Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially[*]

Dan Ariely[†], Anat Bracha[‡] and Stephan Meier [§]

August 6, 2007

Abstract

This paper experimentally examines image motivation—the desire to be liked and well-regarded by others—as a driver in prosocial behavior (doing good), and asks whether extrinsic monetary incentives (doing well) have a detrimental effect on prosocial behavior due to crowding out of image motivation.

By definition, image depends on one's behavior being visible to other people. Using this unique property we show that image is indeed an important part of the motivation to behave prosocially. Moreover, we show that extrinsic incentives interact with image motivation and are therefore less effective in public than in private. Together, these results imply that image motivation is crowded out by monetary incentives; which in turn means that monetary incentives are more likely to be counterproductive for public prosocial activities than for private ones.

Keywords: prosocial behavior, extrinsic incentives, image motivation, experiments

JEL: D64, C90, H41

1 Introduction

Most charitable organizations depend on private contributions, in the form of monetary gifts, volunteer efforts, or other tangible contributions, such as blood donations. The magnitude of private contributions is impressive—in the United States alone 89 percent of households donate, averaging $1,620 per year, and 44 percent of US adults volunteer the equivalent of 9 million full-time jobs (Independent Sector, 2001). This prosocial behavior is striking in light of the economic incentive to free-ride in the provision of public goods. In order to elicit contributions, charitable organizations use many creative efforts to incentivize voluntary giving. For example, the Lance Armstrong Foundation recently launched the LIVESTRONG campaign where individuals contribute by purchasing colorful wrist bands, many organizations give thank-you gifts, such as mugs and calendars, and charities organize walks, concerts, and advertise lists of their donors. Apart from the charitable organizations' efforts to raise donations, the government helps increase charitable giving by offering tax breaks for such donations.

The various types of charitable contributions and the many real-life ways of soliciting such donations suggest that there may be different motives for individuals to behave prosocially. These motives are roughly divisible into three broad categories: intrinsic, extrinsic, and image motivation. *Intrinsic motivation* is the value of giving per se, represented by private preferences for others' well-being, such as pure altruism or other forms of prosocial preferences (for surveys see Fehr and Schmidt, 2003, and Meier, 2007). *Extrinsic motivation* is any material reward or benefit, either monetary or non-monetary, associated with giving, such as thank-you gestures and tax breaks. In persuading potential donors to contribute, the Metropolitan Opera, for example, gives "many wonderful benefits that include: a subscription to OPERA NEWS, member discounts on merchandise, lectures and backstage tours, ticket priorities and much, much more!"[1] (all depending on the contribution amount). *Image motivation*, or signaling motivation, refers to individuals' tendency to be motivated partly

---

[1]http://www.metoperafamily.org/metopera/support/membership_patron/guild_membership/

by how others perceive them. Image motivation therefore captures the rule of opinion in utility, i.e., the desire to be liked and respected by others and by the self. If individuals are looking for social approval in their behavior, then they should try to signal traits which are defined as "good" based on social norms and values (for economic models incorporating social approval, see e.g. George A. Akerlof, 1980, Tore Ellingsen and Magnus Johannesson, 2007, and H. Hollaender, 1990). Being altruistic is often seen as "good", and being greedy or selfish is not. Prosocial behavior is therefore a way to signal to others that one is "good".[2]

The desire for social approval implies that, conditional on prosocial activity yielding a positive image, people will act more generously and prosocially in public than in private settings. Indeed, a number of field and laboratory studies have found such a pattern (e.g., James Andreoni and Ragan Petrie, 2004, Jason Dana et al., 2006, Mari Rege and Kjetil Telle, 2004, Adriaan R. Soetevent, 2005). This may explain why many organizations make individuals' contributions explicitly visible to others by having charity walks, offering wrist bands, posting lists of donors, etc.

While these three motivations (intrinsic, extrinsic, and image) have separate effects on prosocial behavior, they may also interact with each other. Following Titmuss (1970), who claimed that private monetary incentives for blood donations would decrease blood supply, a number of empirical studies have shown that private monetary incentives indeed have negative effects on prosocial behavior (e.g., Bruno S. Frey and Reto Jegen, 2001, Bruno S. Frey and Felix Oberholzer-Gee, 1997, Uri Gneezy and Aldo Rustichini, 2000a,b, Carl Mellstroem and Magnus Johannesson, 2005). While there are a few studies showing detrimental effects of extrinsic incentives, less is known about the mechanisms by which this occurs. A number of explanations have been offered: extrinsic motivations might interact with the intrinsic motivations in the decision to undertake a task (e.g., Edward L. Deci, 1975, Bruno S. Frey, 1997, Bruno S. Frey and Lorenz Goette, 1999); extrinsic incentives might destroy trust in a

---

[2]Behaving prosocially can also be undertaken instrumentally in order to signal one's wealth or status or in order to get future extrinsic rewards, like a political career or admission to a college (e.g., Amihai Glazer and Kai A. Konrad, 1996, William T. Harbaugh, 1998a, b, Blumkin Tomer and Sadka Efraim, 2007).

principal-agent relationship, leading the agent to behave less prosocially (e.g., Armin Falk and Michael Kosfeld, 2006, Ernst Fehr and Armin Falk, 2002, Ernst Fehr and John A. List, 2004); or the introduction of extrinsic motives might shift individuals' decision frame from a social frame, in which prosocial behavior is more prevalent, to a monetary frame in which maximization of self-interest becomes more salient (Uri Gneezy and Aldo Rustichini, 2000a, James Heyman and Dan Ariely, 2004).

This paper experimentally tests another mechanism by which extrinsic incentives can have detrimental effects on prosocial behavior: extrinsic incentives might interact with image motivation by diluting the signaling value of prosocial behavior. Simply put, when observing a prosocial activity, the addition of extrinsic incentives makes it more difficult to answer the question: "Is the individual behaving prosocially to do good or to do well?" This will affect the image gain from prosocial behavior. Think of the following scenario: an individual is considering buying a new hybrid car. This car is more expensive than an equivalent car operating with a standard gasoline engine, but the hybrid car helps in preserving the environment. Driving around a car which is clearly a hybrid car would probably add to one's positive image, especially if she lives in a community that values environmentally-friendly technologies. Suppose now that the government gives a large tax benefit for those who decide to purchase a hybrid car (and everybody knows about this). The tax incentive, of course, reduces the price of hybrid cars and therefore should make the hybrid car more attractive for the individual. However, the tax incentive also decreases the image value of driving the hybrid car. Without the tax incentive, buying the hybrid car definitely shows the individual cares for the environment (positive image), with the tax incentives, it does not. More generally, when extrinsic incentives are provided (such as a tax benefit), it is difficult to conclude whether the prosocial act is due to one's innate good traits (the person's concern for the environment) or due to greed (receiving a tax benefit). Therefore, if image indeed motivates prosocial behavior, introducing extrinsic rewards may reduce image motivation, which can lead to a negative net effect (relative price effect net of crowding out effect) on

4

prosocial behavior.

Benabou and Tirole (2006) provide a nice model, recognizing image motivation as a driver of prosocial behavior, which provides testable hypotheses on image as a motivation, and whether extrinsic incentives crowd-out image motivation (for similar models, see Maarten C.W. Janssen and Ewa Mendys-Kamphorst, 2004, Paul Seabright, 2004). The testable implications make use of the properties of the different incentives driving prosocial behavior: intrinsic and extrinsic motivations are personal—altruism, as an example of intrinsic motivation, is a personal taste, while monetary payment and gifts, as examples of extrinsic motivation, are personal material rewards—yet image motivation is a social motivation that depends on the visibility of one's actions. Importantly, the interaction between extrinsic incentives and image motivation depends on the prevalence of image motivation and therefore on the visibility of the prosocial behavior. In anonymous prosocial decisions, for instance, extrinsic monetary incentives will not crowd-out image motivation, as in this setting there is no image motivation to begin with. If extrinsic motivation does crowd-out image motivation, we expect that extrinsic incentives will be more effective in private decisions to behave prosocially than in decisions which are publically visible. Consequently, extrinsic incentives are more likely to be counter-productive in public rather than private decision-making.

We conduct an experimental study to investigate whether image considerations indeed play a part in motivating prosocial behavior, and how this motivation interacts with extrinsic monetary rewards. Our test of the interaction between extrinsic and image motivation uses the unique property of image concerns — its dependence on visibility. More specifically, in the main experiment, named "Click for Charity," subjects could donate to a charitable organization by clicking two keys on the keyboard. The charities (the American Red Cross or the National Rifle Association) were either associated with positive or negative image consequences. Participants were randomly assigned to make their donation choices in a public or in a private setting, and were randomly assigned to receive, in addition to the donation made on their behalf, monetary incentives that accrued solely to them. The Click for Charity

5

results strongly support the hypotheses that image motivation is important for prosocial behavior, and that private monetary incentives partially crowd-out image motivation. That is, we find that monetary incentives have no effect on contribution effort when contribution decision is made public, while monetary incentives do increase the contribution effort when the decision is private. We were able to further examine some of these hypotheses in a field study (which is more limited due to our setup), and in general the results support what we find in the lab: private monetary incentives seem to interact negatively with image concern, leading to the result that monetary incentives are more effective in motivating private prosocial decisions than ones made in public.

The paper proceeds as follows: Section 2 presents the behavioral hypotheses. Section 3 presents the experimental design, and section 4 the results. Section 5 presents the illustrative field study. Section 6 discusses the implications of our findings and concludes.

2 Behavioral Hypotheses

This section develops behavioral hypotheses to test for image as a driver in prosocial behavior and to test whether extrinsic monetary incentives crowd-out image motivation. We use the model of Benabou and Tirole (2006) as our framework. In this model, an individual's incentives to behave prosocially are divided into three basic motivations: intrinsic (v_a), extrinsic (yv_y), and image motivation $R(a, y)$. Formally, an agent's utility is:

$$U(a) = (v_a + yv_y)a + R(a, y) - C(a),$$

where a is the extent to which the agent behaves prosocially, y is the extrinsic or monetary rewards (valued at v_y), and $C(a)$ is the cost of behaving prosocially. (v_a, v_y) are assumed to be normally distributed, $C(a) = \frac{ka^2}{2}$, and $R(a, y)$ is defined as:

$$R(a, y) = x[\gamma_a E(v_a|a, y) - \gamma_y E(v_y|a, y)].$$

Here, x is the visibility of action a, γ_a is the agent's level of concern for being considered an altruistic or prosocial human-being, $E(v_a|a,y)$, and γ_y is the agent's level of concern for being perceived as greedy, $E(v_y|a,y)$.

Given fixed image concerns, Benabou and Tirole (2006) show that the individual supply of prosocial activity is:

$$a^* = \frac{v_a + v_y y}{k} + xr(y).$$

The aggregate supply, $\bar{a}$, is:

$$\bar{a} = \frac{\bar{v}_a + \bar{v}_y y}{k} + xr(y),$$

where $(\bar{v}_a, \bar{v}_y)$ are the population's averages, and $xr(y)$ is the constant (marginal) image motivation[3].

This model provides testable hypotheses about the conditions under which image motivations should be prevalent and about the interaction between monetary incentives and image value. We discuss each of these hypotheses next. Formal exposition and proofs are left for the Appendix.

2.1 Image Motivation

The model above gives an explicit formulation of image motivation, which depends on the public nature of one's actions, represented in the model by the visibility parameter x. An action can be done anonymously, or publicly, meaning that x is much greater in the latter than in the former case, and it acts as a magnifier: $xr(y)$[4]. Consequently, if image motivation positively affects prosocial behavior at some visibility level x ($r(y) > 0$), then, all else equal, greater public exposure should increase the level of prosocial activity $\bar{a}$. Similarly, if image motivation negatively affects prosocial behavior at some visibility level x ($r(y) < 0$), then, all else equal, greater public exposure should decrease the level of prosocial activity $\bar{a}$. Hence

---

[3] $r(y) = \frac{\partial R(a,y)}{\partial a}|_{x=1}$, independent of a.
[4] In the case of heterogeneous image concerns the impact of visibility may be weaker (Benabou and Tirole 2006).

if image is a driver in prosocial activity, then the following should be true:

Image-Motivation Hypothesis: *Ceteris paribus, changing visibility x changes prosocial activity $\bar{a}$. For a positive image $(r(y) > 0)$, increasing x increases $\bar{a}$. For a negative image $(r(y) < 0)$, increasing x decreases $\bar{a}$.*

Note that the case of self-image does not affect the Image-Motivation Hypothesis if self-image is independent of x. In the experimental set-up, self-image can be thought of as a constant in all conditions and does not affect the conclusions.

2.2 The Effect of Extrinsic Incentives and Crowding-Out of Image Motivation

Introducing extrinsic incentives, $y > 0$, (or increasing thereof) affects prosocial behavior $\bar{a}$ both directly by increasing the extrinsic rewards, and through image motivation $r(y)$. Higher personal benefits associated with activity a tends to decrease the social preference signal and increase the greediness signal, resulting in lower image motivation, $r(y)$.[5] In such a case, offering a greater material reward may backfire, depending on which effect is stronger—higher extrinsic reward or lower image motivation.

Importantly, the effect of extrinsic rewards on image motivation implies that visibility, x, may influence the effectiveness of material rewards. That is, if receiving an extrinsic reward (y) reduces image motivation, then greater publicity (x) reduces the effectiveness of material reward. This leads to our second hypothesis, the Effectiveness Hypothesis:

Effectiveness Hypothesis: *Extrinsic rewards, y, are less effective the greater is the visibility of the prosocial act, x.*

The Effectiveness Hypothesis implies crowding-out of image motivation, $r'(y) < 0$.[6] This

---

[5]This depends on the level of extrinsic reward y, and on whether the correlation between the agent's preferences v_a and v_y is not too negative (see Appendix).

[6]This also holds for the case of heterogenous image concerns provided that the direct effect of visibility on the effectiveness of monetary rewards, $r'(y)$, dominates the indirect effect of visibility via image variance (one may suspect prosocial behavior for image concerns).

conclusion is possible even if extrinsic incentives do not backfire, i.e., $a'(y) > 0$. This occurs whenever the direct increase in utility from extrinsic rewards outweighs any decline in image motivation. As a result, we have three possible cases:

1. $a'(y) \leq 0$: this clearly indicates crowding-out

2. $a'(y) > 0$, and $r'(y) < 0$: crowding-out of image motivation

3. $a'(y) > 0$, and $r'(y) \geq 0$: there is no crowding-out.

Whenever $a'(y) > 0$, it is difficult to conclude whether extrinsic rewards crowd-out image motivation. However, according to the Effectiveness Hypothesis, extrinsic rewards are less effective with more visibility only as long as material rewards reduce image motivation. Therefore, the case of $a'(y) > 0$ in which visibility reduces the effectiveness of extrinsic rewards implies crowding-out. As a result, a detrimental effect of extrinsic incentives is more likely to occur for a visible prosocial effort than for a private one.

In what follows we test these hypotheses experimentally.

3　Experimental Design—Click for Charity

The discussion thus far showed that image uniquely depends on visibility, x, for its motivational effect. We use this property to test for the presence of image motivation and for its interaction with monetary incentives, y. The experimental design has to fulfill three criteria: (1) the prosocial task needs to have clear image implications by allowing one to signal personal traits, (2) one's decision has to be made either public or private (exogenously determined), and (3) extrinsic monetary incentives have to be exogenously determined— either offered or not —and to be publicly known.

In this study, participants could behave prosocially in a real-effort task: sequentially pressing two keys— X and Z —on the keyboard for up to 5 minutes[7]. For every completed

---

[7]Although the task was chosen for the relatively low variance in ability to perform it, still such variance

pair of clicks, we paid money in the participant's name to an assigned charity, hence, the name "Click for Charity." The donations were made according to a decreasing payment schedule: 1 cent was donated for each of the first 200 pairs of X-Z presses, 0.5 cents for each of the next 200 pairs, 0.25 cents for each of the next 200 pairs, ... , and 0.01 cents for each pair above 1,200. Although marginally decreasing, the more pairs pressed, that is, the higher the effort, the greater was the donation to the assigned charity. Before the actual task started, each participant practiced the task for 30 seconds.[8] 161 Princeton undergraduate students participated in this study; none knew the nature of the experiment before entering the lab.

This study used a between-subjects 2x2x2 design, with each of the participants randomly assigned to one of the conditions. The conditions differed according to the nature of the charity, the task's publicity, and the compensation scheme, as explained below.

The nature of the cause was manipulated by randomly assigning subjects to one of two charities: the American Red Cross or the National Rifle Association (NRA). Making efforts to benefit one of these charities is expected to have signaling value. Whether the value of the signal is positive or negative depends at least partially on the norms and values of the community and therefore on the participant's perceptions of others' (observers) opinions (e.g. the two charities might be perceived differently at Princeton and at a University in a pro-gun state). In order to classify the two charities into "good" or "bad," we asked participants about the extent to which, in their opinion, the majority of students at Princeton identify with each of the two charities.[9] This occurred before participants were assigned a charity and it was done on an 11-point scale from -5 "not at all identify" to +5 "very much identify." According to this perception, we classified the assigned charity as "good" if the individual's

---

may exist. That is, some people are better in pressing keys than others, which is equivalent to heterogeneous cost function $C(a, \theta)$, where θ is idiosyncratic. Indeed, Benabou and Tirole (2006) assume a homogeneous cost function. However, since people presumably don't signal their ability in such a task, and this ability is not part of v_a, the theory could be extended to this case as well.

[8]The instructions for the task can be found at:
http://www.bos.frb.org/economic/econbios/meier/Instructions_Task_Click.pdf

[9]In fact, we asked participants about the identification of Princeton students with a number of charities.

perceived identification was positive and as "bad" if her perceived identification was negative. Seven subjects stated a neutral identification (exactly zero). In the main analysis we focused on the remaining 154 subjects; however, including the seven neutral subjects and assuming— according to the average perception of the other participants— that the NRA is "bad" and the American Red Cross is "good" does not change the results. The results also do not change if we classify the NRA for all as "bad" and the Red Cross for all as "good," following the view of the majority of the participants.

The visibility of the task was controlled by randomly assigning subjects to make their effort choice in "public" or in "private". In the private condition, the effort decision was made anonymously, while in the public condition, the choice was disclosed to others. At the end of the task, we asked those in the public treatment to tell the other participants in the lab which charity they were assigned, whether they also earned money for themselves (based on the payment scheme below), and how much money they donated to the charity (and got for themselves).

Finally, *the payment scheme* was such that a random selection of participants could, in addition to the donation, earn money for themselves. The private payment was exactly the same as the donation schedule described above.

The following section presents the results of the study.

4   Results

We present the results in three steps: first, we present the analysis for "good" and "bad" causes separately, then we present the analysis for all causes together, and lastly we test whether a physical limit on the ability to continuously press keys can explain the results.

Figure 1 shows the effect of private monetary incentives in public compared to in private for clicking for a good cause, where a good cause is defined as what participants think that others identify with. The figure shows two important results:

First, without any private monetary incentives, subjects put forth significantly more effort in the public than in the private condition. In the public condition, they pressed, on average, approximately 822 pairs, while in the private condition, subjects pressed only about 548 pairs ($p < 0.05$; t-test[10]). In line with the Image-Motivation Hypothesis, subjects exerted more effort for a good cause in public, where they are able to signal their effort to many others.

[Figure 1 about here]

Second, while monetary incentives do not increase effort in public, they do increase effort significantly in the private condition. In the public condition, there is an insignificant decline in effort from 822 to 702 key pairs on average. In the private condition, there is a significant increase in effort from 548 to 740 key pairs on average ($p < 0.05$). As we show at the end of this section, this result is not driven by subjects in the public condition reaching a physical limit. Therefore, this result supports the Effectiveness Hypothesis that private monetary incentives are less effective in public than in private. This implies that monetary incentives crowd-out the image motivation to behave prosocially.

Figure 2 shows the effect of private monetary incentives for a "bad" cause–a charity that subjects think other Princeton students do not identify with. The figure shows that without incentives, subjects do not contribute less in public than in private as proposed in the Image Motivation Hypothesis. In fact contributions are not significantly different in public and private. But contributions are on a very low level and censoring might be an issue here. The figure, however, shows strong support for the Effectiveness Hypothesis. While private monetary incentives do slightly decrease effort in the public condition (although not statistically significant), in the private condition monetary incentives increase the effort from 204 key pairs to 559 key pairs ($p < 0.05$). That is, offering monetary incentives is less effective at increasing effort in public than in private even for a task with negative image value. As effort for a "bad" cause is at a very low level, physical limit cannot explain the differential

---

[10]We will use t-tests to compare all means, unless otherwise noted.

effect of monetary incentives in the private and public conditions.

[Figure 2 about here]

Analyzing good and bad causes together, Table 1 shows the difference between private and public conditions in an OLS regression where the number of key pairs is the dependent variable. As independent variables, we control for the type of the assigned charity, whether the decision was private or public, and whether subjects received private monetary incentives for their effort. Column 1 shows that, in the private condition, monetary incentives increase effort significantly ($p < 0.01$). In contrast, column 2 shows that for subjects in the public sphere, private monetary incentives do not increase individual effort. The effect is actually slightly negative but not statistically significant. As seen in column 3, the difference in the effectiveness of monetary incentives in the public versus the private condition (i.e. the interaction term of private monetary incentives in the public condition) is substantial and statistically significant at the 95 percent level.[11] The results support the Effectiveness Hypothesis that monetary incentives are less effective in the presence of image motivation, i.e. if prosocial behavior is public.

[Table 1 about here]

One concern while interpreting the differential effect of monetary incentives in the public compared to the private condition is that the effect may be due to a physical limit. Subjects exerting effort for a good cause in the public condition may reach a physical limit even without incentives. In this case, it might be impossible to increase effort with monetary incentives, even if subjects wanted to. While the analysis for the "bad" cause suggests otherwise, since effort is at a low level in this condition, we nonetheless tested explicitly whether participants reached a physical limit in the good-cause condition.

To find where people's physical limit is, we ran an additional control treatment, the "Limit" treatment, in which 26 students from the same subject pool at Princeton partici-

---

[11] As seen in table A.1 in the appendix, the result is robust to adding the subjects who thought that Princeton students neither identify nor don't identify with their assigned charity. We assumed that they were putting effort in for a "good" cause or for a "bad" cause according to the identification of the majority of Princeton participants (i.e., NRA is coded as "bad" and Red Cross is coded as "good").

pated. In this treatment, participants performed the same key-pressing task for 5 minutes, and their effort decision remained anonymous. However, in this condition the task was not a donation task; as the subjects could only earn money for themselves by pressing the keys, and they were paid high financial incentives to do so. Subjects got a flat rate of 0.11 cents per pair, plus a bonus, which increased from $2, if they pressed more than 1,000 pairs, to $20, if they pressed more than 1,500 pairs.

The average number of paired key presses in the Limit treatment was 1,290 (s.d. 44) pairs, while subjects pressed only 822 (s.d. 98) key pairs for a "good" charity in the public sphere. The difference between the key presses in the two conditions is statistically significant ($p < 0.01$) and is illustrated in Figure 3 by the cumulative distributions of key pairs for the "good" cause in the public condition and for the Limit treatment. Note that while in the Limit treatment nobody pressed less than 900 pairs, in the Click for Charity experiment 50 percent of the participants did so. Therefore, the ineffectiveness of monetary incentives in the public condition cannot be attributed to a physical limit, as most subjects could have increased their effort substantially.

[Figure 3 about here]

Thus, the results of the lab experiment Click for Charity show that the Image-Motivation Hypothesis cannot be rejected for a good cause, i.e., image is certainly a significant driver of prosocial behavior. Analyzing the results of introducing private monetary incentives we find a differential effect in the private versus the public condition. For a "good" and "bad" charitable cause, effort dramatically increases with monetary incentives in the private condition, while it does not change significantly in the public condition. Together with the Limit treatment, these results confirm the Effectiveness Hypothesis that private monetary incentives are less effective the greater the publicity. The results also imply that private monetary incentives crowd-out image concerns.

14

5 Field Study—Bike for Charity

The Click for Charity results support the view that extrinsic incentives may have detrimental effects by crowding-out image motivation. This depends on the extent of public exposure of the prosocial activity, which, in turn, has important implications for the use of extrinsic incentives in facilitating contributions to public goods. To strengthen the policy implications of these results, below we present a first-pass illustration of the results in the field.

To illustrate whether the effect found in the Click for Charity experiment also translates to a field setting, we launched a similarly-designed campaign at the MIT gym, titled "Bike for Charity". In the Bike for Charity campaign, participants could donate to a "good" or "bad" charity by cycling on stationary bikes for up to 10 minutes, at a donation rate of $1 per mile. As in Click for Charity we implemented a between-subject 2x2x2 design. The cause of the charity was manipulated as in Click for Charity.[12] The visibility was manipulated by randomly assigning subjects to bike in the public view (on the second floor of the gym, where everyone works out) or in a private location (the third floor of the gym, in a private room). The participants were also assigned to receive either private monetary rewards for themselves or not. Subjects who participated in the first few weeks of the campaign did not receive any private monetary incentives, while those who participated in the last weeks of the campaign, did.

Due to the lack of a randomized compensation scheme, and to the fact that participants knew in advance that they would be randomly assigned to bike for either a "good" or a "bad" charity, selection bias is a concern. More specifically, we expect that individuals whose image is affected the most by monetary incentives will avoid conditions with private monetary incentives. This, in turn, could dampen the interaction between extrinsic incentives and image concerns. Similarly, if the nature of the cause is unclear (the most extreme case

---

[12] In the "Bike for Charity" we used four controversial charities: Two of the four focus on gun control and are either pro or against stronger gun control (the National Rifle Association (NRA) and the Educational Fund to Stop Gun Violence), the other two charities focus on abortion and lobby either for or against it (the Pro-Life Association and the Pro-Choice Action League).

is random assignment, as in our study), the more image-concerned a person is, the less likely she would be to participate in such a campaign—to avoid the chance of being associated with a "bad" cause. Both selection biases act against the image and the crowding-out hypotheses.

Despite the limits of the Bike for Charity design, the results present suggestive field evidence for the Effectiveness Hypothesis. 151 participants were assigned to bike either for a "good" or a "bad" charity, where "good" and "bad" charities are defined according to participant's perception of MIT student's identification, as in the Click for Charity study. For these participants, adding private incentives is, on average, more effective in private than in public. Table 2 shows this effect in an OLS regression with distance biked in miles as the dependent variable. We use the same independent variables as in Click for Charity, plus a variable for the participant's gender. Table 2 shows that extrinsic incentives increases effort in the private condition (Column 1) but not in the public condition (Column 2), and as can be seen in column 3 this difference is statistically significant ($p < 0.1$).[13]

[Table 2 about here]

The effect obtained in the field is not as strong as in the lab. However, as discussed, the lab results are obtained in a fully randomized setting where participants could not self-select into a condition based on its characteristics (a good/bad cause, and with/without private monetary incentives), while in the field they partly could. Therefore, the suggestive evidence is encouraging—in spite of the selection bias we observe differential effects of monetary incentives between the public and the private conditions. Future research should further investigate the interaction between extrinsic incentives and image motivation in the field.

---

[13] Adding individuals to the analysis who believe that the charities are seen by other students as neither positive nor negative and therefore stated a 0 on scale from -5 "not at all identify" to +5 "very much identify," as in Click for Charity, does not change the results qualitatively but increases the standard errors of the estimations.

6   Conclusions

The results of the laboratory experiment, illustrated also in a field study, support our hypothesis that monetary incentives depend crucially on visibility: monetary incentives are more effective in facilitating private, rather than public, prosocial activity. People want to be seen as doing good; without extrinsic incentives, an observer will attribute the prosocial act to one's good trait which motivates people to behave prosocially. But with extrinsic incentives, the signal of a prosocial act gets diluted as one might behave prosocially mainly to do well. The image value decreases, and the incentive becomes less effective. If no one is watching (i.e. the prosocial decision is private) the incentive to be doing well cannot dilute any signal to others, and consequently extrinsic incentives are very likely to increase prosocial behavior.

Our results have important policy implications; for example, if a government considers a tax benefit policy to facilitate the adoption of a new environmentally friendly technology, it should expect the policy to be more successful for a non-visible technology, such as environmentally friendly water boilers, relative to a visible technology, such as hybrid cars. This is because hybrid cars, most of which are clearly visible, may partly be purchased as a signaling medium, while water boilers probably are not. By giving tax benefits, the government might unintentionally damage the signaling value vested in hybrid cars. The net effect of the incentive depends on the strength of the price effect compared with the crowding-out effect (see also Uri Gneezy, 2003).

It is important to note that image crowding-out by monetary incentives is shown here when both the prosocial decision and the extrinsic incentives are public. If extrinsic incentives exist but are not publicly known, others' opinion should be the same as if no extrinsic rewards existed. Consequently, major donors to cultural organizations might not want to make their extrinsic incentives (e.g. favorable seats) public, as it would dilute the image signal of their contribution. If decision-makers, such as fund-raisers or policy makers, anticipate the Effectiveness and Crowding-Out Hypotheses, they should use fewer public extrinsic incentives

for visible prosocial activities. This brings us back to Titmuss and his intuition that monetary incentives might reduce pro-social behavior. The research in this paper contributes to the discussion in showing that Titmuss's claim crucially depends on the visibility of the prosocial decision. Therefore if blood donations are sufficiently public, there is good reason to doubt the effect of extrinsic incentives and even to expect detrimental effect of monetary incentives.

Another related question is whether crowding-out of image motivation depends on the type of the extrinsic incentive used. In this study we explore the effect of monetary incentives, which may be different from the effect of non-monetary incentives. To continue with the hybrid car example, the incentive of using the carpool lane may not dilute the signal as much as monetary incentives would. This is left for future work.

Figure 1: Effect of Private Incentive for "Good" Charity
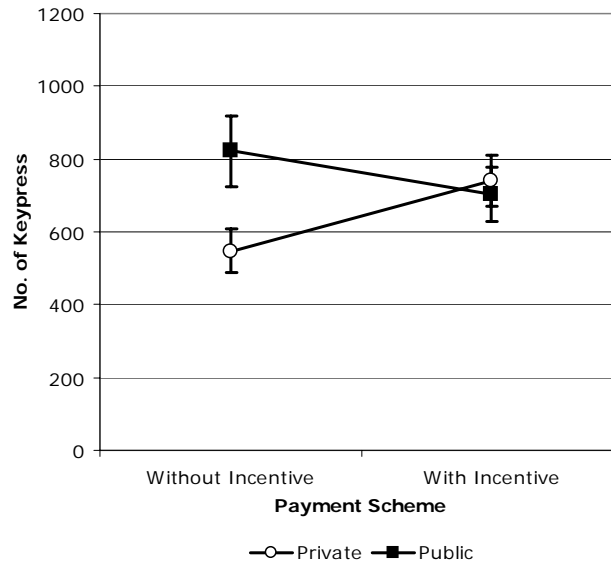
Note: Charity classified according to participants' perception of whether Princeton students identify with charity ("Good" if others' identification is greater than zero on a scale from -5 "not at all identify" to +5 "very much identify"). Error bars are based on standard errors of the mean.

# Figure 2: Effect of Private Incentive for "Bad" Charity
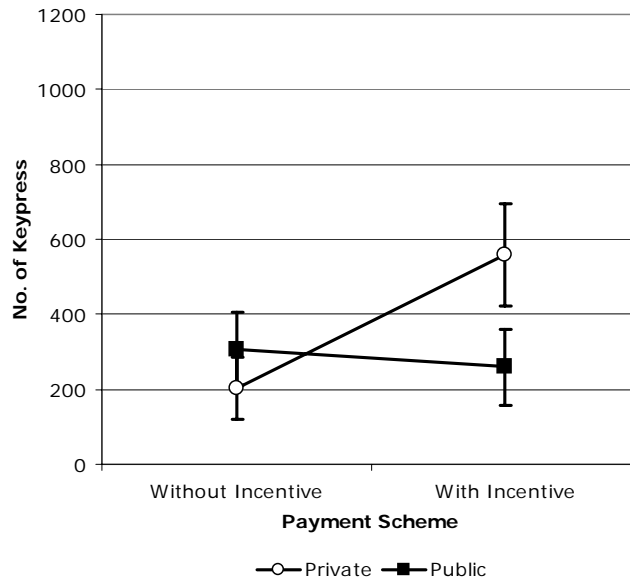
Note: Charity classified according to participants' perception of whether Princeton students identify with charity ("Bad" if others' identification is lower than zero on a scale from -5 "not at all identify" to +5 "very much identify"). Error bars are based on standard errors of the mean.

Figure 3: Cumulative Distribution of Number of Key Pairs for "Good" Cause and "Limit" Treatment
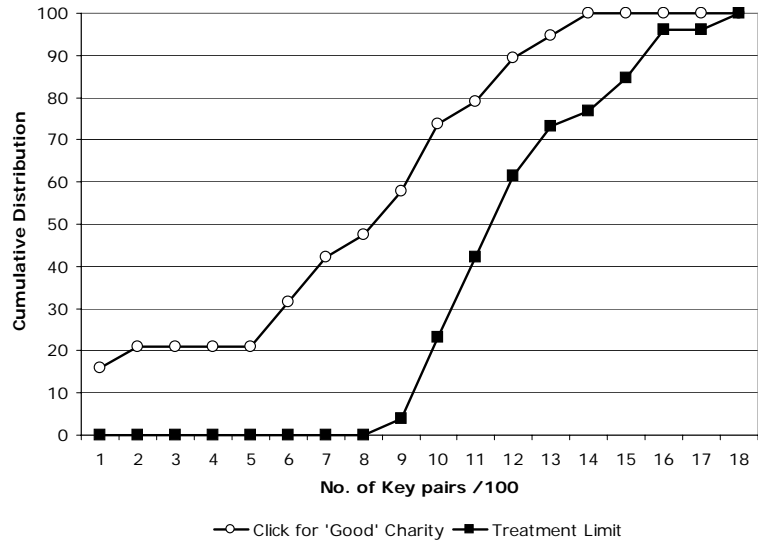
Table 1: Effect of Private Incentives in "Click for Charity"

Dependent variable: Number of key pairs

	(1) Private	(2) Public	(3) All
Private monetary incentive (=1)	252.97 (83.93)***	-90.81 (92.34)	249.87 (84.45)***
Public sphere (=1)			187.52 (85.91)**
Private incentive*Public			-315.83 (123.76)**
'Good' cause (=1)	269.68 (89.92)***	483.51 (95.12)***	375.92 (65.93)***
Constant	249.29 (75.81)***	322.18 (84.37)***	184.51 (64.61)***
N	77	77	154
R squared	0.21	0.25	0.22

Note: OLS regressions. Robust standard errors in parenthesis. Charity classified according to participants' perception of whether Princeton students identify with charity ("Good" if others' identification is greater than zero and "Bad" if others' identification is lower than zero on a scale from -5 "not at all identify" to +5 "very much identify").

Significance level: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

Table 2: Effect of Private Incentives in "Bike for Charity"

Dependent variable: Distance in miles

	(1) Private	(2) Public	(3) All
Private monetary incentive (=1)	0.56	-0.02	0.56
	(0.20)***	(0.23)	(0.20)***
Public sphere (=1)			0.28
			(0.21)
Private incentive*Public			-0.58
			(0.30)*
'Good' cause (=1)	0.84	0.84	0.85
	(0.22)***	(0.26)***	(0.17)***
Female (=1)	-0.54	-0.46	-0.50
	(0.20)***	(0.22)**	(0.15)***
Constant	1.50	1.75	1.48
	(0.24)***	(0.29)***	(0.20)***
N	69	82	151
R squared	0.33	0.18	0.24

Note: OLS regressions. Robust standard errors in parenthesis. Charity classified according to participants' perception of whether MIT students identify with charity ("Good" if others' identification is greater than zero and "Bad" if others' identification is smaller than zero on a scale from -5 "not at all identify" to +5 "very much identify").
Significance level: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

7   Appendix

The appendix lays out the analysis, based on the model of Benabou and Tirole (2006), that leads to the hypotheses in the Behavioral Hypothesis Section. In the analysis leading to the behavioral hypotheses we focus on the case where types $v \equiv (v_a, v_y)$ are distributed normally in the population $(v_a, v_y) \sim N\left(\begin{array}{c} \bar{v}_a \\ \bar{v}_y \end{array}, \begin{bmatrix} \sigma_a^2 & \sigma_{ay} \\ \sigma_{ay} & \sigma_y^2 \end{bmatrix} \right)$, $\bar{v}_a \gtreqless, \bar{v}_y > 0$, and individuals are assumed to have constant image attitudes $(\mu_a, \mu_y) \equiv (x\gamma_a, x\gamma_y) = (\bar{\mu}_a, \bar{\mu}_y)$. Under these assumptions, the optimal level of a is given by Proposition 1 in BT:

Proposition 1 (BT1) *Let all agents have the same image concern* $(\bar{\mu}_a, \bar{\mu}_y)$, *and cost function* $c(a) = ka^2/2$. *There is a unique (differentiable-reputation) equilibrium, in which an agent with preferences* (v_a, v_y) *contributes at the level*

$$a^* = \frac{v_a + v_y y}{k} + \bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y),$$

where $\rho(y) \equiv \frac{\sigma_a^2 + y\sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2}$ *and* $y\chi(y) \equiv 1 - \rho(y)$. *The reputational returns are* $\frac{\partial E(v_a | a^*, y)}{\partial a} = \rho(y)k$ *and* $\frac{\partial E(v_y | a^*, y)}{\partial a} = \chi(y)k$, *resulting in a net value* $kxr(y) = k(\bar{\mu}_a \rho(y) - \bar{\mu}_y \chi(y))$, *independent of* a.

Proof. The proof of this proposition is omitted and can be found in BT. ∎

A straight-forward conclusion is the Image-Motivation Hypothesis, next:

Proposition 2 $\frac{\partial a(y)}{\partial x} > 0 \Leftrightarrow r(y) > 0.$

Further calculation leads to the following proposition:

Proposition 3 *Let* $\sigma_{ay} = 0$ *and define* $\theta \equiv \frac{\sigma_y}{\sigma_a}$. *If* $r(y) > 0$, $\frac{\partial a'(y)}{\partial x} < 0$.

Proof. The change in prosocial behavior as we change the extrinsic reward y is given by:

$$a'(y) = \frac{v_y}{k} + \bar{\mu}_a \rho'(y) - \bar{\mu}_y \chi'(y).$$

24

$$\frac{\partial a^{'}(y)}{\partial x} = \overline{\gamma}_a \rho^{'}(y) - \overline{\gamma}_y \chi^{'}(y)$$

$$
\begin{aligned}
\rho(y) &= \frac{\sigma_a^2 + y\sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2} = \frac{\sigma_a^2}{\sigma_a^2 + y^2\sigma_y^2} \text{ if } \sigma_{ay} = 0 \Rightarrow \rho^{'}(y) = -2\chi(y)\rho(y) \\
\chi(y) &= \frac{y\sigma_y^2 + \sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2} = \frac{y\sigma_y^2}{\sigma_a^2 + y^2\sigma_y^2} \text{ if } \sigma_{ay} = 0 \Rightarrow \chi^{'}(y) = \frac{\sigma_y^2}{\left(\sigma_a^2 + y^2\sigma_y^2\right)}\left(\rho(y) - y\chi(y)\right)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial a^{'}(y)}{\partial x} &= \overline{\gamma}_a \rho^{'}(y) - \overline{\gamma}_y \chi^{'}(y) \\
&= -\chi(y)\left[\overline{\gamma}_a\rho(y) - \overline{\gamma}_y\chi(y)\right] - \overline{\gamma}_a\chi(y)\rho(y) - \overline{\gamma}_y\frac{\chi(y)\rho(y)}{y}
\end{aligned}
$$

If $r(y) > 0$, i.e., $\left[\overline{\gamma}_a\rho(y) - \overline{\gamma}_y\chi(y)\right] > 0$, $\frac{\partial a^{'}(y)}{\partial x} < 0$ ∎

Proposition 3 implies that, provided a positive net image when $x = 1$ (private), the greater the publicity of one's actions, the less effective are the extrinsic rewards associated with giving. Note that $\frac{\partial a^{'}(y)}{\partial x} < 0 \Leftrightarrow r^{'}(y) = \overline{\gamma}_a\rho^{'}(y) - \overline{\gamma}_y\chi^{'}(y) < 0$. This leads us to the Effectiveness Hypothesis, and its implication for crowding-out.

The hypotheses above are for the simple case of no correlation between altruistic taste and personal value of extrinsic rewards ($\sigma_{ay} = 0$). However, this does not affect our hypotheses. The Image-Motivation Hypothesis says that if image motivation is indeed one of the drivers of prosocial behavior and it is positive for some visibility level, then increasing the visibility of one's actions will increase the level of prosocial behavior a. This hypothesis holds equally well even if $\sigma_{ay} \neq 0$. To see this, note that the image motivation, $r(y)$, is influenced by σ_{ay}:

$$
\begin{aligned}
r(y) &= \overline{\gamma}_a\rho(y) - \overline{\gamma}_y\chi(y) \\
&= \frac{\overline{\gamma}_a(\sigma_a^2 + y\sigma_{ay}) - \overline{\gamma}_y(\sigma_{ay} + y\sigma_y^2)}{(\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2)}
\end{aligned}
$$

However if $r(y)$ is positive, increasing the visibility of one's action by $x > 1$ acts only as an amplifier: $xr(y)$.

Similarly, the claim made in the Effectiveness Hypothesis is valid provided that σ_{ay} is not too large. To see this, let's examine the following object:

$$\frac{\partial a'(y)}{\partial x} = \frac{\partial \bar{\mu}_a}{\partial x}\rho'(y) - \frac{\partial \bar{\mu}_y}{\partial x}\chi'(y).$$

Since $\rho'(y)$ and $\chi'(y)$ are functions of σ_{ay}, relaxing the assumption that $\sigma_{ay} = 0$ does influence the Effectiveness Hypothesis. However, it can be shown that, as long as σ_{ay} is not too large, the conclusion still holds.

Proof.

$$\rho(y) = \frac{\sigma_a^2 + y\sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2} \Rightarrow \rho'(y) = \frac{-\sigma_a^2\sigma_{ay} - 2\sigma_a^2 y\sigma_y^2 - y^2\sigma_y^2\sigma_{ay}}{\left(\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2\right)^2}$$

$$\chi(y) = \frac{y\sigma_y^2 + \sigma_{ay}}{\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2} \Rightarrow \chi'(y) = \frac{\sigma_y^2\sigma_a^2 - y^2\sigma_y^4 - 2\sigma_{ay}^2 - 2y\sigma_y^2\sigma_{ay}}{\left(\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2\right)^2}$$

$$\frac{\partial a'(y)}{\partial x} = \overline{\gamma}_a\rho'(y) - \overline{\gamma}_y\chi'(y)$$

$$= \overline{\gamma}_a\frac{-\sigma_a^2\sigma_{ay} - 2\sigma_a^2 y\sigma_y^2 - y^2\sigma_y^2\sigma_{ay}}{\left(\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2\right)^2} - \overline{\gamma}_y\frac{\sigma_y^2\sigma_a^2 - y^2\sigma_y^4 - 2\sigma_{ay}^2 - 2y\sigma_y^2\sigma_{ay}}{\left(\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2\right)^2}$$

If $\sigma_{ay} \neq 0$:

$$\frac{\partial a'(y)}{\partial x} = \frac{\sigma_{ay}^2(2\overline{\gamma}_y) + \sigma_{ay}\left(2\overline{\gamma}_y y\sigma_y^2 - \overline{\gamma}_a\sigma_a^2 - \overline{\gamma}_a y^2\sigma_y^2\right) - 2\overline{\gamma}_a\sigma_a^2 y\sigma_y^2 - \overline{\gamma}_y\left(\sigma_y^2\sigma_a^2 - y^2\sigma_y^4\right)}{[\sigma_a^2 + 2y\sigma_{ay} + y^2\sigma_y^2]^2}$$

$$\Rightarrow \frac{\partial a'(y)}{\partial x} < 0$$

$$\Leftrightarrow (2\overline{\gamma}_y)\sigma_{ay}^2 + \sigma_{ay}(2y\overline{\gamma}_y\sigma_y^2 - \overline{\gamma}_a\sigma_a^2 - \overline{\gamma}_a y^2\sigma_y^2) - \overline{\gamma}_a 2y\sigma_a^2\sigma_y^2 - \overline{\gamma}_y[\sigma_a^2\sigma_y^2 - y^2\sigma_y^4] < 0$$

$$(2\overline{\gamma}_y)\sigma_{ay}^2 + \sigma_{ay}(2y\overline{\gamma}_y\sigma_y^2 - \overline{\gamma}_a\sigma_a^2 - \overline{\gamma}_a y^2\sigma_y^2) - \overline{\gamma}_a 2y\sigma_a^2\sigma_y^2 - \overline{\gamma}_y[\sigma_a^2\sigma_y^2 - y^2\sigma_y^4] = 0$$

$$(A)x^2 + x(B) - (C) = 0.$$

For $\sigma_{ay} = 0$ $\frac{\partial a'(y)}{\partial x} < 0$; this implies that there is a solution to the above quadratic equation in σ_{ay} $(C < 0 \Rightarrow B^2 - 4AC > B^2 > 0)$. Therefore, as long as σ_{ay} is small enough, $\frac{\partial a'(y)}{\partial x} < 0$.

∎

Table A.1: Effect of Private Incentives in "Click for Charity"

Dependent variable: Number of key pairs

	(1) Private	(2) Public	(3) All
Private monetary incentive (=1)	239.03 (81.45)***	-88.52 (90.90)	234.61 (82.49)***
Public sphere (=1)			153.06 (85.24)*
Private incentive*Public			-287.66 (121.93)**
'Good' cause (=1)	216.19 (86.17)**	486.74 (92.97)***	346.28 (64.14)***
Constant	302.29 (75.85)***	317.65 (78.97)***	226.5 (65.67)***
N	82	79	161
R squared	0.17	0.26	0.20

Note: OLS regressions. Robust standard errors in parenthesis. Charity classified according to participants' perception of whether Princeton students identify with charity ("Good" if others' identification is greater than zero and "Bad" if others' identification is lower that zero on a scale from -5 "not at all identify" to +5 "very much identify"). For those stating 0, the assigned charity got classified according to the perception of the majority: NRA is perceived as a "bad" cause and Red Cross as a "good" cause.

Significance level: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

References

[1] Akerlof, George A. "A Theory of Social Custom, of Which Unemployment May Be One Consequence." Quarterly Journal of Economics, 1980, 94(4), pp. 749-75.

[2] Andreoni, James and Petrie, Ragan. "Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising." Journal of Public Economics, 2004, 88(7-8), pp. 1605-23.

[3] Benabou, Roland, and Tirole, Jean. "Incentives and Prosocial behavior." American Economic Review, 2006, 96(5), pp. 1652-1678.

[4] Blumkin, Tomer, and Sadka, Efraim. "A Case for Taxing Charitable Donation." Forthcoming Journal of Public Economics, 2007.

[5] Dana, Jason; Cain, Daylian M. and Dawes, Robyn M. "What You Don't Know Won't Hurt Me: Costly (but Quiet) Exit in Dictator Games." Organizational Behavior and Human Decision Processes, 2006, 100(2), pp. 193-201.

[6] Deci, Edward L. Intrinsic Motivation. New York: Plenum Press, 1975.

[7] Ellingsen, Tore and Johannesson, Magnus. "Generosity," Mimeo. Department of Economics, Stockholm School of Economics, 2007.

[8] Falk, Armin and Kosfeld, Michael. "The Hidden Cost of Control." American Economic Review, 2006, 96(5), pp. 1611-30.

[9] Fehr, Ernst and Falk, Armin. "Psychological Foundations of Incentives." European Economic Review, 2002, 46, pp. 287-724.

[10] Fehr, Ernst and List, John A. "The Hidden Costs and Returns of Incentives - Trust and Trustworthiness among Ceos." Journal of the European Economic Association, 2004, 2(5), pp. 743–71.

[11] Fehr, Ernst and Schmidt, Klaus. "Theories of Fairness and Reciprocity - Evidence and Economic Application," M. Dewatripont, L. P. Hansen and S. J. Turnovsky (Eds), Advances in Economics and Econometrics - 8th World Congress, Econometric Society Monographs. Cambridge: Cambridge University Press, pp. 2003, 208-57.

[12] Frey, Bruno S. Not Just for the Money. An Economic Theory of Personal Motivation. Cheltenham, UK and Brookfield, USA: Edward Elgar, 1997.

[13] Frey, Bruno S. and Oberholzer-Gee, Felix. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out." American Economic Review, 1997, 87(4), pp. 746-55.

[14] Frey, Bruno S. and Goette, Lorenz. "Does Pay Motivate Volunteers?," Mimeo. Institute for Empirical Research in Economics, University of Zurich, 1999.

[15] Frey, Bruno S. and Jegen, Reto. "Motivation Crowding Theory: A Survey of Empirical Evidence." Journal of Economic Surveys, 2001, 5(5), pp. 589-611.

[16] Glazer, Amihai and Konrad, Kai A. "A Signaling Explanation of Charity." American Economic Review, 1996, 86(4), pp. 1019-28.

[17] Gneezy, Uri. "The W Effect of Incentives." Mimeo. The University of Chicago Graduate School of Business, 2003.

[18] Gneezy, Uri and Rustichini, Aldo. "A Fine Is a Price." Journal of Legal Studies, 2000a, 29(1), pp. 1-18.

[19] ____. "Pay Enough or Don't Pay at All." Quarterly Journal of Economics, 2000b, 115(3), pp. 791-810.

[20] Harbaugh, William T. "The Prestige Motive for Making Charitable Transfers." American Economic Review, 1998a, 88(2), pp. 277-82.

[21] ____. "What Do Donations Buy? A Model of Philanthropy Based on Prestige and Warm Glow." Journal of Public Economics, 1998b, 67, pp. 169-284.

[22] Heyman, James and Ariely, Dan. "Effort for Payment: A Tale of Two Markets." Psychological Science, 2004, 15(11), pp. 787-93.

[23] Hollaender, H. "A Social Exchange Approach to Voluntary Cooperation." American Economic Review, 1990, 80(5), pp. 1157-67.

[24] Independent Sector. Giving & Volunteering in the United States. Mimeo. Independent Sector, 2001.

[25] Janssen, Maarten C.W. and Mendys-Kamphorst, Ewa. "The Price of a Price: On the Crowding out and in of Social Norms." Journal of Economic Behavior and Organization, 2004, 55(3), pp. 377–95.

[26] Meier, Stephan. "A Survey on Economic Theories and Field Evidence on Pro-Social Behavior," Bruno S. Frey and Alois Stutzer (Eds), Economics and Psychology: A Promising New Cross-Disciplinary Field. Cambridge: MIT Press, 2007, pp. 51-88.

[27] Mellstroem, Carl and Johannesson, Magnus. "Crowding out in Blood Donation: Was Titmuss Right?" Mimeo. Stockholm School of Economics, 2005.

[28] Rege, Mari and Telle, Kjetil. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations." Journal of Public Economics, 2004, 88(7-8), pp. 1625-44.

[29] Seabright, Paul. "Continuous Preferences Can Cause Discontinuous Choices: Application to the Impact of Incentives on Altruism," Mimeo. IDEI, Toulouse University, 2004.

[30] Soetevent, Adriaan R. "Anonymity in Giving in a Natural Context - an Economic Field Experiment in 30 Churches." Journal of Public Economics, 2005, 89(11-12), pp. 2301-23.

[31] Titmuss, Richard M. The Gift Relationship. London: Allen and Unwin, 1970.