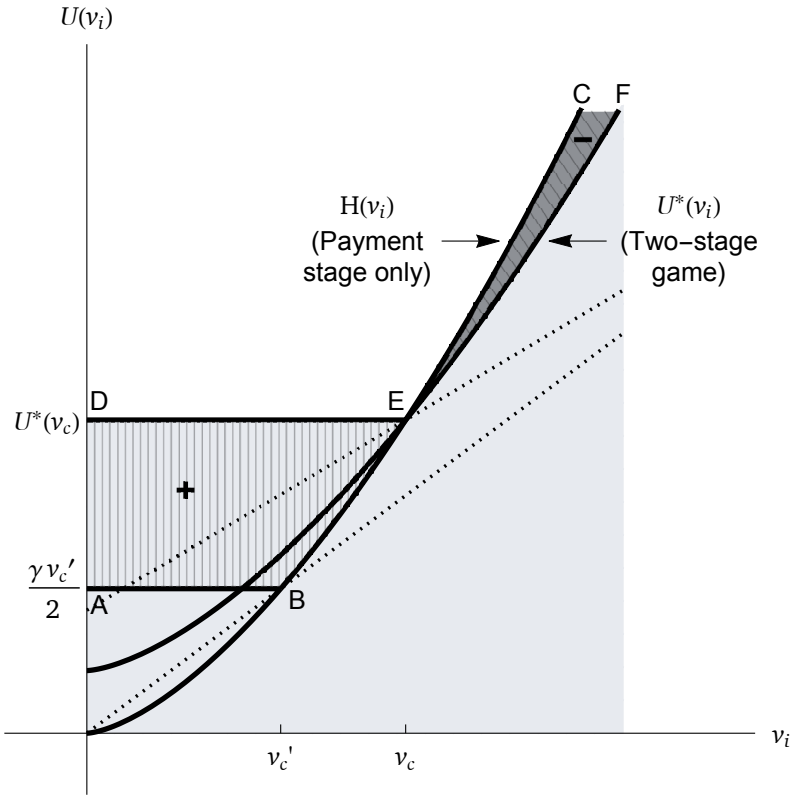

Figure A3. Effect of availability of sorting on equilibrium behavior

Notes: The graph plots the expected payoff to the donor in equilibrium as a function of the donor type. The line ABC traces the payoff in the two-stage game (sorting available), and the line DEF traces the payoff in the payment-stage-only game (sorting not available). Donor types between v_c' and v_c are the reluctant givers: they pool at sorting out in the two-stage game, but give positive amounts in the payment-stage-only game. The area ABED is the gain in utility from the availability of sorting, resulting from the possibility that non-participation is interpreted by the observer as unintentional. The area EFC is the loss in utility from the availability of sorting, resulting from the possibility that donors that intended to donate experience an unexpected event that prevents them from participating.

A.5. Motivated reasoning and social information

In the model, the donor and the observer are different agents. But one could take a dual-self interpretation and view them as different sides of the same individual. In this interpretation, the individual lacks complete awareness of what motivates her non-participation, and has a tendency to find reasons to interpret her non-participation self-servingly. The parameter p would represent the individual's ability to make self-serving interpretations, which may depend on the decision context. One important contextual element might be information about what peers are doing. Thus,

by manipulating social information, the fundraiser may be able to affect sorting in either direction, by facilitating or impeding the availability of excuses to sort out.

REFERENCES

Allcott, H., 2011. Social norms and energy consumption. *Journal of Public Economics*, 95.9: 1082–1095.

Alpizar, F., Carlsson, F., and Johansson-Stenman, O., 2008. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. *Journal of Public Economics*, 92.5: 1047–1060.

Andreoni, J., and Bernheim, B.D., 2009. Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77.5: 1607–1636.

Andreoni, J., and Payne, A.A., 2013. Charitable Giving. In *Handbook of Public Economics*, vol. 5, eds. Auerbach, A.J., Chetty, R., Feldstein, M., and Saez, E., North Holland.

Andreoni, J., Rao, J.M., and Trachtman, H., 2017. Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy*, 125.3: 625–653.

Andreoni, J., and Scholz, K., 1998. An econometric analysis of charitable giving with interdependent preferences. *Economic Inquiry*, 36.3: 410–428.

Argo, N., Klinowski, D., Krishnamurti, T., and Smith, S., 2019. The completion effect in charitable crowdfunding. *Working paper.*

Bardsley, N., 2008. Dictator game giving: Altruism or artefact? *Experimental Economics*, 11.2: 122–133.

Battigalli, P., and Dufwenberg, M., 2007. Guilt in games. *American Economic Review*, 97.2: 170–176.

Batson, C.D., 2011. *Altruism in Humans*. Oxford University Press.

Baumeister, R.F., and Vohs, K.D., 2007. *Encyclopedia of Social Psychology*. Sage Publications.

Becker, G., 1974. A theory of social interactions. *Journal of Political Economy*, 82.6: 1063–1093.

Bekkers, R., and Ottoni-Wilhelm, M., 2016. Principle of care and giving to help people in need. *European Journal of Personality*, 30: 240–257.

Bénabou, R., and Tirole, J., 2006. Incentives and prosocial behavior. *American Economic Review*, 96.5: 1652–1678.

Bénabou, R., Falk, A., and Tirole, J., 2018a. Eliciting moral preferences. *Working paper*.

Bénabou, R., Falk, A., and Tirole, J., 2018b. Narratives, imperatives, and moral reasoning. *NBER Working Paper 24798*.

Bernheim, D.B., 1994. A theory of conformity. *Journal of Political Economy*, 102.5: 841–877.

Bergstrom, T., Blume, L., and Varian, H.R., 1986. On the private provision of public goods. *Journal of Public Economics*, 29: 25–49.

Bodner, R., and Prelec, D., 2003. Self-signaling and diagnostic utility in everyday decision making. In Brocas, I., and Carrillo, J.D., eds., *The Psychology of Economic Decisions*, vol. 1. Oxford University Press.

Bohnet, I., and Zeckhauser, R., 2004. Social comparisons in ultimatum bargaining. *Scandinavian Journal of Economics*, 106.3: 495–510.

Broberg, T., Ellingsen, T., and Johannesson, M., 2007. Is generosity involuntary? *Economics Letters*, 94.1: 32–37.

Byrne, K.A., Silasi-Mansat, C.D., and Worthy, D.A., 2015. Who chokes under pressure? The Big Five personality traits and decision-making under pressure. *Personality and Individual Differences*, 74: 22–28.

Camerer, C., 2003. *Behavioral Game Theory*. Princeton University Press.

Cason, T. N., and Mui, V.-L., 1998. Social influence in the sequential dictator game. *Journal of Mathematical Psychology*, 42: 248–265.

Castillo, M., Petrie, R., and Wardell, C., 2014. Fundraising through online social networks: A field experiment on peer-to-peer solicitation. *Journal of Public Economics*, 114: 29–35.

Castillo, M., Petrie, R., and Wardell, C., 2017. Friends asking friends for charity: The importance of gifts and audiences. *Working paper.*

Cho, I.K., and Kreps, D.M., 1987. Signaling games and stable equilibria. *Quarterly Journal of Economics*, 102.2: 179–221.

Cialdini, R.B., and Schroeder, D.A., 1976. Increasing compliance by legitimizing paltry contributions: When even a penny helps. *Journal of Personality and Social Psychology*, 34.4: 599–604.

Coffman, L.C., Featherstone, C.R., and Kessler, J.B., 2017. Can social information affect what job you choose and keep? *American Economic Journal: Applied Economics*, 9.1: 96–117.

Corless, R.M., Gonnet, G.H., Hare, D.E., Jeffrey, D.J., and Knuth, D.E., 1996. On the Lambert W function. *Advances in Computational Mathematics*, 5.1: 329–359.

Dana, J., Cain, D.M., and Dawes, R.M., 2006. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100.2: 193–201.

Dariel, A., and Nikiforakis, N., 2014. Cooperators and reciprocators: A within-subject analysis of prosocial behavior. *Economics Letters*, 122.2: 163–166.

DellaVigna, S., List, J.A., and Malmendier, U., 2012. Testing altruism and social pressure in charitable giving. *Quarterly Journal of Economics*, 127.1: 1–56.

De Quidt, J., Vesterlund, L., and Wilson, A.J., 2019. Experimenter demand effects. In *The Handbook of Research Methods and Applications in Experimental Economics*, eds. Schram, A., and Ule, A., Edward Elgar Publishing.

Eckel, C.C., and Grossman, P.J., 2003. Rebate versus matching: Does how we subsidize charitable contributions matter? *Journal of Public Economics*, 87.3: 681–701.

Ellignsen, T., and Johanesson, M., 2008. Price and prejudice: The human side of incentive theory. *American Economic Review*, 98.3: 990–1008.

Exley, C.L., 2016. Excusing selfishness in charitable giving: The role of risk. *Review of Economic Studies*, 83.2: 587–628.

Exley, C.L., and Kessler, J.B., 2018. Motivated errors. *Working paper*.

Exley, C.L., and Petrie, R., 2018. The impact of a surprise donation ask. *Journal of Public Economics*, 158: 152–167.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10.2: 171–178.

Fischbacher, U., Gächter, S., and Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71.3: 397–404.

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.

Geanakoplos, J., Pearce, D., and Stacchetti, E., 1989. Psychological games and sequential rationality. *Games and Economic Behavior*, 1.1: 60–79.

Gill, D., and Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102.1: 469–503.

Gino, F., Norton, M.I., and Weber, R.A., 2016. Motivated Bayesians: Feeling moral while acting egoistically. *Journal of Economic Perspectives*, 30.3: 189–212.

Grossman, Z., and Van der Weele, J.J., 2017. Self-image and willful ignorance in social decisions. *Journal of the European Economic Association*, 15.1: 173–217.

Hallsworth, M., List, J.A., Metcalfe, R.D., and Vlaev, I., 2017. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148: 14–31.

Haisley, E.C., and Weber, R.A., 2010. Self-serving interpretations of ambiguity in other-regarding behavior. *Games and Economic Behavior*, 68.2: 614–625.

Hoffman, M.L., 2000. *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge University Press.

John, O.P., Donahue, E.M., and Kentle, R.L., 1991. *The Big Five inventory—Versions 4a and 54*. University of California, Berkeley, Institute of Personality and Social Research.

Kamdar, A., Levitt, S.D., List, J.A., Mullaney, B., and Syverson, C., 2015. Once and done: Leveraging behavioral economics to increase charitable contributions. *Working paper.*

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.

Keser, C., and van Winden, F., 2000. Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, 102.1: 23–39.

Kessler, J.B., 2017. Announcements of support and public good provision. *American Economic Review*, 107.12: 3760–3787.

Klinowski, D., 2018. Gender differences in giving in the dictator game: The role of reluctant altruism. *Journal of the Economic Science Association*, 4.2: 110–122.

Krupka, E.L., and Croson, R.T.A., 2016. The differential impact of social norms cues on charitable contributions. *Journal of Economic Behavior & Organization*, 128: 149–158.

Kunda, Z., 1990. The case for motivated reasoning. *Psychological Bulleting*, 108.3: 480–498.

Lazear, E.P., Malmendier, U., and Weber, R.A., 2012. Sorting in experiments with application to social preferences. *American Economic Journal: Applied Economics*, 4.1: 136–163.

Linardi, S., and McConnell, M.A., 2011. No excuses for good behavior: Volunteering and the social environment. *Journal of Public Economics*, 95: 445–454.

List, J.A., 2007. On the interpretation of giving in dictator games. *Journal of Political Economy*, 115.3: 482–493.

List, J.A., and Lucking-Reiley, D., 2002. The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy*, 110.1: 215–233.

Lotz, S., Schlösser, T., Cain, D.M., and Fetchenhauer, D., 2013. The (in)stability of social preferences: Using justice sensitivity to predict when altruism collapses. *Journal of Economic Behavior & Organization*, 93: 141–148.

Martin, R., and Randal, J., 2008. How is donation behavior affected by the donations of others? *Journal of Economic Behavior & Organization*, 67.1: 228–238.

Meer, J., 2011. Brother, can you spare a dime: Peer pressure in charitable solicitation. *Journal of Public Economics*, 95.7: 926–941.

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.

Muris, P., Roelofs, J., Rassin, E., Franken, I., and Mayer, B., 2005. Mediating effects of rumination and worry o the links between neuroticism, anxiety and depression. *Personality and Individual Differences*, 39: 1105–1111.

Ottoni-Wilhelm, M., and Bekkers, R., 2010. Helping behavior, dispositional empathic concern, and the principle of care. *Social Psychology Quarterly*, 73.1: 11–32.

Potters, J., Sefton, M., and Vesterlund, L., 2007. Leading-by-example and signaling in voluntary contribution games: An experimental study. *Economic Theory*, 33.1: 169–182.

Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review*, 83.5: 1281–1302.

Romano, R., and Yildirim, H., 2001. Why charities announce donations: A positive perspective. *Journal of Public Economics*, 81.3: 423–447.

Sachdeva, S., Iliev, R., and Medin, D.L., 2009. Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, 20.4: 523–528.

Samek, A.S., and Sheremeta, R.M., 2014. Recognizing contributors: An experiment on public goods. *Experimental Economics*, 17.4: 673–690.

Shang, J., and Croson, R., 2009. A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *Economic Journal*, 119.540: 1422–1439.

Shaw, L.L., Batson, C.D., and Todd, R.M., 1994. Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences. *Journal of Personality and Social Psychology*, 67.5: 879–887.

Smith, S., Windmeijer, F., and Wright, E., 2015. Peer effects in charitable giving: Evidence from the (running) field. *Economic Journal,* 125.585: 1053-1171.

Sugden, R., 1984. Reciprocity: The supply of public goods through voluntary contributions. *Economic Journal*, 94.376: 772–787.

Trachtman, H., Steinkruger, A., Wood, M., Andreoni, J., Murphy, J.J., and Rao, J.M., 2015. Fair weather avoidance: Unpacking the costs and benefits of "avoiding the ask". *Journal of the Economic Science Association*, 1.1: 8–14.

Varian, H.R., 1994. Sequential contributions to public goods. *Journal of Public Economics*, 53.2: 165–186.

Vesterlund, L., 2003. The informational value of sequential fundraising. *Journal of Public Economics,* 87.3: 627–657.

Vesterlund, L., 2006. Why do people give? In *The Nonprofit Sector: A Research Handbook,* Second Edition, eds. Steinberg, R., and Powell, E., Yale University Press.

Vesterlund, L., 2016. Using experimental methods to understand why and how we give to charity. In *The Handbook of Experimental Economics*, vol. 2, eds. Kagel, J.H., and Roth A.E., Princeton University Press.

Yagasaki, M., 2018. Pride, shame, and social comparisons. *SSRN Working Paper 3414146*.

Zhong, C., and Liljenquist, K., 2006. Washing away your sins: Threatened morality and physical cleansing. *Science*, 313.5792: 1451–1452.