————

Thèse de Doctorat (NR) de Sciences Economiques

Présentée et soutenue publiquement par

Luigi Butera

le 1 Juillet 2013

en vue de l'obtention du grade de docteur de l'Université de Lyon

délivré par l'Université Lumière Lyon 2

————

Three Essays on the Economics of Image Motivation

————

Directrice de Thèse :

Marie Claire Villeval - Directrice de recherche CNRS, Université de Lyon

Jury :

Lorenz Goette - Professeur, Université de Lausanne (Switzerland) - Rapporteur

Louis Levy-Garboua - Professeur, Université Paris I - Rapporteur

Daniel Houser - Professeur, George Mason University (USA)

Michael Kosfeld - Professeur, University of Frankfurt (Germany)

Acknowledgements

The essence of all beautiful art, all great art, is gratitude.

(Friedrich Nietzsche)

As with all things in life, this work would have been impossible without the visible and invisible hand of many people who taught me, encouraged me, comforted me, pushed me, and sometimes, of course, scolded me.

I am deeply indebted to my advisor and mentor Marie Claire Villeval, who provided me with invaluable guidance and constant support during these years. Marie Claire, with wisdom (and a lot of patience I shall admit!), taught me to aim high while keeping my feet on the ground, and challenged me to investigate human behavior with creativity, humbleness and rigor. I hope to be able to live up to Marie Claire's standards, both as a scholar and as a teacher. I know that I gave her quite a few "mauvais quart d'heures", and I could not be more grateful for the way she kept me in line while also "letting me go". Although I feel my training with her is far from being over, I wish to tell her from the bottom of my heart: thank you, not only you trained me to be a scholar, but you helped me to become a better human being.

The same gratitude, affection and respect goes to Daniel Houser, advisor and mentor for my current adventures and for hopefully many more to come. Dan has done something more than teaching me, he has protected me, valued me and trusted me. I have rarely met a person so connected with the noblest calling of a scholar: If I will ever become even 10% of the scholar he is, I'll call it a day. Besides, he is contagiously fun and eclectic, which makes working and studying with him even more enjoyable!

Thanks to Lorenz Goette and Louis Levy-Garboua, who have kindly accepted to be discussants of this thesis, and to Dan Houser and Michael Kosfeld who agreed to be members of the jury.

Quite a few professors and scholars helped, inspired, and guided me through this path. Among the many, I am grateful to Marco Castillo, Giorgio Coricelli, John Nye, Kevin McCabe, Ragan Petrie, Stephane Robin, and Thomas Stratmann. I am glad and excited you represent my reference work community, it gives me confidence that I have made the right professional choice.

This work would have not been possible without the amazing coauthors with whom I have had the honor to work with on these projects: Jason Aimone, Jean Claude Dreher, Jeffrey Horn, Elise Metereau, and Marie Claire Villeval. Thank you for being so wonderful, I am sure this is just the beginning!

The Gate Lyon-St Etienne and Lyon University have been a wonderful environment to work and live in. I am particularly grateful to Flavio, Isabelle, Julie 1, Julie 2, Nelly, Ranoua, Stephane, Tai, and of course, Zio Nando. Thank you for enriching my life.

In the Gate gang, one person in particular deserves full recognition of my appreciation and love: Bruno Crevat. Bruno hasn't just animated and illuminated my life at Gate and Lyon with his exquisite taste for life, friendship, and music, but he has become truly part of my family. You are a gentle soul, and your innate attitude for compassion and love will always inspire me.

I am greatly in debt to George Mason University, and ICES in particular, which has hosted me and has become the place I call home. I shall not dwell upon all the wonderful people I have the chance to work and live with, since they will soon enough get a chapter on their own.

In the eternal debate on whether nature or nurture makes a man who he is, my parents and family sit firmly and comfortably on the top of both. I had the privilege of being born into a family of extraordinary people: loving, righteous, knowledgeable

and imbibed with the finest sense of self-irony (I'm sure my dad, brother and sister will giggle here..!). The bias, I admit, is pretty evident, but I have never met people so capable of facing their own weaknesses and strengths to use them, as wise artisans, to build an ever new and better version of themselves. If there is an animal that captures the essence of my family, that would be the Phoenix.

To my parents who, as we speak, are still struggling with their metaphysical adolescence, I wish to say: *Grazie Mamma, Grazie Papa'*.

Friends. Well, what can I say? Ten pages for each of you would not give justice to the love and gratitude I feel for you. Whether you are in Italy, U.S., France, U.K, Germany or Switzerland, I want you to know you are my greatest strength, and ultimately my greatest achievement. You keep making me the richest person in the world.

But most of all, this work is dedicated to Giulia, the love of my life. Without her this challenge would have not been possible, and even if so, it would have borne the signature of a sadder and ultimately worse man. Thank you for letting me discover a new part of you every day. A new part of me, after all. We will keep changing, becoming many new persons in the years to come. I am looking forward to be there. For all of you.

Andiamo and see you at Trio tonight!

Washington D.C., April 29th 2013

viii

Contents

List of Tables

List of Figures

General Introduction

> *Behavior is the mirror in which everyone shows their image*
>
> (Johann Wolfgang von Goethe)

Social approbation and self-esteem represent a fundamental drive for human actions. The cognition of what is socially acceptable and what is not develops in children at very early stages in life, as well as the understanding of our own actions as a central definer of who we are (see i.e. Call and Tomasello 1998; Meltzoff 2002; Decety and Sommerville 2003). In any culture, at any given time in history, individuals have always shown tendency to ponder their actions according to how those would be perceived by others, engaging in self-sacrifice to maintain their reputation, honor, and respect. Not only people care about the opinion others hold of them, but also about what their own actions signal to themselves; behaving according to morals, sacrificing for a principle or ideal, acting "as if" we were always under judgment, are all mechanisms humans put in place to maintain their own self-image or self-esteem.

Economic decisions make no exception. Adam Smith was one of the firsts, and for a long time one of the only economists to formally recognize the importance of image concerns for economic exchange. In his Theory of Moral Sentiments (1759), Smith highlights people's desire to maintain and enhance their social-image, and points out how individuals naturally understand their own actions as the vehicle through which their peers and themselves can evaluate their virtues. In his own words:

> *"Man naturally desires, not only to be loved, but to be lovely; or to be that thing which is the natural and proper object of love. He naturally dreads, not only to be*

1

hated, but to be hateful; or to be that thing which is the natural and proper object of hatred. He desires, not only praise, but praise-worthiness; or to be that thing which, though it should be praised by nobody, is, however, the natural and proper object of praise. He dreads, not only blame, but blame-worthiness; or to be that thing which, though it should be blamed by nobody, is, however, the natural and proper object of blame". (TMS,Part II, Ch.2)

Smith's essential point was that while individuals struggle for praise and recognition, they understand that the very praise they are after has to be consistent with their own actions in order to provide *utility*. To paraphrase him, individuals are guided in their actions by (the metaphor of) an *impartial spectator* who, in a recursive and endogenous process, influences both our judgments of the actions of others and judgments of, and actions by, ourselves.

Despite the interest philosophers, historians, psychologists, and early economists as Smith have manifested towards the role of image in human decisions, economists have long disregarded it. The rationale for this disaffection is both philosophical and methodological.

From a philosophical perspective, the work of early economists as Adam Smith has been somehow misunderstood: as the Wealth of Nations became rapidly the reference text of Smith's thinking, his complex and multifaceted notion of self-interest has been translated into the more pragmatic concept of *monetary* self-interest: the *homo oeconomicus* was born. It is only in the last century that the economic profession has started to (re)expand the notion of self-interest to accommodate a wide range of behaviors that would be otherwise hard to explain using the homo oeconomicus paradigm: charitable giving, provision of public goods, volunteering, tax compliance, discrimination, authority acceptance within firms, efficiency wages, are all examples of inter and intra-personal strategic decisions that would be hard to explain by restricting the notion of interest to a pure monetary one.

In terms of methodology, the introduction of social cognition, self-image, and beliefs in the economics discourse has been made possible by last century's substantial innova-

tions in theoretic and mathematical tools, as well in empirical methods. On one hand the development of modern game theory and bayesian statistics expanded our ability to generate theoretic predictions about beliefs' formation and strategic interactions; on the empirical side, the revolution in modern statistics started by the work of Ronald Fisher and the introduction of experimentation in economics, made possible to generate and test hypotheses that would be otherwise hard to falsify using natural occurring data. Moreover, since fifteen years neuroeconomics, a new interdisciplinary field that combines economic experiments with neuroscience techniques, is expanding our knowledge of the way individuals evaluate and process monetary and non-monetary rewards. Such evidence is particularly important for economic decisions involving image concerns, as understanding how the brain computes image gains and losses in different situations will help unravel why some policies and practices aimed to encourage prosocial behavior work and others don't[1].

Understanding how individual decision-making problems are affected when there is an image component in the utility function is important not only from a philosophical perspective, but from a much more empirical one. Because pro-sociality is considered a desirable trait, public policies and private practices often generate incentives to encourage such decisions: donors are offered matching grant mechanisms, governments encourage charitable giving by allowing tax deductions, subsidize environment-friendly products (e.g. low emission vehicles), and use punishment mechanisms to deter anti-social behaviors like tax evasion or polluting habits. While such practices are intended to create personal (often monetary) incentives that offset personal costs suffered by pro-social individuals, they may sometimes backfire. Similarly, unintended consequences may arise as governments or private entities try to use social pressure to leverage pro-social behavior. Benabou and Tirole (2006) for instance, hypothesize that while some individuals are genuinely other-regarding, others view good deeds like charitable giving as investments in identity-building or demonstration goods. These investments are

---

[1]For a comprehensive review on the role of neuroscience in economic research see Camerer, Loewenstein and Prelec 2005.

made to establish or maintain social esteem, and to maintain their own view of what kind of person they are. For these individuals thus, the presence of extrinsic incentives or a greater visibility of actions, changes the meaning attached to prosocial (or antisocial) behavior, reducing the reputational incentive to engage in it. As in fact that incentive mechanism attracts more materialistic individuals, the signaling value of prosocial actions gets diluted.

Of similar importance is to understand how payoff-strategic decisions are affected by image or pride concerns. In the last twenty years, several empirical papers have given more and more ground to the idea that workers are not solely motivated by monetary interest, but see their work as a source of self-realization. As a consequence, providing monetary incentives or controlling subordinates may come with hidden costs. It has been shown that higher wages sometimes induce better performances (Fehr, Kirchsteiger and Riedl 1993; Gneezy and Rustichini 2000; Fehr and Goette 2007), and that stronger monetary incentives sometimes induce worse performances (Frey and Oberholzer-Gee 1997; Fehr and Rockenbach 2003; Fehr and List 2004). Similarly, the use of authority within organizations has important behavioral effects which would be hard to interpret without taking explicitly into account self-image or pride concerns (Falk and Kosfeld 2006). Principals retain their authority even when this choice is monetarily sub-optimal with respect to delegation (Fehr, Herz and Wilkening 2013), and subordinates oppose control for reasons other than monetary incentives (e.g. social distance, spite for being distrusted, illegitimate authority). Many of these puzzles can be explained by taking explicitly into account the fact that people care about their self-realization (Benabou and Tirole 2006), self-esteem, and pride (Ellingsen and Johannesson 2008).

The three essays presented here concentrates on two areas of decision making that have important connection with social and self image concerns: charitable giving and hierarchical relationships. The first essay investigates whether social pressure works in the same way for decisions that involve *doing good* and *avoiding doing bad*, studying both behavioral and neural responses to social pressure. The second essay answer the question of whether donors care about the quality of the recipients and explores the

4

relationship between *quality* and *quantity* of giving for donors motivated by prestige. The third essay investigates the effects of authority in hierarchical relationships characterized by aligned monetary interests, exploring how different ways of exercising control interact with self-image of both subordinates and principals.

Why individuals sacrifice personal resources in favor of someone else's wellbeing has been widely investigated in economics literature. As the traditional idea that people are motivated solely by pure altruism failed to explain several empirical observations about charitable giving, different theories have been called to explain private generosity. Following the notion that individuals may derive a direct utility from giving (see i.e. Becker 1974), internalized norms (Arrow 1971; North 1981; Levy-Garboua et al. 2006), social approval(Hollander 1990), warm-glow (Andreoni 1990; ; Ribar and Wilhelm 2002; Harbaugh, Mayr and Burghart 2007), conditional cooperation (e.g. Fischbacher, Gachter, and Fehr, 2001), reciprocity (e.g. Sugden, 1984), and prestige (Harbaugh 1998a, 1998b; Bracha, Heffetz and Vesterlund 2009), have been identified as powerful motivations for giving. In particular, the empirical evidence on the importance of social image and prestige is vast (See i.e. Vesterlund 2003; Andreoni and Petrie 2004; Rege and Telle 2004; Ariely, Bracha and Meier 2009; Shang and Croson 2009).

The first two essays contribute to this literature.

In the first essay we use functional magnetic resonance imaging to investigate the neural correlates of two types of pro-social decisions: *doing good* and *avoiding doing bad*. In particular, we investigate two types of situations where moral and monetary rewards are at odds: (1) people can incur a monetary loss to allow a donation to a positively evaluated organization (i.e. costly donation to a good association), and (2) can refuse to earn money in order to avoid a monetary transfer towards a negatively evaluated organization (costly oppose a negatively valued association).

Such neural evidence is particularly useful from a policy perspective as it helps understanding to what extent and why people are ready to trade off personal monetary losses with social image gains.

By systematically varying the price of giving and the degree of social visibility, we show that for these two types of decisions moral and monetary rewards are weighted and evaluated by different brain regions, depending on whether social image represents an issue or not. To do so we implement a 2x2 within-subject design, where individuals decide to accept or reject transfers to two different organizations, either in presence or absence of observers. While facing the positively evaluated charity, subjects decide whether to accept or not various donations made by the experimenter to the organization at a variable cost for themselves, taken out from their initial endowment. In trials involving the negatively evaluated charity instead, subjects decide whether to accept or not various potential gains for themselves, which are associated with various monetary transfers made by the experimenter to the anti-charity.

In our design thus, the only way to earn money was to let the experimenter transfer money to a negatively evaluated organization, while any transfer to a positively evaluated organization involved a monetary loss. This setup allows us to investigate whether, for these two types of decisions, moral and monetary rewards are weighted and evaluated in the brain in a similar way, and how social image concerns affect behavior and its neural correlates. Our results show that indeed doing good and avoiding doing bad involve partially different neural processes, and that the effect of social exposure depends both on the nature of the pro-social decision, and the payoffs involved.

We find that social image gains deriving from pro-social public decisions and monetary gains deriving from selfish private decisions do not share the same neural networks, and are partially different for the two decisions : while social image gains recruit for both decisions the Anterior Cyngulate Cortex and the Putamen (regions known to respectively monitor conflicts and compute reward-related events), monetary gains deriving from selfish "secret" decisions are dissociated in the Orbitofrontal Cortex (OFC): refusing to make a charitable contribution (and thus *saving* money) correlates with the right median OFC, a region connected to the appraisal of the reward value of reinforcers (e.g. money, food). Differently, accepting to *earn* money and allow a transfer to a negative association correlates with the lateral OFC, a region involved in the anticipation of punishment and

the feeling of moral disgust and reprehension. Moreover, we show that social exposure has little impact when doing bad involves low moral losses (i.e. low payoffs for the negative association).

Taken together, these results have important implications for economists: when the concern is with people doing good, our work shows that policies aimed to increase social pressure on individuals are effective when relatively high personal sacrifices generate even higher positive externalities. However, when the objective of a policy is to deter private behaviors that generate social negative externalities, our experiment suggests that the efficacy of "social dissuaders" will mostly depend on the magnitude of the damage created: in our experiment, social visibility deterred negative behavior only for decisions that would involve a high transfer to the negative association. For any other combination of personal gains and gains for the negative association instead, the effect of social exposure was negligible, even when perspective payoffs for the individual were low. On the neural side, we find that *not doing good* without image drawbacks recruit regions related to the anticipation of reward, while *doing bad* correlates with regions connected to the anticipation of punishment and the feeling of moral reprehension, even when being selfish has no negative image consequences. This gives us a glimpse into how different could be the effect of social exposure for different types of pro-social behaviors. We hypothesize that these results could be explained by different social norms underlying the two types of behaviors.

This may help explaining why it is hard to change people's habits when it comes to choices that, though harming for society at the aggregate level, are hard at an individual level to be thought as "despicable" (e.g. not recycling, crossing streets with red light, avoiding flu vaccinations, polluting).

While the first essay explores the relationship between image concerns and *quantity* of giving, the second essay investigates how giving is affected by information about the *quality* of giving, and how this relationship depends on image concerns.

In the last decade in fact, the amount of information available about the financial responsibility and efficiency of charities has increased dramatically. More and more

7

charity watchdogs are developing *homogeneous* measures aimed to help donors to compare *heterogeneous* charities and make better informed decisions. Examples of these non-profit watchdogs are Charity Navigator, Give Well, and the Urban Institute. The availability of these platforms has important consequences on the market for charitable giving, especially when it comes to small donors. While in fact large donors always had both the incentives and the means to gather this type of information, determining the worthiness of a particular charity required nontrivial research efforts for small donors: gathering appropriate government documents, phoning charities to request information, interviewing administrators and board members, or even serving on boards themselves. The emergence of these information clearinghouses however, gives now charities an incentive to report on their own, since being unrated or un-reviewed might look suspicious to donors. As these mechanisms both increase the amount of information readily available and reduce the cost of access, donors of all levels may now seek information at relatively little expense. While indeed some donors may choose to remain ignorant about these measures, the effect this information has on those that do encounter it is poorly understood.

The objective of the second essay is thus twofold: (i) investigate how new information about real charities' efficiency affects small donors' contributions and (ii) understand whether the public visibility of that information matters for donors, that is, how donors motivated by social image would take into account information that, although it can be ignored, it cannot be hidden from other people. Otherwise said, we investigate whether for small donors the size of the gift and the efficiency of the recipient are *complements* or *substitutes*, and how the response to information depends on that information being public or not.

To do so we implement a two-phase experiment, with the second phase unknown to the subjects in the first phase. In phase one, subjects freely choose three charities from a list of more than 5,000 charities rated by Charity Navigator, and play three independent dictator games with the selected charities. Subjects are also allowed to indicate a favorite charity among the three, a decision that increases the probability of

that specific charity being randomly selected for final payment[2]. While the first phase serves to assess the unconditional willingness-to-give of subjects, in the second phase we assess how decision made in the first phase are revised in light of new information about the charities. In phase two, subjects are incentivized to correctly guess what they believe is the true efficiency of their charities, and then receive privately the true efficiency values. We consider that a subject discovering that a charity is more efficient than expected receives good news, and that a subject discovering a charity is less efficient than expected receives bad news[3]. Finally, subjects are given the opportunity to revise their initial donation decisions and to indicate a new favorite, if desired. This means that participants can switch favorite charity across phases, or indicate a favorite in phase 2 while they did not in phase 1 (and vice versa). It's important to notice that decisions from phase 2, which are the only ones taken into account for payments, cannot influence those made in phase 1, as subjects during phase 1 are unaware of the existence of a subsequent and final phase. This also means that participants do not know whether others did change their decisions across phases. As we are interested in comparing responses to efficiency information under different levels of social exposure, we implement three between-subjects treatments: (T0) donation decisions and efficiency are private information, (T1) the donation to the final charity is publicly revealed but information about efficiency remains private, and (T2) both efficiency and donation amount of the randomly-implemented final decision are publicly revealed.

Our results show that both the type of information and the degree of its visibility have important consequences for donors' behavior. We find that as long as charity efficiency remains private information, individuals reward better-than-expected charities (good news) by increasing their donations. Differently, when charities are worse-than-expected, as far as giving happens under full anonymity, individuals disregard these bad news about their charities. However, when charity efficiency is revealed to others, some charities receive less although news were good, and some receive more despite news

---

[2]In fact, only one of the three decisions is randomly implemented for payments.
[3]We assume (and show in the essay) that individuals do care about this type of information.

were bad. This emergent behavior accounts for 30% of subjects that do respond to new information. We suggest that the latter behavior is attributable to image motivated donors treating efficiency and donation amounts as *substitutes* in terms of social image payoffs, reallocating their gifts in order to maintain the same expected image payoff.

Taken together, these results show two important effects of information in the market for giving: as far as information about ones' own charities' efficiency remains private (in the sense that an image motivated donor can announce his gifts without implicitly signaling how efficient the charity is), the diffusion of homogeneous measures would increase the size of the market for those charities that outperform in terms of efficiency. However, when information becomes bundled with charities' names, charities that are better than average in terms of efficiency may paradoxically loose part of their funding, as image motivated donors would trade off the *quantity* of their gifts with the *quality* of the recipients. Otherwise said, with common knowledge information, donors motivated by prestige would treat the quantity of their gifts and the quality of the recipients as *substitutes*.

While the first two essays focus on pro-social decisions and their connection with the value people attribute to their social image, the third essay explores the role self-image has in firm relationships. We investigate, in a firm-like environment, how different hierarchical institutions affect subordinates' willingness to comply with principals' formal authority. More precisely, we explore whether agents' and principals' use of real and formal authority depends on the signaling value of restrictions and delegation. The focus of the essay is on environments where principals and agents face aligned monetary interests, and compares institutions where principal's authority is exercised impersonally (e.g. rules apply to every subordinate, regardless of his behavior) with institutions where authority is exercised via direct monitoring (e.g. ad-hoc restrictions whenever deemed necessary).

The objective of this study is to understand whether hidden costs of control do arise even in cases where principal's exercise of control does not conflict with agent's payoff. We hypothesized that, if agents have preferences over the signaling value of principal's

10

exercise of authority, they should then oppose restrictions more fiercely whenever they are personal, as they signal the principal disagrees with their choices, and that he is directly taking actions to limit their real authority.

Traditional agency theory in fact predicts that when agents' and principals' monetary interests are not aligned, agents will try to act in their own interest, making thus the exercise of control a possible way for principals to limit agents' opportunistic decisions. Recent empirical findings however have pointed out that for principals, controlling subordinates may come with hidden costs. Agents who feel distrusted (Falk and Kosfeld 2006), or perceive principal's authority as illegitimate (Schneidler 2010), appear to punish control by providing less effort they would if principals had delegated them. More generally, social distance (Frey 1993; Dickinson and Villeval 2008), intentions behind principals' decisions (Charness and Levine 2007), and punishment's threats (Fehr and List 2003), represent important factors determining the outcome of control.

In this essay we test the hypothesis that control may come with hidden costs even when principal and agent monetary interests are *aligned*. We introduce a new principal-agent-client investment game where control takes the form of limiting agent's reciprocity towards a third party (i.e. an investor, or client) who entrusts the agent: the principal can exert formal authority on the agent (e.g. provides an upper limit to agent's trustworthiness towards the client), while the agent, the trustee, retains real authority (e.g. implements the effective decision within the allowed boundaries). Since whatever is not sent back to the client is split evenly between the agent and the principals, they have aligned monetary interests. By varying principal's monitoring ability, we assess how agents' trustworthiness towards their client depends on whether principals' exercise of formal authority represents a *personal restriction* to agents' actions or an *impersonal rule* (applied ex-ante to all type of agents) that merely signals principality preferences.

Our results show that also with aligned monetary interests, direct control does involve hidden costs: restricted agents are relatively more generous with the client when the restriction is *personal* (principal informed) compared to when the restriction is *impersonal* (principal uninformed). The latter result provides evidence that, even with

11

aligned monetary interests, hidden costs of control may exist whenever restrictions represent a direct opposition to agents' actions, and that these costs can be mitigated by exercising control via impersonal rules. More explicitly, agents appear to be willing to costly punish the principal (by sending more to the client) whenever the restriction imposed by the latter represents an ad hoc or intentional restriction (e.g. I restrict you because *you were* doing wrong); when restrictions instead only signal principal's preferences (e.g. I restrict you because *I am* greedy), agents significantly reduce their trustworthiness towards their client.

Surprisingly however, despite ex-post relatively higher payoffs, we find that principals tend to recur to formal authority relatively less often under an impersonal rule system. We speculate that principals may be reluctant to be "first movers" on unfair decisions, and prefer to signal their greed only when strictly necessary: if in fact principals have preferences for appearing fair and trustworthy (although not necessarily for being fair and trustworthy)(see i.e. Dana, Weber, Kuang 2007; Hao and Houser 2011), then personal restrictions allow them to signal their greed only when necessary (e.g. when the agent is too generous); on the other hand imposing a restricting impersonal rule signals their type (e.g. untrustworthy) also to agents that "didn't need to know", as they are already not overly generous. This result appears to be in line with recent studies on delegation and blame aversion under principal-agent aligned monetary interests. Bartling and Fischbacher (2012) and Grossman (2012) have shown how delegation can be an effective instrument principals can use to shift the blame (and punishment) on their employees for unfair outcomes received by third parties. This essay complements this literature by showing how delegation of unfair choices may be important for principals even when they do not face the threat of being punished by clients, and when their accountability in face of latter cannot be manipulated.

Essay 1

Audience effect on two types of pro-social behavior: an fMRI Investigation[1]

1.1   Introduction

In everyday life people are often called to sacrifice resources to improve the well-being and wealth of others, squaring with the temptation of indulging over one's own private benefit.

Costly actions that benefit others come essentially in two forms: people can spend resources to directly increase the well-being of others, or can renounce to something that benefits themselves but generates a negative externality. Charitable giving represents a good example of the first type of pro-social action; tax compliance, refusing bribes, recycling, or not littering are all examples of the second type. While these decisions are sometimes taken in private, away from other people's eyes, they are often actions that, being visible to others, they contribute to establish and maintain a positive *social image*.

---

[1]This chapter is a joint work with Jean Claude Dreher (principal investigator), Elise Metereau (first author) and Marie Claire Villeval (third author).

Recent economic theories (see i.e. Harbaugh 1998a, 1998b; Benabou and Tirole 2006) and empirical evidence (e.g. Andreoni and Petrie 2004;Ariely, Bracha and Meier 2009; Grossman 2010) have in fact suggested that human decision making may be seen as a combination of altruism (either pure or impure), self-interest, and concerns for social reputation: humans may exhibit preferences for pro-social behavior not (just) because they are intrinsically good but also because of concerns for their image.

An important consequence is that monetary incentives provided to encourage such behaviors, like matching grants or tax deductions, may not produce the desired effect.

Benabou and Tirole (2006) for instance, hypothesized that while some individuals are genuinely other-regarding, others view good deeds like charitable giving as investments in identity-building or demonstration goods. As these investments are made to establish or maintain social esteem, the presence of extrinsic incentives may change the meaning attached to pro-social behavior, reducing the reputation incentive to engage in it.

On the empirical side, Ariely, Bracha and Meier (2009) have shown that people tend to behave more pro-socially when their decisions are visible to others, and that monetary incentives are more effective in private than in public[2], that being true also for actions that are subjectively perceived as morally inappropriate[3]. In a similar vein, Della Vigna, List and Malmendier (2009) find that social pressure is an important determinant of door-to-door giving, suggesting that people may donate not only because they like to, but also because they dislike saying no when hard-pushed. In absence of social motivators, monetary incentives that reduce the cost of giving appear to work, though imperfectly: Karlan and List (2007) find that matching grants increase willingness to give; however, if a one-for-one match significantly increases contributions, higher matches do not.

While evidence thus exists on the interaction between personal monetary incentives and social image for decisions involving *doing good*, less however is known about its effect on pro-social decisions that involve *avoiding doing bad*[4].

---

[2]See also Goette, Stutzer and Zender 2011.
[3]e.g. supporting the National Rifle Association.
[4]One exception, again, is Ariely et al. 2009.

More importantly, for both types of dilemmas, it is unclear how social image affects pro-social decisions under different combinations of monetary and non-monetary stakes. In real life situations in fact, the benefits and costs a person may face by acting pro-socially are different: while in some cases the personal costs of being pro-social are low compared to the social image benefits one may get by being such, in others high image benefits come with relatively high personal costs.

The effect of image concerns on pro-social decisions has been studied in the context of *constant costs of giving*[5].

This essay builds on that literature and explores how actions that involve *doing good* and *avoiding doing bad* are influenced by social image, and how the latter influences dilemmas that involve different combinations of monetary rewards and more abstract rewards, such as prestige or warm glow.

Two factors in fact may affect individuals' pro-sociality under public exposure: the relative cost of being pro-social, and the absolute cost (and absolute positive externality) of being such. We test thus the hypothesis that, for both types of decisions, public exposure would increase the likelihood of pro-social decisions when the relative cost of being such decreases, and that the same likelihood would decrease when the absolute cost of generosity increases. Otherwise said, we test the extent to which the effect of social exposure on giving is a function of the relative cost of giving and the absolute monetary and abstract stakes involved by the decision.

To do so we implement a 2x2 within-subject design, where individuals decide to accept or reject transfers to two different non-profit organizations, one positively and one negatively valued, either in presence or absence of observers[6]. While facing the positively evaluated charity, subjects decide whether to accept or not various donations made by the experimenter to the organization at a variable cost for themselves, taken out

---

[5]Karlan and List (2007) for instance explore willingness to give under different price lists, but their study doesn't involve public exposure. Ariely, Bracha and Meier (2009) seminal work tested the interaction between social visibility and monetary incentives using real effort tasks. In their design thus the (monetary) cost of giving is zero.

[6]We used Resto du Coeur as the positively valued charity, a French charity that provides free meals for homeless. The negatively evaluated organization was Groupe d'Action Royaliste, a non-profit group that lobbies for the reintroduction of monarchy in France.

from their initial endowment. In decisions involving the negatively evaluated charity instead, subjects choose whether to accept or not various potential gains for themselves, which are associated with various monetary transfers made by the experimenter to the negative association.

Subjects' brain activity is tracked during the experiment using functional magnetic resonance imagery (hereinafter fMRI). If in fact much can be learned from observed behavior, neural activity may help to shed light on the reason why social exposure is documented to increase giving, and explain how abstract and monetary payoffs are weighted in the brain when people have to choose between money and "moral acts". Such evidence would be particularly important for economic decisions that involve image concerns, as understanding how the brain computes image gains and losses in different situations would help to unravel why some policies and practices aimed to encourage pro-social behavior work and others don't. While in fact social exposure is known to have a sizable impact on people's willingness to perform good deeds, the evidence of its role in preventing behaviors that benefit the individual at a societal cost is less systematic. We hypothesize that the way the brain computes monetary payoffs and more abstract ones may differ depending on the nature of the underlying pro-social decision, as these decisions may differ in terms of personal values and social norms people perceive to be attached to them.

Here thus we use functional magnetic resonance to investigate how image motivation modulates cerebral activity in presence of two types of sacrifices: incurring a monetary cost to allow a donation to a positively evaluated organization, and refusing to earn money to avoid a monetary transfer to a negatively evaluated organization. This design allows us to investigate (i) how the moral and monetary stakes involved by these two different types of transfers are computed in the brain, and (ii) whether social image gains deriving from pro-social public decisions share or not the same neural networks with monetary gains earned through selfish private actions.

We do indeed find patterns of behavior consistent with previous evidence: people tend to behave more pro-socially under public exposure. However we find that the re-

sponse varies with the possible combinations of abstract and monetary payoffs available for the two types of decision.

For decisions concerning the negatively evaluated association, we find that social pressure changes individual behavior only when social costs associated to a selfish action are relatively high (i.e. high payoffs for the negatively evaluated association), independently though from the personal gain an individual may get. This suggests that people may refrain from earning money that generates a negative externality only when the latter is sufficiently high, but that social pressure is ineffective in limiting selfish though relatively "harmless" decisions. Differently, for donations to a positive charity, we find that people take different decisions in private and public only when the conflict between personal monetary losses and societal gains is, in our framework, the highest (e.g. high individual costs, high social benefits); this means that for decisions involving a *similar cost of giving* but lower stakes, people behave in public and private in the same way.

Taken together, our results show that social image works differently, depending on weather being pro-social means to actively give away resources for a good cause, or renouncing to earnings that involve negative social consequences.

Our magnetic resonance results sheds light on the neural mechanisms underlying these two types of decisions, and their sensitivity to image concerns.

If in fact a large body of behavioral and economic evidence consistent with social motivation theories exists, there is still no clear consensus on how the brain computes costs and benefits associated with moral and monetary trade-offs. Moreover, whether social exposure causes similar or different neural networks to be involved in the weighting process of these different types of rewards is still poorly understood.

Indeed recent studies have shown how brain regions normally implicated in reward processing of money, food and drinking are similarly activated when people decide to donate to a charitable organization (warm-glow), as well as when these transfers are mandatory instead of voluntary (pure altruism) (Moll et al., 2006; Harbaugh et al., 2007; Hare et al. 2010). The values that people attach to payoffs involved in these decisions

appear to be represented in the ventromedial prefrontal cortex (vmPFC) when decisions concern primary, monetary or social rewards (Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006; Kable and Glimcher, 2007; Plassmann et al., 2007; Hare et al., 2008; Chib et al., 2009; Hare et al., 2009; Berns et al. 2012).

Consistently with these studies we find that vmPFC activity correlates with the gains the good association may get in case of acceptance, supporting the notion that social and abstract rewards share similar mechanisms with primary rewards; on the other hand we find activity in the Insula for trials concerning the negative association, a region part of a network known to superintend social cognition and emotions' processing [7].

This may suggest that different evaluation mechanisms may be in place for the two types of decisions.

We find further evidence of different neural networks involved in *doing good* and *avoiding doing bad* when we compare pro-social public decisions (which generate positive image payoffs at a personal cost) and selfish "secret" ones (which create monetary payoffs without social image losses).

Pro-social decisions made in public correlate for both associations with activity in the ACC and Putamen, two regions known to respectively monitor conflicts and compute reward-related events.

Different processes however appear to be involved when people make selfish decisions in private: when people *earn* money and allow a transfer to a negative association, we find activation in the lateral OFC, a region related to the evaluation of punishment and the processing of moral emotions like disgust and indignation. On the other hand, *saving* money by not allowing a transfer to a positive charity recruits the median OFC, a region connected to the appraisal of rewards, such as food or money.

These results support the hypothesis of a functional dissociation in frontal and striatal areas with respect to the value of monetary and non-monetary payoffs, and show how payoffs associated with selfish and pro-social decisions are computed differently,

---

[7]For a comprehensive review of reward and decision making processes in the brain see Dreher and Tremblay 2009.

18

depending on the salience of social image. Although in fact the neural mechanisms underlying pro-social public behavior are the same for the two associations, we do find that selfish decisions made under the guarantee of anonymity show different patterns: if *saving* money by "not doing good" when unseen by others correlates with a region associated with reward anticipation, *earning* money in secrecy by "doing bad" is instead associated with activation in lateral OFC, a region known to correlate with the evaluation of punishment and moral disgust.

A possible explanation is that internalized social norms (e.g. Arrow 1971; North 1981) may be different for the two decisions. As in France and most European countries, direct charitable giving does not represent a widespread practice as it is, for instance, in the United States, refusing to make an anonymous donation may not interfere with the social and moral norms embedded in that environment, allowing thus subjects to anticipate a reward when money is spared without norms' violation. Differently, a negative sentiment against monarchic institutions is highly embedded in French culture: earning money by allowing a transfer to an association that lobbies for it may be subjectively perceived as a violation of a social and moral norm, generating thus an anticipation of disgust and punishment.

A second hypothesis relies on the idea that the brain would process differently selfish decisions that have a "direct" negative consequence (e.g. "by earning money I am doing bad"), and selfish decisions that have an "indirect" negative consequence (e.g. "by saving money I am not doing good, but I am not doing bad either").

Although our design is not intended to disentangle between the two theories, both point to a substantial difference between *doing good* and *not doing bad*.

Finally we look at the neural correlates for different combinations of payoffs for the individual and the charity, and compare the effect of image exposure for the two types of decisions. For combinations of monetary and moral payoffs for which public exposure alters behavior, we find that the Inferior Parietal Cortex and Frontopolar Cortex are more active during private decisions than public ones. The known role of these regions in processing individual sense of agency suggests that individuals may feel "forced" to

19

put themselves in someone elses' shoes when choosing in public. This intuition appears corroborated when we look at neural activity in the counterfactual situation: two regions part of what is known as the "menthalizing network" are in fact more active for decisions taken in public than in private, the Insula and the Superior Temporal Gyrus.

Overall, our results show that the effect social exposure has on decisions involving *doing good* and *avoiding doing bad* is partially different for the two decisions, both at a behavioral and neural level. For decisions that involve doing good, social exposure alters behavior when the combination of personal costs and abstract rewards involves "hard" dilemmas, that is, when relatively high individual expenses allow even higher positive social consequences. Differently, social exposure for decisions that involve avoiding doing bad appears to be mostly a function of the magnitude of the negative externality generated by a selfish decision (e.g. money going to the negative association). At a neural level, we find that costs and (image) benefits deriving from public pro-social decisions are computed similarly in the brain, suggesting that both *doing good* and *avoiding doing bad* under social exposure are processed by the brain as rewards. On the other hand, we find that *not doing good* (and thus *saving* money) under secrecy recruits brains regions associated with reward, while *doing bad* (and thus *earning* money) under secrecy correlates with activity in regions known to superintend moral disgust and anticipation of punishment. We argue that those differences could be related to the different social norms associated with these types of pro-social behaviors, or to differences in terms of how "consequential" is the damage created by a selfish decision. Both hypotheses suggest that while the brain computes social image gains from pro-social choices somehow independently from the nature of these choices, monetary gains deriving from selfish (secret) decisions are instead weighted in the brain according to how these decisions violate or not a social norm or an "ethic" framework.

The remainder of the essay is organized as follows: section 1.2 presents the design; section 1.3 describes behavioral results, section 1.4 discusses imaging results, and section 3.5 concludes.

1.2 Experiment Design

1.2.1 Subjects

Seventeen healthy subjects (all men, age 22.47±2.62 years) with no history of neurological or psychiatric illness participated in a simple decision making experiment performed in a fMRI machine. All subjects were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and presented no symptoms of depression as assessed by the 13-item version of the Beck Depression Inventory (Beck and Beck, 1972). All subjects gave written informed consent to be part of the experiment, which was approved by the local ethics committee (CPP Centre Léon Bérard).

1.2.2 Pre-testing and recruitment procedure

Prior to the fMRI experiment, we implemented an incentivized behavioral study involving 48 healthy volunteers to help designing stimuli and task procedures. To guide the selection of the two organizations, we asked participants to complete a questionnaire that presented brief descriptions and logos of fourteen non-profit organizations. Both organizations with a commonly recognized positive and negative valence were presented. For each one, participants were asked to rate their feelings towards them on a scale from -10 to 10. The order of the organizations in the questionnaire was randomized for each subject, and in each session of the pre-test participants then played the task described in the next section with two random organizations (one positive, one negative) from the list of fourteen presented in the questionnaire. Participants from the pre-test earned on average 15.94 Euro [8]. The two organizations that received the worst (mean = -5.73, SD = 3.68) and the best (mean = 8.40, SD = 2.04) feeling rates in all sessions were chosen to be used in the fMRI experiment. The two associations are respectively Resto du Coeur, a French charity that provides free meals for homeless, and "Groupe d'action Royaliste", an organization that lobbies for the restoration of monarchy in France.

---

[8]We run a total of three sessions for the pre-test. Average earnings were respectively 19.3, 13.77, 14.75 Euro

The follow-up fMRI experiment was advertised within the Lyon University campus, and new participants were recruited using a two-steps procedure. In a first interview, perspective participants were asked to rate their feelings towards each of fourteen organizations we used in the pre-test on a scale ranging from -5 to 5. We selected and re-invited only subjects that rated Resto du Coeur with a score superior to zero, and Groupe d'Action Royaliste with a negative score, that is, subjects that expressed a negative view of Groupe d'Action Royaliste and a positive one of Resto du Coeur.

1.2.3 Experimental Task

In our 2x2 within-subject design, participants decided to accept or reject transfers to two different organizations (one negatively and one positively evaluated), either in presence or absence of observers (Fig. 1A). At the beginning of the experiment, subjects received an initial endowment of 14 Euros, and were told they would receive a check of 50 Euro in the following days by mail. During the experiment, subjects were faced with a series of offers regarding either "Resto du Coeur", the positively evaluated charity, or "Groupe d'Action royaliste", the negatively evaluated non-profit. For the positive charity, subjects had to choose between accepting or rejecting various donations made by the experimenter to the charity at a variable cost for them, taken out from their initial endowment. For the negative association, subjects had to decide whether to accept or not various potential gains for themselves, associated with various monetary transfers made by the experimenter to the negative organization. In our design thus, the only way to earn money was to let the experimenter transfer money to a negatively evaluated organization, while any transfer to a positively evaluated organization involved a monetary loss.

Potential transfers to the organizations, and potential costs/earnings for the subjects were varied independently across trials. In each trial, the organization potential gains ranged from 4 to 32 Euro, by increases of 4 Euro. Subjects' potential payoffs (in the case of the negatively evaluated organization) or costs (in the case of the positively evaluated

organization) varied from 1 to 8 Euro, by increases of 1 Euro[9]. This manipulation resulted in 64 different dilemmas (Fig. 1B).

This simple setup was implemented under two visibility conditions: a *Private* treatment, and a *public* treatment. In private trials, subjects' decisions were guaranteed to be anonymous, while in public trials, decisions to accept or reject were visible to observers witnessing the experiment from the fMRI's control room. To induce either anonymity or public visibility, subjects were shown that computer screens in the control room would display the same images they would see once in the fMRI.

As shown in figure 1C, each trial began with the presentation of an offer that the subject could either accept or reject by pressing two buttons they had below their right hand while in the fMRI. To induce visibility during public trials, the chosen alternative (accept or reject) was highlighted on screen for 1.5 seconds, while the other option disappeared: this means that in public trials, observers from the control room were able to see what option was chosen, and this was common knowledge among participants. In private trials instead, no changes were shown on screen after the response, making impossible for anyone (including the subject) to infer participants' decisions from the screen. Visual stimuli were back-projected on a screen located at the head of the scanner bed and presented to the participants through an adjustable mirror located above their head. The presentation of the stimuli were controlled by Presentation © software (Neurobehavioral Systems), which also recorded trigger pulses from the scanner signaling the beginning of each volume acquisition.

For both organizations, subjects took a decision over each possible combination of personal's and charity's payoffs twice, once in private, and once in public. Our participants thus took 256 decisions in total, 128 related to the negatively evaluated organization and 128 related to the positively evaluated one. In each category, half of the decisions were made in public and the other half in private.

---

[9]Note that subjects could donate at maximum 50% of their initial endowment of 14 Euro.
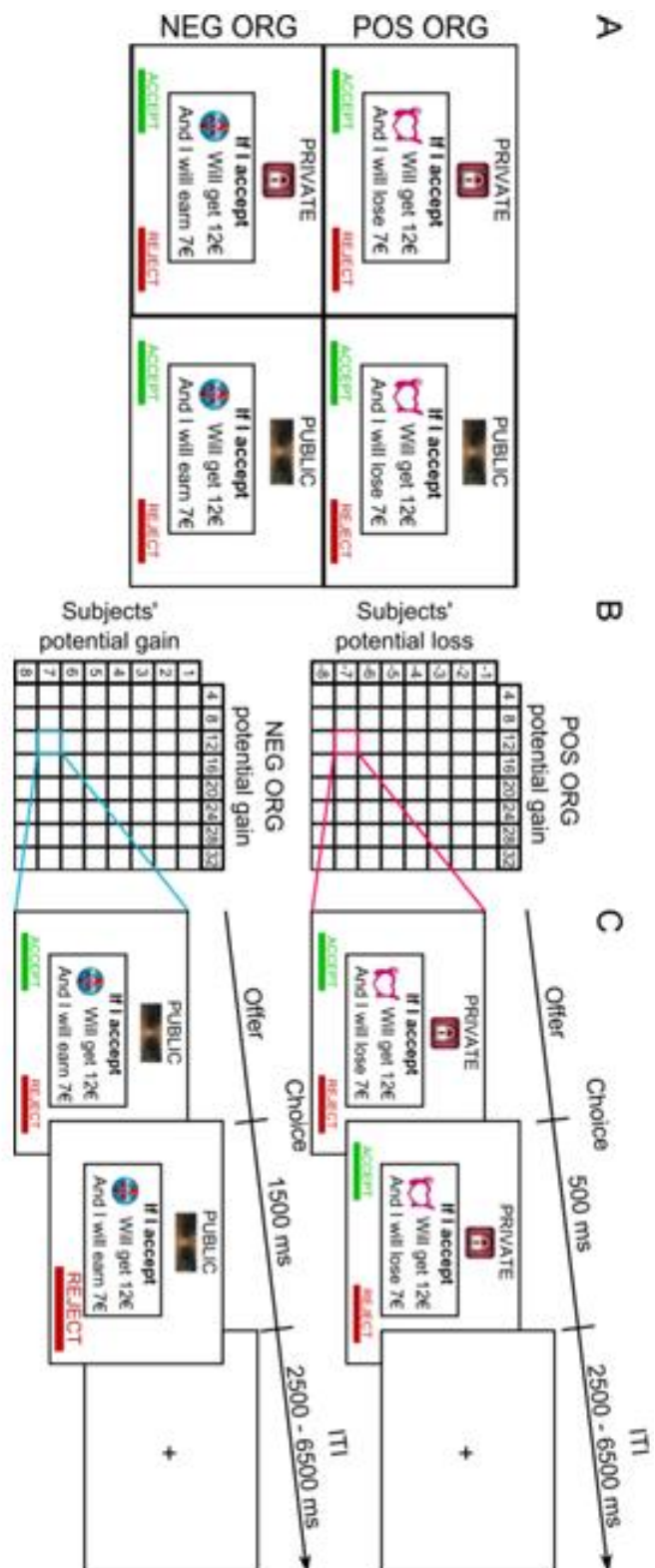
Figure 1.1: Game Description

At the end of every trial, a fixation cross was displayed for a random number of seconds, drawn from a uniform distribution between 2.5 and 6.5 seconds. This was done to ensure better sampling of the hemodynamic response, and to avoid sleepiness that may arise from a monotonous pace. Subjects were encouraged to make their decision within 3 seconds, although we emphasized in the instructions that the accuracy of their responses was more important than their speed.

The scanning session was divided into four blocks of 64 trials, each corresponding to one of the 2x2 conditions [10]. Within each block, the order of presentation of different combinations of payoffs for the subject and the association was fully randomized. The order of presentation of the four possible combinations was instead semi-randomized: while we randomized the order of presentation of the type of association, we let subjects play first the two conditions (private/public) for one association, and then the two for the other. For each type of association, we randomized the order of presentation of the two private and public blocks. We chose not to mix the organizations' type and the visibility condition within each block as pre-testing results shown that this required an excessive cognitive effort to readjust at every trial to new conditions over two different dimensions. The order of presentation of the organization and of the public/private condition was counterbalanced across subjects to control for potential order effects.

Only one decision from the public condition and one decision from the private condition were randomly drawn and implemented at the end of the experiment[11]. While only the public decision was drawn in front of the experimenters, both were sent automatically to the person in charge of paying subjects. To further increase the salience of the public conditions, subjects were informed at the beginning of the experiment that they would be videotaped while announcing what was the randomly drawn "public" offer, and what decision they took for that offer. No-one complained about nor refused the videotape procedure, and all participants signed a separate consent form.

---

[10]e.g. Good/private, Good/public, Bad/private, Bad/public
[11]This means that the two randomly selected trials could both be about the positive association or the negative one.

1.3   Behavioral Results

Our results are organized as follows: first we present the general analysis on the main treatment effects for both associations. Second we show which combinations of abstract and monetary rewards are driving differences between public and private treatments. Third we discuss how neuroimaging results can help us to interpret our behavioral evidence.

Result 1: *For any level of personal costs (gains), the higher the transfer to the good (bad) association from the experimenter, the higher (lower) the fraction of accepted offers.*
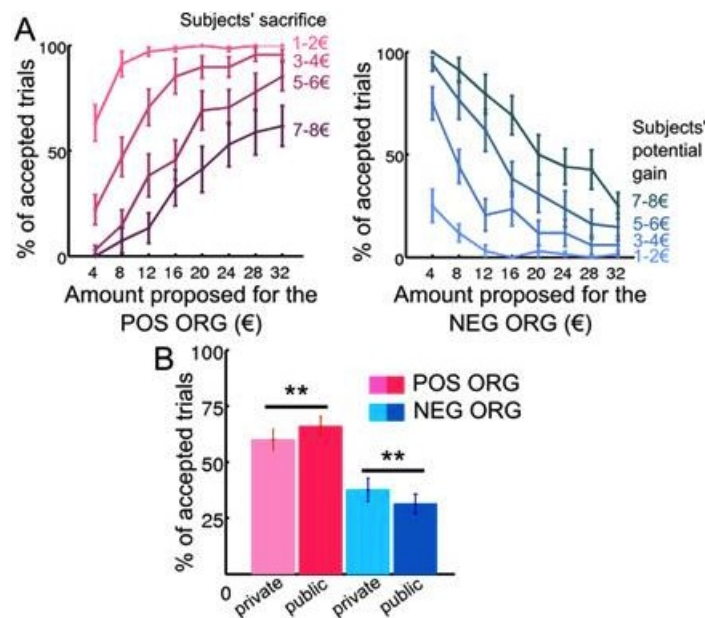
Note: *=0.05; **=0.01; ***=0.001

Figure 1.2: Acceptance rates

Part A of figure 1.2 graphs the percentage of accepted trials for the two associations for all possible combinations of association's and individual's payoffs. This figure shows a first general result: for the positive association and for each level of individual costs, the higher the associated payoff for the association is, the higher the acceptance rate.

Similarly, for all potential gains in the bad association trials, the higher the payoff of the association, the lower the acceptance rate. Differences in acceptance rates with respect to individual's payoffs tend to be smaller the higher the association's payoff is. This first result confirms that the associations we presented to subjects were truly perceived as good and bad respectively, and shows how aggregate preferences over different bundles are well behaved. Subjects earned on average 64,27 Euro (50 Euro fix show up fee plus average earnings of 14.27 Euro).

Result 2: *On aggregate, public visibility increases the fraction of accepted (rejected) offers for the positive (negative) organization.*

Part B of figure 1.2 shows results from paired t-tests on the effect public exposure has on acceptance rates. At an aggregate level people accept significantly more to donate money when decisions are taken in public, and accept less to earn money in the bad association trials under public exposure. This suggests that at a general level, social image matters for both types of decisions.

Table 1 shows similar results from two random effect probit regressions with cluster-robust s.e., separated by the type of association (marginal effects are shown).

We use the following specification:

$$Accept_{ik} = \alpha + \beta_1 \cdot \pi(Asso)_{ik} + \beta_2 \cdot \pi(Subject)_{ik} + \beta_3 \cdot PUBLIC_{ik} + \beta_4 \cdot TIME_{ik}$$

$$+\beta_5 \cdot GOODFIRST_{ik} + \epsilon_{ik}$$

Where the dependent variable is the probability for individual *i* of accepting an offer *k*, regressed on the payoff for the association, the gain or loss for the subject, a dummy for the treatment (PUBLIC=0 is our private treatment), the number of the trial (time), and whether the subject encountered first offers concerning the positive association (GOODFIRST=1) or the negative association (GOODFIRST=0).

Table 1.1: Random effect Probits (marginal effects shown)

	Resto (1)	GAR (1)	Resto (2)	GAR (2)
DEP. VARIABLE	accept	accept	accept	accept
asso_gain	0.0397***	-0.0326***	0.0397***	-0.0326***
	(0.0116)	(0.0106)	(0.0116)	(0.00865)
subject_payoff	-0.179***	0.146***	-0.179***	0.145***
	(0.0522)	(0.0470)	(0.0521)	(0.0385)
public_treatment	0.116***	-0.101***	0.116***	-0.101***
	(0.0402)	(0.0377)	(0.0401)	(0.0329)
time	7.15e-05	-0.000595*	6.65e-05	-0.000547*
	(0.000301)	(0.000323)	(0.000302)	(0.000295)
good_first			0.275	0.252***
			(0.214)	(0.0947)
Observations	2.176	2.176	2.176	2.176
Number of id	17	17	17	17

For the good association, an increase in the payoff the association may receive in case of acceptance, raises the probability of accepting the offer by 3.9 % (p=0.001), while an increase in the cost of the donation reduces the probability by 17.9% (p=0.001). Deciding under public exposure raises the probability of accepting the donation by 11.6% (p=0.001). Finally, for decisions concerning the good association, we do not find evidence that the order of presentation of the associations matters.

For the bad association, as one may expect, results are reversed: increases in the gains for association reduce by 3.2% (p=0.001) the probability of accepting the offer, while an increase in subject's gains raises it by 14.5% (p=0.001). The probability of accepting is significantly reduced (10.1%, p=0.001) when decisions have to be taken in public. Differently from good associations' trials, we do observe that having taken decisions over the good association first, increases significantly and substantially the probability of accepting.

Although this last result does not control for which decisions are effectively taken with the (previously played) good association, it evokes the idea that people may feel more entitled to act selfishly (i.e. accepting to earn money and allow a donation to the

bad association) whenever they have been generous in another context. The very same does not seem to apply for decisions regarding the good association, which are invariant to the order of associations.

So far we have seen that our public treatment had a significant effect on individual willingness-to-give. Our parametric design however, allows us to say more about which combinations of monetary and abstract rewards are driving the latter result. To explore how public exposure affects choices involving different combinations of payoffs, we divide payoffs for the association and for the individual in two classes: for individual's payoffs we categorize as low payoffs those ranging from 1 to 4 Euro, and as high all payoffs above (5 to 8 Euro). Similarly, we divide associations' payoffs in low (4 to 16 Euro) and high (20 to 32 Euro)[12].

Result 3: *Social exposure's effect on behavior is different for the two decisions.*

By taking into account the relative magnitude of the stakes involved, we find that social exposure has different effects on the two decisions different patterns for doing good and avoiding doing bad: paired t-tests show that for the good association, the only combination of abstract and monetary rewards that produces significant differences between private and public trials is when both individual's losses and association's gains are high (p=0.063)[13] (e.g. cost to for the subject greater than 4 Euro, gains for Resto du Coeur greater than 16 Euro).

For all other combinations of monetary and abstract payoffs, people do not behave differently in private and public. Interestingly, we do not find evidence of a systematic role of the cost of giving in shaping public and private pro-social decisions: for decisions involving similar ratios of individual's and association's payoffs, but lower stakes, social exposure appears not to alter behavior [14].

For the negative association instead, we observe that social exposure has a positive significant impact on pro-social decision making only when the benefits the negative

---

[12]We chose the median payoff as cutoff point. When using cutoff points shifted one position to the left or to the right, however, results don't change.

[13]Unless mentioned otherwise, all p-values presented come from two-tailed t-tests.

[14]By testing only trials with low-low payments and high-high payments with exactly the same cost of giving (element on the diagonal of the matrix of payments) results are substantially the same.

association can get are high (above 16 Euro) (p=0.007 when individual payoff is high (strictly above 4 Euro); p=0.023 when individual payoff is low (below 4 Euro)).

Although the general acceptance rate is lower the lower personal gains are, these results suggest that policies aimed to discourage people from *doing bad* through social exposure may at best be ineffective when the negative externality generated is relatively negligible.

In this section we have shown that for certain combinations of abstract and monetary payoffs, social visibility encourages individuals to make pro-social decisions which, although being monetarily costly for them, they generate abstract rewards such as warm glow or prestige.

The next section presents our neural results and explores how abstract rewards deriving from being pro-social are computed in the brain compared to monetary rewards deriving from money-maximizing decisions . The objective is twofold: understand whether the sense of reward that may derive from being pro-social is processed in the brain similarly to monetary rewards, and understand how this processing is modulated by social image concerns.

1.4 Imaging Results

1.4.1 Pro-social and selfish decisions

Our results are organized as follows: this section presents evidence on which regions do correlate with pro-social and selfish decisions, independently from whether decisions are taken in private or in public. Section 1.4.1.1 compares pro-social decisions made in public and selfish decisions made in private. Finally, section 1.4.2 compares decisions taken in private with decisions taken in public, regardless of the type of decision (e.g. pro-social or selfish)[15].

---

[15] All data from neural activity are recorded at the moment when the proposition is presented to subjects.

Result 4: *For the positive association, pro-social decisions correlate with activity in the ventromedial prefrontal cortex (vmPFC). For the negative, pro-social decisions correlate with the Insula.*

By looking at which brain regions correlate with pro-social decisions and with variations in abstract and monetary payoffs, we find that the vmPFC is more activated when people take the pro-social decision with the positive association (p<0.05) (sacrificing resources to let a good association earn money), compared to the case when the selfish option is chosen (refusing to lose money to make a benefit to the charitable association). This result is in line with previous studies on the role vmPFC has in moral reasoning (Greene et al., 2001; Moll et al., 2002a; Moll et al., 2002b; Heekeren et al., 2003; Greene et al., 2004; Heekeren et al., 2005; Luo et al., 2006). In addition our results show that when people opt for the pro-social alternative, activations in the latter area are positively correlated with the amount the association receives from the experimenter; however the very same does not apply for the monetary costs experienced by subjects. It is less and less controversial the fact that for every decision the brain attributes values to every option available. vmPFC has a central role in this coding process, as shown by several studies both on animals (Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006) and human subjects (Kable and Glimcher, 2007; Plassmann et al., 2007; Tom et al., 2007; Valentin et al., 2007; Chib et al., 2009; Hare et al., 2009).

Our results suggest that, conditional on having chosen to donate, the coding process of costs and benefits is inelastic with respect to variation in monetary costs experienced by subjects, while varies greatly with the benefits the association will get.

These results appear to be specific to decisions regarding the positive association as, for the negative one, pro-social decisions correlate with activity in the Insula, a region known to be involved in social cognition and emotions' processing (p<0.01) (Craig, 2002; Craig, 2009; Naqvi and Bechara, 2009).
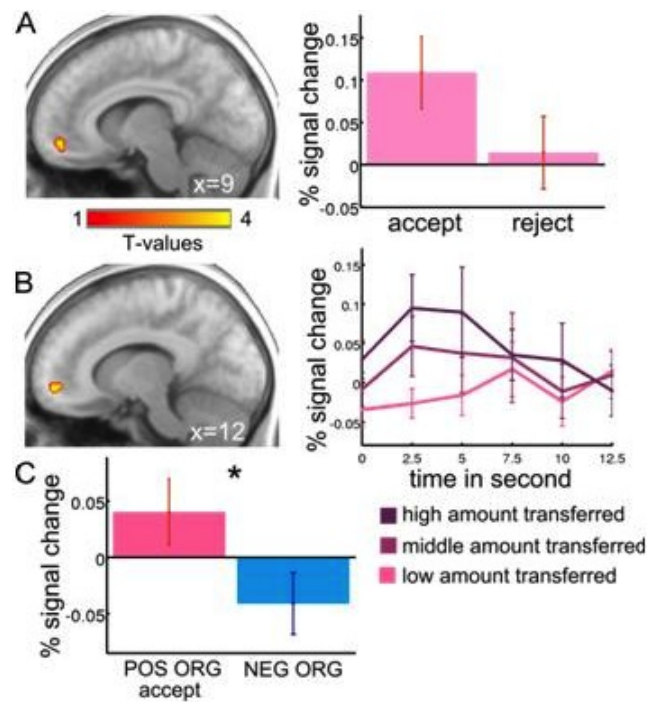
Note: *=0.05; **=0.01; ***=0.001

Figure 1.3: vmPFC

Although these results are in line with previous literature and do confirm that moral rewards share some neural patterns with monetary rewards, they do not inform us about the role social image has in the coding process of monetary and abstract rewards. Section 1.4.1.1 addresses this question.

1.4.1.1 Prosocial *public* decisions and selfish *private* decisions

This section compares, for the two types of decisions (doing good and avoiding doing bad), brain activity for pro-social public decisions and selfish private ones. The objective is to compare how the brain computes image gains with monetary gains.

We obtain three main results: first pro-social public decisions and selfish private ones do correlate with different regions. Second, those regions that correlate with being publicly pro-social are the same for both *doing good* and *avoiding doing bad*. Third, one

region correlates with selfish private decisions for both choices but, within that region, doing good and avoiding doing bad are dissociated.

Result 5: *Pro-social "public" decisions correlate with activity in the Anterior Cyngulate Cortex (ACC) and Putamen for both associations. Selfish "private" decisions are dissociated in the Orbitofrontal Cortex*

For the positive association, we do find that deciding to privately refuse to lose money and make a charitable contribution correlates with the right median Orbitofrontal Cortex (OFC), a region connected to the appraisal of the reward value of reinforcers (e.g. money) (p<0.05). Differently, donating money in public entails activity in the Anterior Cyngulate Cortex (ACC) and Putamen, regions known to respectively monitor conflicts and compute reward-related events (respectively, p<0.001 and p<0.01).

This suggests that in charitable giving contexts, selfish private decisions and public pro-social behaviors involve different regions, though all connected to reward appraisal.

As it is for doing good, for decisions regarding the negative association we observe a higher activation in ACC and Putamen when people decide to be pro-social in public, that is, to renounce to earn money and avoid a monetary transfer to the negative association (for both, p<0.05). Under public exposure, thus, donating to a good charity and renouncing to personal gains to oppose a negative one, correlate with the same brain regions, ACC and Putamen.
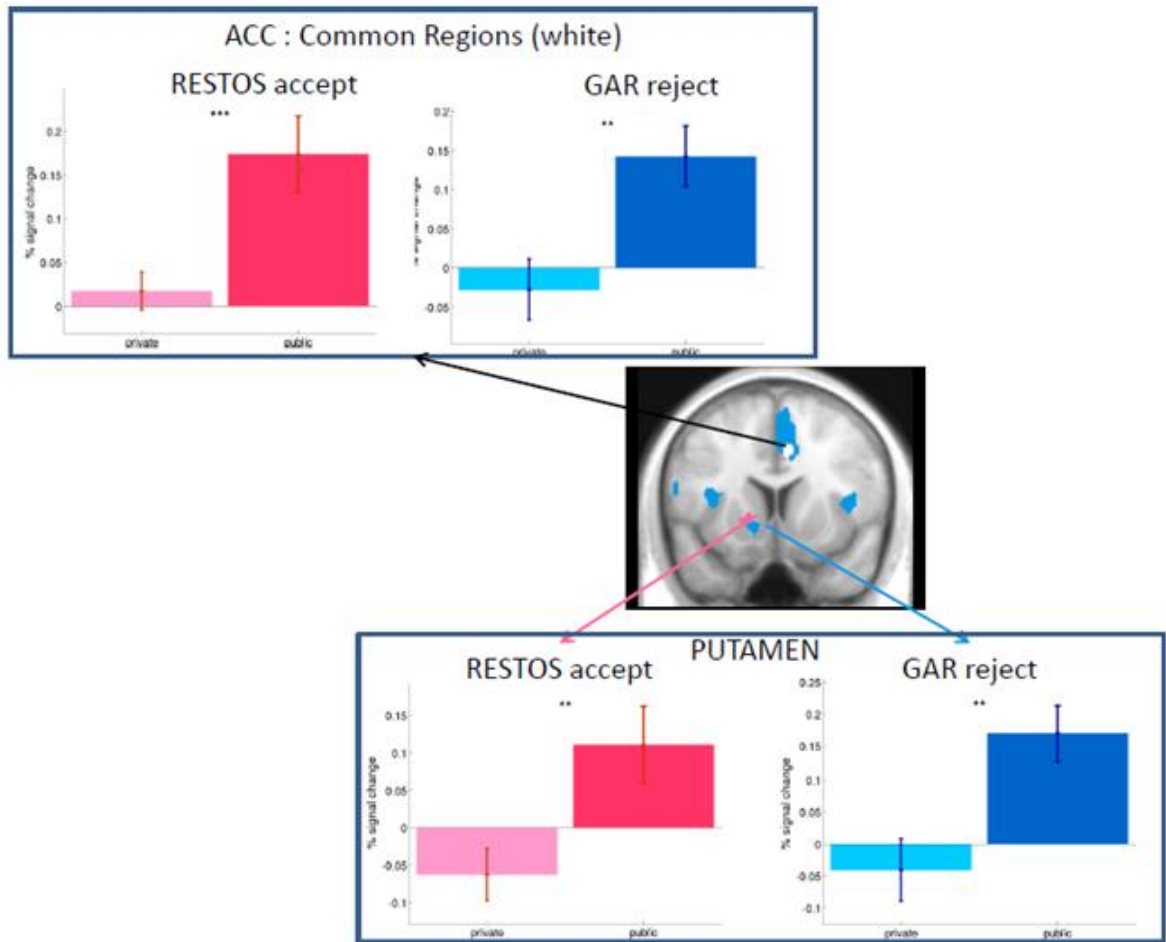
Note: *=0.05; **=0.01; ***=0.001

Figure 1.4: ACC and Putamen

On the other hand, differences between the two organizations are found for selfish private decisions. While ***not*** *doing good* in private correlates with the median OFC, *doing bad* in private correlates with the lateral OFC (left lateral p<0.001; right lateral p<0.01).

Although both types of decisions correlate with the same macro region, recent studies have suggested that the lateral and median OFC may superintend different sub-processes.

The Orbitofrontal Cortex has been found in previous studies to be active in guessing tasks, simple delayed matching tasks and sentence completion. Elliot, Dolan and Frith

(2000), for instance, shown that the OFC is most likely observed when the information available is not sufficient to determine the appropriate course of action, that is, when the likely reward value of stimuli and responses, rather than their nature or type must be taken into account to make a correct decision. Based on these results, O'Doherty et al. (2001) have shown that lateral OFC could be involved in the evaluation of potential punishment, while the median OFC may play a role in the anticipation of rewards. Furthermore, other studies have shown the importance of the lateral OFC in aversive mechanisms (Kringelbach M.L. 2005), such as anger and moral disgust (Blair RJ, Morris JS, Frith CD, Perrett DI, Dolan RJ 1999; Moll J, Zahn R, de Oliveira-Souza R, Krueger F, Grafman J 2005).

Thus, the activation of the lateral OFC we find when subjects choose to *"do bad"* in secrecy suggests that, even without social-image drawbacks, accepting to earn money that creates a negative social externality may push people to anticipate a form of punishment and moral reprehension. The fact that both sub-regions are more active during private trials than public ones can be explained by the fact that selfish decisions are significantly more frequent when subjects choose in secrecy.
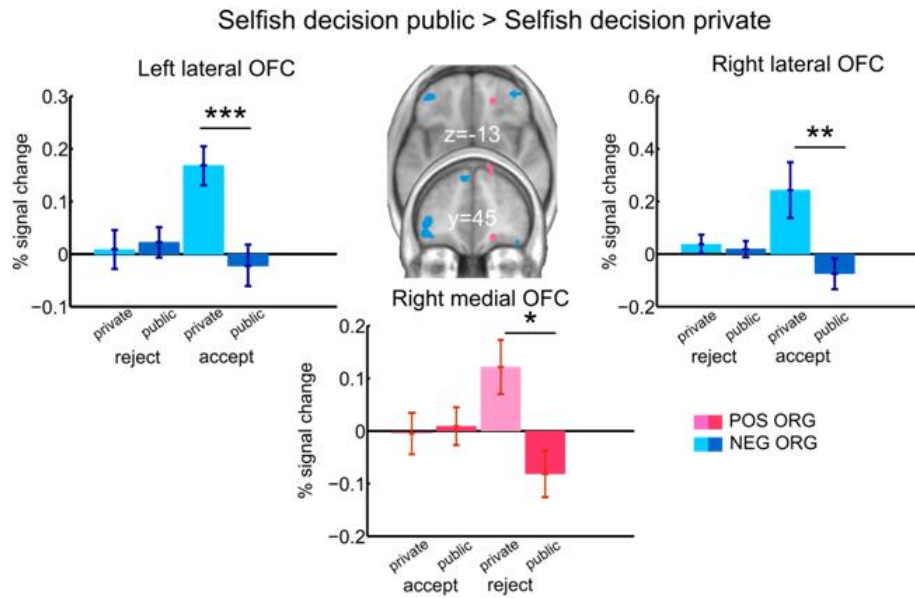
Note: *=0.05; **=0.01; ***=0.001

Figure 1.5: Orbitofrontal Cortex

Two hypotheses can be made about this result: the first one relies on the notion that the violation of shared social norms generates an anticipation of punishment and feelings of moral disgust. Supporting monarchy, in fact, represents in France a strong taboo. Thus, if subjects perceived that being (indirectly) supportive of monarchy represents a violation of an embedded social norm, then accepting a proposition that brings money to the individual but also creates revenues for an association that lobbies for it would generate a sense of moral reprehension and punishment[16]. On the other hand, direct charitable giving does not represent in France a practice as common as it is in other Western countries like the United States[17]. This hypothesis thus would explain why the same result does not apply when it comes to "missed" donations, which may not be perceived as a violation of a widespread practice or social norm.

---

[16] Aversion to monarchy dates back to the French Revolution in 1789, which led in 1793 to the beheading of king Louis XVI.

[17] In 2010 for instance, 90% of US households donated, with an average annual contribution of 2,213$ and a total 298.3 billion dollars donated. In France 56% of households donated in 2010, with an average annual contribution of 366 Euro, totaling 1.9 billion Euro for that year.

36

A second, complementary explanation relies on a more general notion of ethical behavior and consequentialism. Many studies on a thought experiment that has become known as the "Trolley Problem" (see i.e. Greene 2007) have in fact shown that individuals tend to treat isomorphic moral problems differently, depending on whether the negative consequences of an act can be thought as "intentional" or just a "side effects"[18]. In our experiment, by ***not doing good*** an individual saves money; as a "side effect" of this action, a positive charity will not receive a donation. Differently, when an individual chooses to *do bad*, he accepts that in order to earn money he has to intentionally let someone send money to an association he considers bad. Otherwise said, a person who is not being generous is of course not doing good, but after all he is not doing bad either. On the other hand, by accepting to earn money the same person is accepting a *direct* consequence of his act, that is, he is accepting to *do bad*.

Although our experiment is not designed to disentangle between the two theories, both nevertheless suggest that when social image is not a concern, the monetary rewards (and "moral" losses) associated with a selfish decision are weighted differently for the two types of decisions.

To summarize, in section 1.4.1 we have shown that regardless of the treatment condition (public/private), donating to a good association correlates with activity in the vmPFC, a region part of the "reward network", while opposing a negative association with activity in the Insula, a region involved in emotions' processing.

While looking at prosocial public decisions and selfish private ones, section 1.4.1.1 showed that pro-social public decisions correlate for both associations with ACC and Putamen, regions known to respectively monitor conflicts and compute reward-related event. For selfish private decisions instead, we find a dissociation in the OFC for the

---

[18]Many versions of this experiment exist. One commonly used version presents subjects with the following two scenarios: [A] A trolley with five people is doomed to crash if the subject does not activate a rail switch. The switch however is far, as the individual is on a bridge, and the only way to activate it is to push off the bridge a bystander. By doing so the bystander dies but the lives of the five people on the trolley will be spared. [B] The same trolley is running towards a dead end. The individual has his hand on the rail switch. If he activates the switch the trail will change direction and kill a bystander who is sitting on the other trail; however the five people will be saved. If he does not activate the switch the trolley will continue his ride and derail.

two types of decisions: refusing a donation to a good association correlates with activity in the right median OFC, a region connected to the appraisal of rewards, as money and food. Differently, accepting money and allowing a transfer to a negative cause correlates with activity in the lateral OFC, a region related to the evaluation of punishment and negative moral emotions like disgust and reprehension. For both types of decision however, the involvement of OFC in selfish private decisions suggests the presence of an "internal" moral dilemma, which cannot be explained by social image concerns. Social norms' violation, or internal moral codes may be potential explanations.

The next section investigates which regions are more active in private and public treatments, regardless however of the effective decision taken. The aim of the section is to isolate the general effect social exposure has on subjects' neural activity.

1.4.2 Agency and Menthalizing

In this section we investigate which regions are more active when decisions are taken in private or in public, independently from the effective decision taken. In our experiment, subjects are given the possibility to donate their money to a positive association or to oppose a negative one at a personal cost. In both scenarios, participants experience losses (or missed gains). We have seen that for decisions concerning the positive association, the only combination of abstract and monetary rewards that produced significant differences between private and public trials is when both individual's losses and association's gains are high. For all other combinations of monetary and abstract payoffs, people do not behave significantly differently in private and public. On the other hand, for decisions concerning the negative association we observe that social exposure has a positive significant impact on pro-social oppositional decision making only when the benefits the negative association can get are high.

It looks straightforward thus to ask which regions are more active while people take decision in public or in private.

Result 6: *The Insula and Superior Temporal Gyrus (menthalizing network) are more active in public than in private trials for both types of decisions.*

38

In line with previous evidence and with our behavioral results, we find that both Insula and Superior Temporal Gyrus (STg) are more activated when decisions are taken in public. Insula is a region classically involved in emotional appraisal (Craig, 2002; Craig, 2009; Naqvi and Bechara, 2009) , while STg has been found being implicated in shifting attention to focus on another's perspective (Fehr and Camerer, 2007; Hampton et al., 2008; Hare et al., 2010; Nicolle et al. 2012). Taken together, these results suggest that in our context, people's mentalization's process is directed towards the audience, and not to the association itself; if the latter was the case in fact, we should observe higher activations in Insula and STg when propositions are accepted compared to propositions rejected, which is not the case in our experiment. Not surprisingly, for both regions, differences between private and public treatments are significantly stronger for those combinations of abstract and monetary rewards in correspondence of which social exposure alters individual decision making.
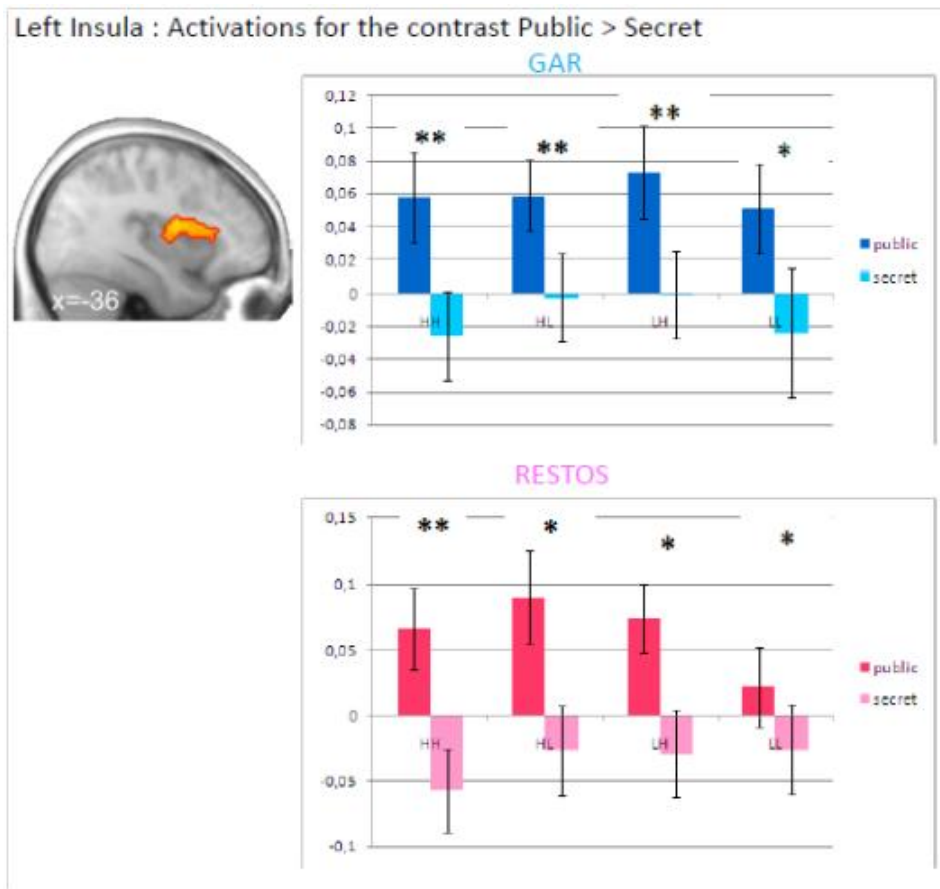
Note: *=0.05; **=0.01; ***=0.001

Figure 1.6: Left Insula
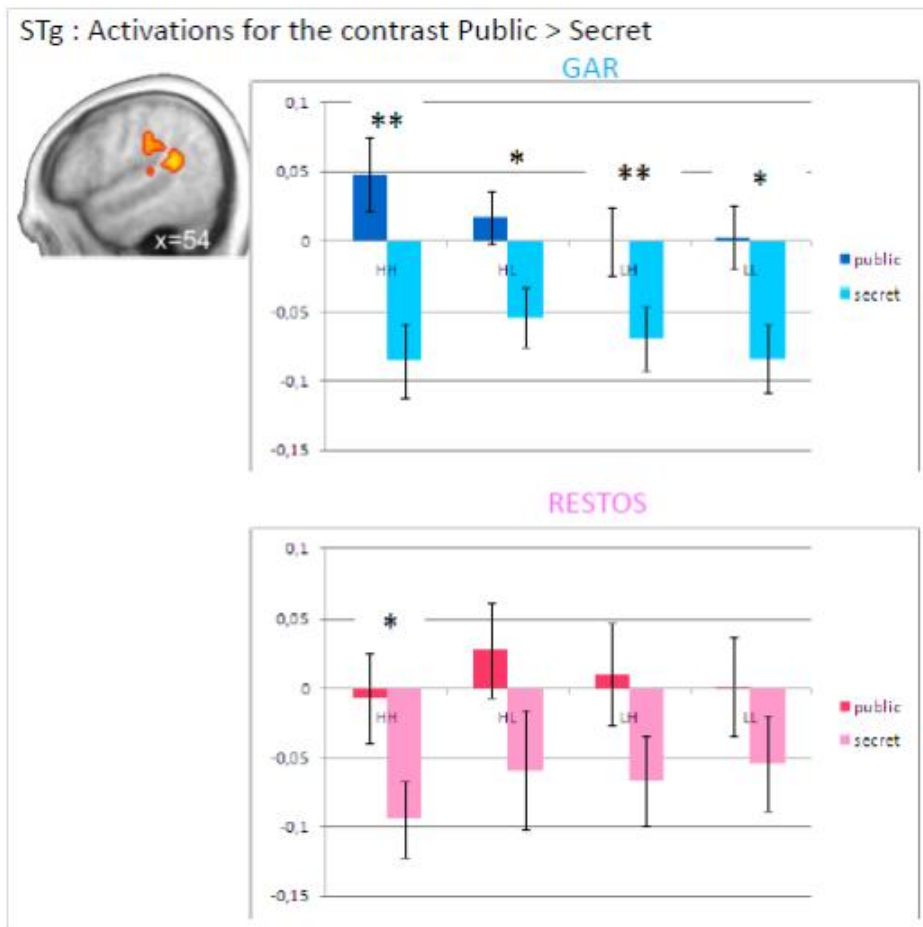
Note: *=0.05; **=0.01; ***=0.001

Figure 1.7: STg

Result 7: *the Inferior Parietal Cortex and the Frontopolar Cortex (agency network) are more active in private than in public trials for both types of decisions.*

While looking at cerebral correlates processing private decisions more than public ones, we find activity in the Inferior Parietal Cortex (IPC) and the Frontopolar Cortex. Those regions are largely known to be involved in the process of distinction between oneself and the others (Decety and Sommerville, 2003; Blakemore and Frith, 2004; D'Argembeau et al., 2010). Activity in the region in fact is more pronounced for self-produced actions (Chaminade and Decety, 2002). The robustness of these findings is

41

corroborated by evidence on patients with lesions of the parietal cortex, which are found be unable to distinguish their own actions from those produced by others (Sirigu et al., 1999). More recently den Ouden et al. (2005) have highlighted the role the right parietal cortex has in perspective memory, which is the ability to keep track of actions that have to be taken in the future. Similarly, Goldberg et al. (2008) show that the inferior parietal cortex is more activated when people take a free decision compared to decisions taken following an exogenous queue. It is worth to note that the very same two regions are more activated during the resting state of MRI sessions (Gusnard and Raichle, 2001).
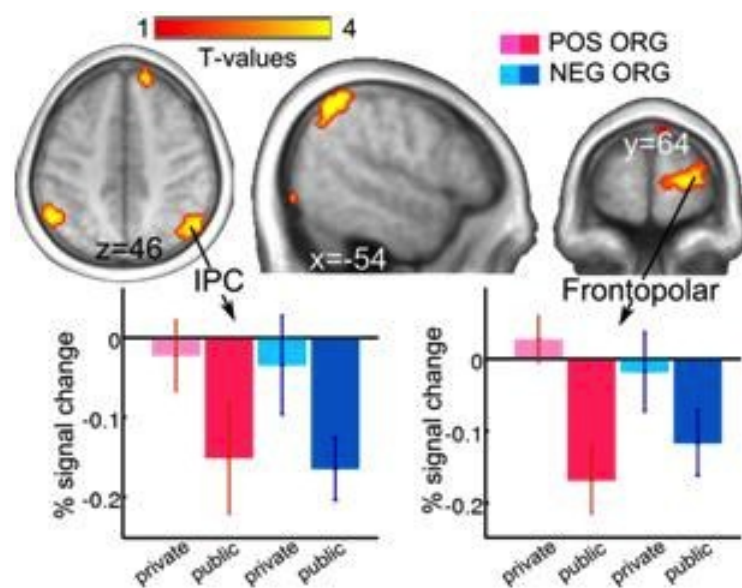
Note: *=0.05; **=0.01; ***=0.001

Figure 1.8: IPC and Frontopolar Cortex
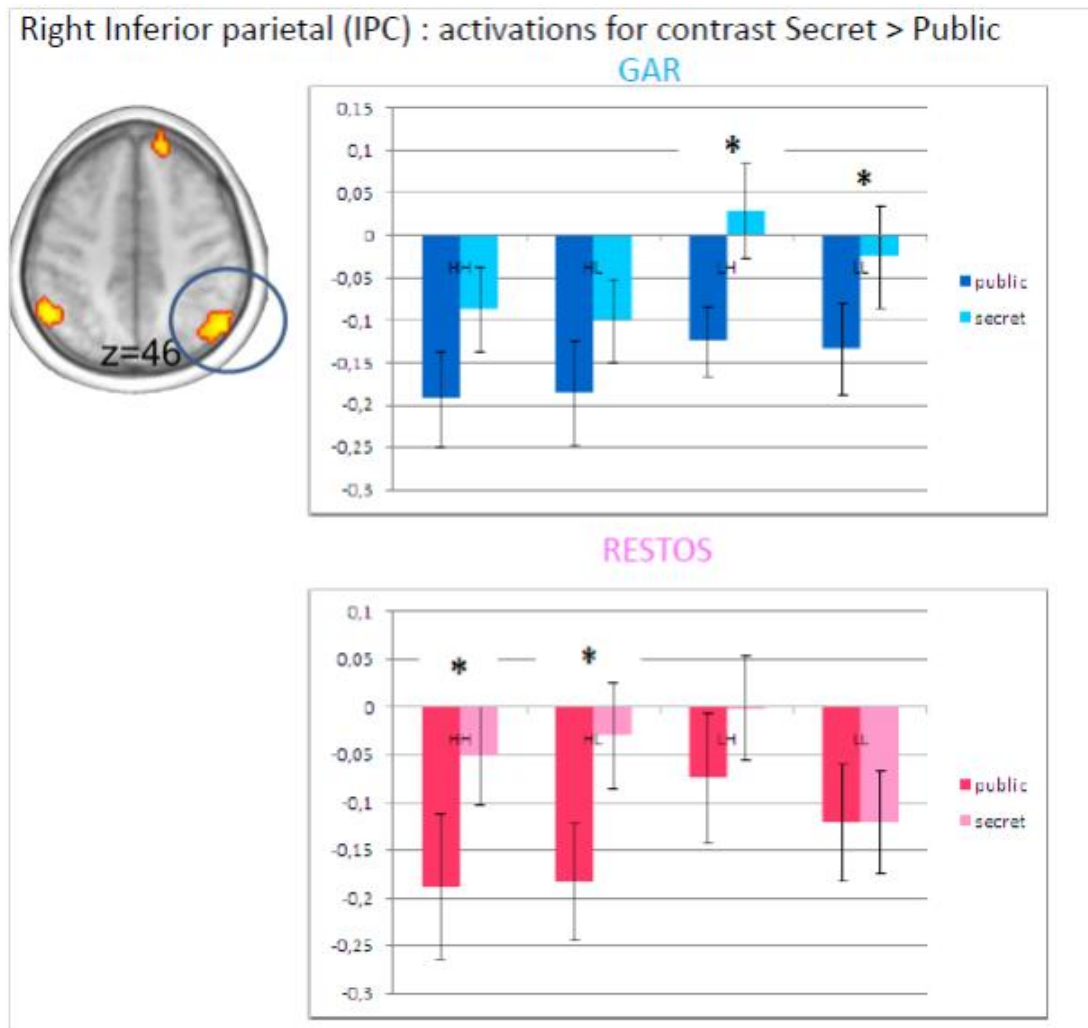
Note: BOLD signal changes by combination of association's and individuals' payoffs. e.g. HL = High payoff for the association, Low payoff for the subject.

Figure 1.9: IPC by different combinations of abstract and monetary rewards

Our parametric design allows us once again to break down which combinations of abstract and monetary payoffs are more likely to drive latter results. If the IPC is a region implicated in the sense of agency, one may expect enhanced activation in private for those combinations of abstract and monetary payoffs in correspondence of which people alter their behavior when observed. We thus expect to see higher activation of IPC in private than in public when both individual's and association's payoff are high

with decisions on the positive association (respectively, 4 to 16 Euro, and 20 to 32 Euro), and when gains for the negative association are high (20 to 32 Euro), independently from subject's gains.

We do find indeed that IPC's stronger activation in private is led for the positive association by trials where the perspective gains for the association in case of acceptance are high ($p < 0.05$). This is no surprise, since these correspond to those cases where private behavior significantly differs from public one. While looking at decisions regarding the negative association, however, we do find evidence that is apparently at odds with our behavioral results: the IPC results more activated in private for trials where perspective gains for the negative association are low ($p < 0.05$). For those trials however, we have seen that individuals' behavior does not change between private and public treatments, raising the question of why for similar behavior in private and public we observe a different activation. One speculation is that when self-serving decisions have a limited negative consequence for others, social exposure is not only ineffective in promoting pro-social behavior, but it crowds out sense of agency associated with selfish though "harmless" decisions. This conjecture is further supported by the enhanced activity of the Striatum observed for the same cases in private treatments.
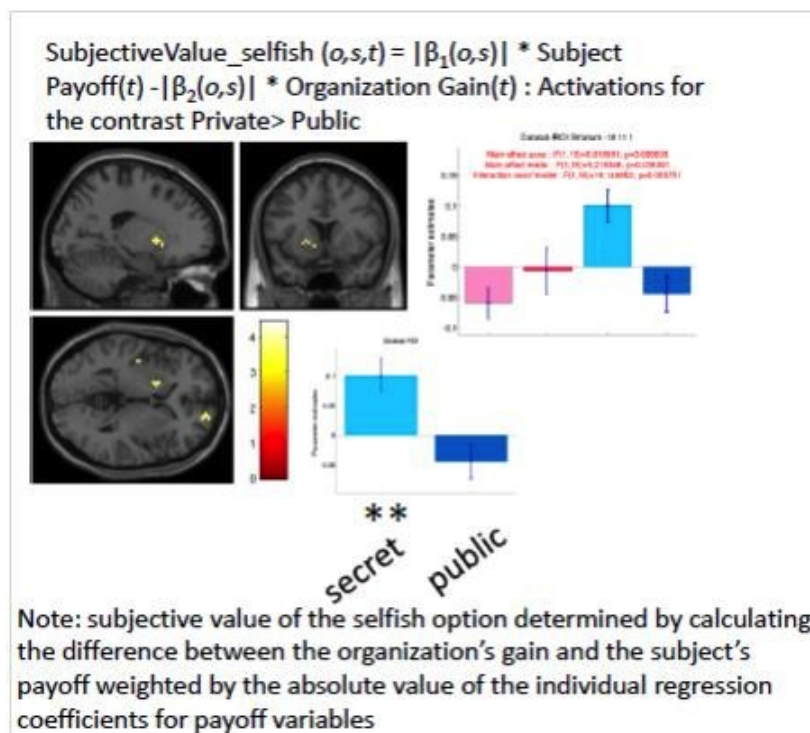
Figure 1.10: Subjective Decision Value

1.5 Conclusion

This essay uses fMRI imagery to compare two types of pro-social decisions where monetary and moral payoffs are at odds: incurring a monetary cost to allow a donation to a positively evaluated charity, and refusing to earn money to avoid a monetary transfer to a negatively valued association. These decisions are taken under two conditions: full anonymity, and social exposure (individuals' decisions are observed by others). We systematically vary the cost of giving by proposing participants variable amounts for themselves (either cost or reward) and variable amounts for the association that would be transferred if the offer is accepted.

For decisions regarding the positive charity, the difference between private and public decisions is driven by dilemmas that are "hard to solve", that is, when losses for subjects are high but gains for the charity are even higher. Differently, for decisions concerning

the negative association, we observe difference between private and public treatments only when the perspective gains for the negative association are high, regardless of the perspective gains for the subjects.

These results suggest that social image may work differently, depending on whether being pro-social means *doing good* or *avoiding doing bad*. When the concern is with people doing good, our work shows that policies aimed to encourage good deeds are most effective when the stakes are high, both for the individual and the recipient[19]. However, when the objective of a policy is to deter private behaviors that generate social negative externalities, our experiment suggests that the efficacy of "social dissuaders" will depend on the magnitude of the damage created.

This gives us a glimpse into how different could be the effect of social exposure for different types of pro-social behaviors. Our results show that social exposure affects decisions that involve not doing bad only when the negative externality generated is relatively large. This may help explaining why it is hard to change people's habits when it comes to choices that, though harming for society at the aggregate level, are hard at an individual level to be thought as "despicable" (e.g. not recycling, crossing streets with red light, avoiding flu vaccinations, polluting).

Our fMRI analysis provides evidence of the neurological and psychological processes underlying these decisions. At a general level, we show that for the positive association, pro-social decisions correlate with activity in the ventromedial prefrontal cortex (vmPFC), a region involved in the evaluation and control of temptation, and moral reasoning. For the negative association we find a correlation between pro-social decisions and the Insula, a region part of a network known to superintend social cognition and emotions' processing. This shows an implication of the vmPFC and Insula in the evaluation of moral benefits deriving from pro-social decisions, confirming the important role the mesolimbic system has in social and motivational mechanisms (Moll, Krueger, Zahn et al. 2006).

---

[19]A paradigmatic example are distinguished chairs in academia, or buildings named after big benefactors.

We then compare pro-social decisions taken in public and selfish decisions taken in private, and find that the two types of associated rewards (respectively "moral" and monetary) are correlated with different brain regions. Pro-social "public" decisions correlate with activity in the Anterior Cyngulate Cortex (ACC) and Putamen for both associations, regions known to monitor conflicts and compute reward-related events. Selfish "private" decisions on the other hand correlate with activity in the Orbitofrontal Cortex (OFC), a region involved in the evaluation of rewards and punishment in contexts where payoffs are uncertain. Within that region, *doing bad* in private correlates with lateral OFC, a part of the OFC connected to the perspective evaluation of punishment and aversive mechanisms (e.g. moral disgust), while ***not doing good*** with the median OFC, a sub-region more connected with the evaluation of rewards. Our study suggests that internal moral taboos or internalized social norms may act as an inhibitor for decisions that *actively* generate a negative externality, while "avoiding to be generous" is processed as a perspective gain, although it generates a "missed" positive externality.

Finally, our study is the first that shows how regions implicated in agency feelings (IPC and Frontopolar Cortex) are more active when people choose in private whether to do good or to avoid doing bad. Symmetrically, choices made in public recruit the Superior Temporal Gyrus and the Insula, regions recognized in previous literature to be involved in social cognition, mentalizing and emotions' processing.

Overall, our work provides important insights for social scientists on the neural and psychological mechanisms underlying human altruism and social awareness. Individuals appear to be motivated by prestige in their pro-social actions, evaluating the representations other may have of their own actions, while feeling more "in command" when those decisions that are sensitive to social image concerns are taken in private. This applies to both decisions that imply a positive externality or a "missed" negative externality. Different mechanisms though appear to regulate these two pro-social behaviors (and associated rewards): our data suggest that self-image or internalized social norms play a crucial role in the evaluation of consequences from *doing bad*, in particular for "secret" actions, while both *doing good* in public and *not doing good* in private appear to be

connected with anticipation of a reward (either a social or a monetary one). These results expand our understanding of the neurobiology of human pro-sociality, suggesting that the way the brain computes rewards and costs (both moral and monetary) depends on the underlying nature of the pro-social decision.

What represents the difference between these types of pro-social behaviors remains an open question.

Further work in fact is needed to understand whether our results are evidence for a general difference between "active" pro-social decisions and "passive" ones (e.g. donating money vs not accepting bribes), or represent evidence for how the brain differentiates payoffs according to whether they violate or not social norms.

Essay 2

Good News, Bad News, and Social Image: The Market for Charitable Giving[1]

2.1   Introduction

The last fifteen years have seen a remarkable increase in the availability of information about charities' finances, practices, and transparency policies. Several non-profit organizations, such as Charity Navigator[2], the Urban Institute[3], and GiveWell[4], have since started to monitor and review charities, developing indices and measures aimed to help donors to easily compare different charities and make more informed decisions.

Although both qualitative and quantitative factors are important in the analysis of a charities' activity, synthetic measures of financial efficiency currently represent a cornerstone for the evaluation of charities' ability to deliver what they promise[5]. While large

---

[1]This chapter is a joint work with Jeffrey Horn.

[2]http://charitynavigator.org/

[3]http://www.urban.org/nonprofits/index.cfm

[4]http://www.givewell.org/

[5]For example, Charity Navigator provides a measure called "program expenses", which measures the share of the total budget that is spent on programs that further the stated the mission of the charity. Charity Navigator provides several other measures of financial health, including fundraising effectiveness and

donors always had both the incentives and the means to gather this type of information, determining the worthiness of a particular charity required nontrivial research efforts for small donors: gathering appropriate government documents, phoning charities to request information, interviewing administrators and board members, or even serving on boards themselves. The emergence of these information clearinghouses however, gives now charities an incentive to report on their own, since being unrated or un-reviewed might look suspicious to donors. As these mechanisms both increase the amount of information readily available and reduce the cost of access, donors of all levels may now seek information at relatively little expense.

While indeed some donors may choose to remain ignorant about these measures[6], the effect this information has on those that do encounter it is poorly understood.

Borrowing a terminology that has become classic in the literature (see i.e. Andreoni 1989), a *pure altruist* donor, for instance, may respond to new positive information about his charities' efficiency by reducing his contribution, as now less money is needed to meet his previous expected effective donation[7]. Otherwise said, if efficiency is perceived as a proxy for the *efficacy* of giving, then the *implicit cost of giving* would decrease as positive information is received. Conversely, an *impure altruist* may increase his contribution in response to unexpected good news, as this positively contributes to his *warm-glow*. If for these individuals efficiency represents a proxy for *deservedness*, good news would be rewarded and bad news punished.

Studying the effect of information is important also for another type of impure altruist: *image motivated donors*. In fact, as information becomes bundled with charities' names, donors who value the social prestige associated with charitable giving may well take it into account as they choose *how much* and *to whom* donate. Suppose that a donor gave money for years mostly to look good to his peers: if the name of his charity becomes

---

administrative expense. In Charity Navigator's methodology, financial health determines at least half of the charities' overall four-star rating.

[6]Whether small donors seek this information or remain rationally ignorant is an important empirical question we will not address in this current work.

[7]i.e. the percentage of his contribution he was expecting, given his priors, to be effectively used for the stated mission of the charity.

more and more publicly associated with high quality and efficiency, he may actually *reduce* the size of his gifts, as now a new positive signal contributes to the prestige of his generosity, reducing thus the *(subjective) cost of looking good* (i.e. "I (only) gave 50$, but my charity is very efficient!"). Roughly speaking, an image motivated donor may *trade-off* the size of his gift with the efficiency of the recipient. Alternatively, he may indeed decide to increase his contribution if he expects some scale economies for his image payoff. Understanding the interaction between information and prestige is thus critical: cheap, widely accessible information about charities' efficiency in fact makes it progressively harder for a donor to signal his own generosity without indirectly providing information about the efficiency of his gift. Even if this information is sometimes ignored, it cannot be hidden from others when donors advertise how generous they are.

Whether and *how* the response to information depends on its social visibility thus, represents a second important unanswered question.

Indeed, several studies have pointed out the importance of social-image motives in charitable giving (see i.e. Harbaugh 1998a, 1998b; Andreoni , 1988, 1990). Several studies have shown how the presence of indirect social approval incentives increases individual contributions (e.g. Andreoni and Petrie 2004; Rege and Telle 2004). It is not merely generosity per se that matters, but also the relative cost of giving (lower than one) (Andreoni and Miller 2003; Karlan and List 2007) and its social visibility (Ariely, Bracha and Meier 2009).

Few papers however have studied how information about the quality of the recipients affects generosity (see i.e. Fong 2007; Fong and Oberholzer-Gee 2011), and very little is known about the effect this information has when it is common knowledge.

The objective of the present study is thus twofold: (i) investigate how new information about real charities' efficiency affects small donors' contributions, that is, studying donors' reaction to a *real* prices of giving *greater than one*, and (ii) understand whether the public visibility of that information matters for donors, that is, if impure altruists motivated by *warm glow* respond to private information as impure altruists motivated by *prestige* would when that information cannot be hidden from others.

To do so we implement a two-phase experiment, with the second phase unknown to the subjects in the first phase. In phase one, subjects freely choose three charities from a list of more than 5,000 charities rated by Charity Navigator, and play three independent dictator games with the selected charities[8]. Subjects are also allowed to indicate a favorite charity among the three, a decision that increases the probability of that specific charity being randomly selected for payment. While the first phase serves to assess the unconditional willingness-to-give of subjects, in the second phase we assess how decision made in the first phase are revised in light of new information about the charities. In phase two, subjects are incentivized to correctly guess what they believe to be the true efficiency of their charities[9], and then receive privately the true efficiency values. In what follows, we consider that a subject discovering a charity is more efficient than expected receives good news, and that a subject discovering a charity is less efficient than expected receives bad news. Finally, subjects are given the opportunity to revise their initial donation decisions and to indicate a new favorite, if desired[10].

As we are interested in comparing responses to efficiency information under different levels of social exposure, we implement three between-subjects treatments: (T0) donation decisions and efficiency are private information, (T1) the donation to the final charity is publicly revealed but information about efficiency remains private, and (T2) both efficiency and donation amount of the randomly-implemented final decision are publicly revealed.

As mentioned, predicting a priori donors' reaction to new information about efficiency is not trivial. First, individuals may perceive this information as a signal of *deservedness*. As a consequence, people may reward better-than-expected charities by

---

[8]For each charity, participants are asked to choose how many E\$ of their endowment they want to send to each charity, knowing that only one decision will be randomly implemented at the end of the experiment.

[9]For our measure of efficiency, we chose "program expenses" from Charity Navigator, which captures the percentage of total budget used to implement programs serving the stated mission of the charity. Charity Navigator calculates all financial measures on the basis on the financial information each charity provides in their public tax returns (IRS Form 990).

[10]This means that participants can switch favorite charity across phases, or indicate a favorite in phase 2 while they did not in phase 1 (and vice versa). It's important to notice that decisions from phase 2, which are the only ones taken into account for payments, cannot influence those made in phase 1, as subjects during phase 1 are unaware of the existence of a subsequent and final phase. This also means that participants do not know whether others did change their decisions across phases.

donating more, and punish worse-than-expected charities by donating less. Second, people may interpret information as a signal of *efficacy*, or *implicit cost of giving*. In this case, donors may increase donations to worse-than-expected charities to offset high overhead, and may donate less to better-than-expected charities when it turns out they "need" less money that expected to meet a certain objective. Third, people may not care at all about synthetic measures of efficiency, considering them as merely reductive means to compare charities that differ considerably in nature[11]. Finally, the way donors respond to information may depend on whether the latter is private or common knowledge.

Clearly, a person that trades off efficiency with donation amounts for image purposes would behave exactly like someone that maintains constant the effective efficacy of the donation for pure altruism reasons (i.e. reducing gift in response to good news in order to maintain the same expected effective gift). Similarly, if efficiency is perceived as a signal of deservedness or quality, people *doing good* and people *doing well* would be indistinguishable.

The important economic question here however is if, given a heterogeneous population of donors, the response in terms of giving to information that is private[12] is behaviorally similar to the response to information that is common knowledge.

Our results suggest that this is not the case.

As studies on self-serving biases and imperfect belief-updating show (see i.e. Bradley 1978; Babcock and Loewenstein 1997; Svenson 1981; Eil and Rao 2011; Sharot et al. 2011; Sharot et al. 2012), people tend to avoid or ignore information that carries negative value for them. If donors believe their causes are deserving, when confronted with information that suggests otherwise, such as a lower-than-expected efficiency score, donors might discount or ignore the information.

We expected such behavior in our first treatment, where all information and decisions are private, as there can be no image effect associated with ignoring bad news. In other

---

[11]Even for charities that share a similar purpose, measures of efficiency may be perceived as useless, as other characteristics that distinguish one from the other may be very salient for individuals (i.e. different location, religious vocation, helping children in Ghana or in Brazil etc.).

[12]Or at least that can be hidden from others.

words, ignoring bad news is relatively cheap. Indeed, we find that in our private treatment, participants show no significant response to bad news. Good news on the other hand significantly increases donation amounts across phases. The latter is true also for our second treatment, when donations amount are made public, but information about efficiency remains private.

Second, although we were agnostic on whether good news would increase or decrease giving for image-concerned donors, we expected however to see a significant change in behavior when charity efficiency was made public. Indeed, we find that the way individuals adjust their contributions in response to new information is sensitive to efficiency being publicly revealed. We find that in treatments where efficiency is private information, all participants increased their giving in response to good news. However, in the treatment where efficiency is publicly announced, 34% of those that reacted to information did so by donating *less* in response to good news, and donating *more* in response to a bad one.

We show that an interpretation that explains different behavior in public and private treatments is that efficiency and donation amounts may be *substitutes* in terms of social image. This means that an image-motivated donor would be indifferent between announcing a large donation (without specifying anything else) and announcing a smaller donation that goes to a highly efficient charity.

Our results suggest that as values of efficiency measures become common knowledge, the most efficient charities are likely to lose some part of their support, due to the interaction between social image concerns and visibility of information. Without social image concerns, providing efficiency information donors will increase donations, since approximately 75 percent of donors do under-estimate charity efficiency. However, advertising high efficiency widely will decrease donations from certain types of donors. Considering the whole market, we find that the relatively strengths of these effects are equal and total volume in the market for giving is unchanged by making efficiency public.

2.2 Background

Economics literature has long addressed the question of why people donate money to private charities. As the traditional idea that people are motivated only by pure altruism fails to explain several empirical observations about charitable giving (e.g. large participation, incomplete government crowd out, average contributions non decreasing in the number of contributors) (see i.e. Andreoni 1988), several theories have been called to explain private charitable giving. In line with the intuition that individuals may derive a direct utility from giving (see i.e. Becker 1974), internalized norms (Arrow 1971; North 1981), social approval(Hollander 1990), warm-glow (Andreoni 1990; ; Ribar and Wilhelm 2002; Harbaugh, Mayr and Burghart 2007), conditional cooperation (e.g. Fischbacher, Gachter, and Fehr, 2001), reciprocity (e.g. Sugden, 1984), and prestige (Harbaugh 1998a, 1998b; Bracha, Heffetz and Vesterlund 2009), have been identified as powerful motivations for giving.

In particular, the empirical evidence of the importance of social image and prestige is vast. The possibility of direct and indirect social approval incentives generally increase individual contributions (Andreoni and Petrie 2004; Rege and Telle 2004). As individuals appreciate the positive image consequences of giving, their generosity depends also on the cost of giving - or nominal price of giving - (Andreoni and Miller 2003; Karlan and List 2007) and how others would perceive ones' own generosity given its cost (Ariely, Bracha and Meier 2009). Social influence and imitation as well play an important role (List and Lucking-Reiley, 2002; Shang and Croson 2009; Vesterlund 2003; Potters, Sefton and Vesterlund 2007; Bracha, Menietti and Vesterlund 2011).

Recent studies show that the evidence of giving for self-signaling reasons is weak compared to social-signaling motives (see i.e. Grossman 2010). Although much is known about how warm glow and social image modulates at a general level individual giving, less is known about how information about the recipient of the gift affects generosity[13].

---

[13]For empirical evidence on the effects of identification of versus information on the recipient see Small and Loewenstein 2003.

In this direction, Fong and Oberholzer-Gee (2011) use real individual recipients and costly information to show that a significant fraction of their subjects is willing to pay to gather information about the recipient and achieve a distribution of income that matches their preferences, and that they use this information to withhold resources to less preferred recipients. Overall however, with costly information not all donors are willing to invest resources to find preferred recipients. Differently from previous literature, our work uses real charities instead of individual recipients and explores the effect on generosity of a *real* price of giving *greater than one*. We investigate how cheap available information (as it is nowadays) affects giving, and how the public visibility of that information matters for donors.

2.3   Experimental Design

2.3.1   Overview

We recruited 99 subjects from George Mason University to participate in a simple, individual decision-making experiment. The mean age was 22.25, with 53.95% of men and 46.05% of women; 70% of subjects took at least one course in Economics (with 57% with more than 2 courses)[14].

Subjects were seated at private computer terminals and were instructed not to interact or communicate during the experiment. All decisions were made privately. The experiment lasted approximately one and a half hours, and consists of two phases, with the second phase revealed to subjects only at the end of the first phase.Data was collected from May 2012 to July 2012 using pencil-and-paper, but subjects used a computerized search interface for part of the experiment. In all treatments subjects were paid by a third person unfamiliar with the experiment[15].

---

[14]67.4% of subjects declared to have donated money at least once in the last year (any sum to anyone), 69.8% to have volunteered, and 12% to have tithed.

[15]As detailed further, this means that in all treatments the experimenter is unaware of the overall decisions made by participants and the characteristics or name of their recipients.

2.3.2   Phases

We designed three treatments for this experiment, each with two phases. In the first phase of all treatments, subjects were asked to choose three charities from a list of more than 5000 charities. With each of the charities, subjects chose how to split their endowment between them and the charity, knowing that only one split (and thus one charity) would be randomly selected for final implementation. In the second, surprise phase of all treatments, subjects received new information about their charities and were allowed, should they wanted to, to adjust their initial decisions in response to this new information. One of three decisions from the second and last phase was implemented according to a compound lottery.

2.3.3   Treatments

Our three treatments differed in whether the implemented decision was publicly revealed and whether the news each subject received was publicly revealed. In no treatment was the name of the randomly chosen charity revealed, nor was the name of the subject revealed, nor was any personal information about the subject revealed.

In our control, Treatment 0 (**T0**), all decisions and information were private. In Treatment 1 (**T1**), subjects were required to stand up at the end of the experiment and announce only how much they donated to the randomly chosen charity. In Treatment 2 (**T2**), subjects were required to stand at the end of the experiment and announce both the amount donated to the randomly chosen charity and the information received about that charity in the second phase [16].

Note that phase 1 decisions are fully comparables across treatments T1 and T2, as in both subjects' information set is the same (i.e. they only know that the donation to the randomly selected charity will be publicly revealed).

---

[16]Subjects in T1 and T2 were explicitly told that the name of the charity should never be revealed.

Table 2.1: Experiment Design

	Phase One	Phase Two	End
T0	Pick charities, familiarity quiz, comprehension quiz, initial donation decisions, and choice of favorite (if ever). Participants aware all decisions will be private.	News explained; comprehension quiz; subjects guess their charities' efficiency; real efficiency revealed; final donation decisions, and choice of favorite (if ever). Explained that efficiency of the final charity will be private.	One of three decisions implemented by compound lottery.
T1	As above, but subjects aware that final donation will be made public.	As above. Subjects reminded donations will be made public. Explained that efficiency of the final charity will be private.	As above. Subjects must stand and announce the implemented donation amount.
T2	As above, subjects aware that final donation will be made public.	As above. Subjects aware that both donation and real efficiency of final charity will be made public.	As above. Subjects must stand and announce both donation amount and news.

2.3.4 Detailed Procedure

Subjects were endowed with 25 experimental credits (EC; equivalent to $17 US). Participants were presented with a web-based search interface for a database of approximately 5,400 charitable organizations, which represents the full list of charities rated by the charity watchdog Charity Navigator (CN)[17].

Subjects selected three charities from the database and answered questions about their familiarity with and attitude toward each charity (Appendix B).

After completing a comprehension quiz, subjects decided how to split their initial endowment of 25 EC between themselves and each charity[18].

---

[17]Charity Navigator is a nonprofit organization which analyzes and rates charities based on their financial records and transparency (http://www.charitynavigator.org). Information about CN is provided to participants only at the beginning of Phase 2. CN rates charities that have both local branches and national or international offices. Subjects were only allowed to choose one office, either local or (inter)national, but not both.

[18]Any integer amount from zero up to and including 25 EC could be chosen to send to each charity.

Subjects were instructed that the three decisions were independent and that at the end of the experiment only one decision would be randomly selected for implementation. We emphasized that only one of the three charities would receive the donation, and that subject would be paid according to the split chosen by the subjects for the randomly-implemented charity. Subjects were assigned a random ID number, and were told that donations would be made on their behalf by the experimenter, using the ID number as a donor name. Later, subjects were able to collect the receipt of the donation, and verify that the correct amount was sent. If subjects were in Treatment T1 or Treatment T2, they were informed that the donation amount they allocated to the randomly-implemented charity would be publicly revealed at the end of the experiment.

Although we could have asked participants to pick only one charity, we decided that three independent decisions would allow us to assess both the absolute and relative effect of revealed information (news) on donations[19].

In addition, we were interested in knowing whether subjects were willing to increase the probability that one of three charities would be implemented, both before and after news was revealed. By having such selection mechanism in both phases, we are able to assess what charity among the three subjects prefer the most prior information about efficiency is revealed, and the extent to which subjects in phase 2 would be willing to alter this selection after information is disclosed. As we move along treatments, private information becomes more and more public: a cross treatment comparison of changes in favorites thus, allows us to infer the degree to which information alters intrinsic ordinal preferences of individuals (e.g. in T0 and T1) or alters extrinsic motives for giving (e.g. T2). Therefore, subjects were allowed to designate one charity as their "favorite".

A compound lottery consisting of a coin toss and a die roll determined which decision would be implemented. If no favorite was indicated, each phase 2 decision stood a one in three chance of being implemented (33 percent)[20]. However, conditional on a favorite

---

[19]For a detailed discussion of this experimental procedure please see section 2.5

[20]At the end of the experiment we rolled a die in front of everyone. An outcome of one or two indicated the decision associated with the first charity on each subject's list would be chosen; three or four, the second; and five or six, the third.

being chosen, that favorite stood an increased chance of being implemented of two in three (66 percent)[21].

Since subjects were not informed that this was a two-phase experiment at this point, indicating a favorite indicated a personal preference for the selected charity. In Phase 2, a different charity could be chosen as a favorite, indicating either a personal preference or a reaction to news, since choosing a favorite at this point was conditional on having received news about all three charities.

At the beginning of Phase 2, subjects were asked to hide their decision sheets and return the initial list of charity names to the experimenter. This list was used to create the final decision worksheet for Phase 2. We used a computer program to collect the names of each charity profiled by CN[22], as well as a homogenized measure of charity efficiency called "program expenses"[23]. Program expenses is ratio of dollars spent providing services in pursuit of the charity's stated mission or purpose over the charity's total functional expenses. The residual below unity can usually be thought of as the percentage of a charity's budget spent on fundraising and administrative expenses, although a few charities have nontrivial expenses in other categories[24]. Collecting this data in advance allowed us to use the search interface we designed to quickly and accurately find the program expenses rating associated with the charities on each subject's worksheet[25]. These ratings were handwritten by a lab assistant on the worksheets.

While we were preparing the final decision worksheet, we explained what "program expenses" measured and gave an example. We asked subjects to guess each charity's actual program expenses ratio, as well as how confident they were in their guess by

---

[21]We also flipped a coin at the end of the experiment. An outcome of "heads" indicated the donation decision associated with each subject's favorite charity would be implemented. If the outcome was "tails", the result of the die roll determined which decision was implemented.

[22]Horn, Jeffrey R. Charity Navigator Scraper. https://github.com/jrhorn424/charity-navigator-scraper

[23]Charity Navigator,section "How Do We Rate Charities' Financial Health?".

[24]A footnote about average program expenses (or the distribution), admitting they vary by sector, is appropriate here.

[25]Ratings were returned almost instantaneously due to two design factors. First, data was stored in the web page (charity names for subjects; both names and ratings for experimenters) rather than on a server, eliminating the lag from Internet communications. Second, the search scope was narrowed with each keystroke; so-called "fuzzy matching". This works well for small datasets, but requires longer page load times as web pages approach megabytes in size.

revealing the subjective likelihood that the true ratio fell within each of five quintiles. Subjects were informed that they would later receive the true values of efficiency for their charities: if their guess was within +/- 5% from the true value, they received additional 6ECU at the end of the experiment, provided that the charity for which the guess was correct was randomly selected for payments.

If subjects were in Treatment T1 or Treatment T2, we reminded them that the donation amount they allocated to the randomly-implemented charity would be publicly revealed at the end of the experiment. If subjects were in Treatment 2, they were informed for the first time that the program expenses rating of the randomly-implemented charity would also be revealed.

Subjects in T0 were reminded that all their decisions and information would be kept private, and subjects in T1 that only the amount sent to the final charity would be revealed. Afterward, we returned the final decision worksheets with the charity news (program expenses ratio) to the subjects, at which point subjects were given the opportunity to adjust their initial split and to change favorites, if desired. Finally, one of the three decisions was implemented randomly according to the previously described compound lottery[26].

2.4   Results

We organize our results as follows: in section 3.4.1 we present general results on (i) overall donations, (ii) real efficiency levels and individual guesses about efficiency, and (iii) decisions about favorite. In section 2.4.1.2 we combine individual guesses and real efficiencies to examine how the quality of information (bad/good news) about charities' efficiency affects donation levels and choices of the favorite charity across treatments.

Before proceeding, it may be useful to recall how information sets in phase 1 and 2 differ across treatments. In phase 1, participants are never told anything about efficiency:

---

[26]For treatments T1 and T2, once the final charity was randomly selected, subjects where orally reminded to cover the other 2 decisions as the experimenter would stand by them when they announced their donation (both treatments) and the charity's efficiency (only T2). As this procedure was orally explained also in phase 1, subjects could not misreport their decisions or expect to do so.

by consequence treatment T1 and T2 look exactly the same from phase 1 perspective: in both, the only information provided is that the donation amount to the randomly selected charity will be announced to others.

Differently, in treatment T0 participants' decisions are kept private. The outcome variables from phase 1 in treatments T1 and T2 can thus be aggregated, and compared with treatment T0 results. On the contrary, in phase 2 the information sets are different in all treatments. In treatment T0, both donation amount and real efficiency of the final charity remain private information; in treatment T1 the donation amount to the final charity is revealed to other participants while its real efficiency remains private; finally, in treatment T2 participants know that both information will be made available to other participants.

2.4.1 General Results

Table 2.2 reassumes our principal descriptive statistics.

Table 2.2: Summary Statistics

Summary Statistics	N	mean	sd
T0			
Donation Phase 1	81	8.86	8.18
Donation Phase 2	81	9.45	8.48
Favorite Phase 1	81	0.66	0.47
Favorite Phase 2	81	0.62	0.48
Efficiency Guess	81	68.93	18.07
Real Efficiency	81	79.50	15.14
T1			
Donation Phase 1	84	10.30	7.88
Donation Phase 2	84	10.49	8.24
Favorite Phase 1	84	0.78	0.41
Favorite Phase 2	84	0.85	0.35
Efficiency Guess	84	74.15	11.39
Real Efficiency	84	82.34	8.41
T2			
Donation Phase 1	132	8.84	7.92
Donation Phase 2	132	8.86	8.23
Favorite Phase 1	132	0.70	0.46
Favorite Phase 2	132	0.77	0.42
Efficiency Guess	132	68.98	18.13
Real Efficiency	132	80.04	13.63

Note: favorites are % of individuals indicating a favorite.

2.4.1.1 Donations

Overall, participants did donate a significant part of their endowment of 25E\$ to the charities they choose. With only 16.8% of subjects donating zero in phase 1, and 17% donating zero in phase 2, participants gave on average 9.25 E\$ in phase 1, and 9.48 E\$ in phase 2. This leaves to subjects average final earnings of 17 US dollars (including show up fee).

A two-sided Jonckheere-Terpstra test shows that donations in phase 1 are not significantly different across treatments (p=0.724); similarly, donations in phase 2 are overall

similar across treatments (p=0.392)[27]. Finally, the average difference of donations between phases is not different across treatments (p=0.282).

One may be surprised to notice that making donations public does not increase the average individual contributions: by comparing differences in donations between phases in fact, we find that the average difference in private treatment (T0) is not significantly different from the average of the two treatments where donation amounts are visible (T1 and T2) (p=0.286)[28].

Indeed numerous papers have shown that public visibility generally increases donors' contributions. However, in most of those papers individuals face only one decision while in our experiment each participant takes three simultaneous decisions. By consequence, an immediate comparison of our results with previous literature may be misguiding[29].

2.4.1.2   Efficiency and guesses about efficiency

We now turn to how people formed guesses about efficiency in our experiment.

As explained in section 2.3, participants received in phase 2 information about their charities' efficiency, after having guessed their values. Clearly one of the concerns could be that differences across treatments in terms of donation amounts would be hard to interpret if either real efficiency values or guesses were systematically different across treatments.

Real efficiency values encountered both within and between treatments are fully comparable. As the total average of real efficiency was 80.5% (s.d. 12.85), averages of real efficiency were not significantly different across the three treatments (p= 0.792).

---

[27]Because more information is revealed to other participants when one moves along treatments, at a general level of analysis it seems natural to compare outcomes using the distribution free Jonckheere-Terpstra test, which tests whether a variable significantly differs as we move along k ordered treatments. The interpretation of the results is similar as for companion non-parametric tests like Mann Whitney with one exception: p-values for two sided test in this context represent the probability that a variable consistently differs as one either ascend or descend the ordered treatments.

[28]Unless mentioned otherwise, all p-values from pairwise comparisons come from two-sided Wilcoxon-Mann-Whitney tests or Wilcoxon matched-pairs signed-ranks. For comparisons of three treatments, all reported p-values come from two-sided Jonckheere-Terpstra tests.

[29]When comparing within subjects donations to the three charities (i.e. how much to the first charity, to the second, etc.), we do not find significant differences in either phase 1 or phase 2 (respectively p=0.431 and p=0.512).

This means that across treatments, subjects selected charities very similar in terms of efficiency. As participants choose three charities however, one may wonder if efficiencies are comparable also between subjects, regardless of the treatment they are in[30]. We do not find evidence that some subjects systematically received much skewed draws of charities.

We can draw essentially the same conclusion by looking at participants' guesses. For the latter one additional concern is in place: in treatment T2 in fact, differently from the other two treatments, participants were guessing the efficiency knowing that its true value would be revealed to others: the way in which people do form beliefs thus may have been biased by this information. By confronting average guesses in our T2 treatment versus the other two however, we cannot reject the hypothesis that indeed subjects did form their guesses in the same way across treatments (p =0.432). Finally, also average guessing errors are not significantly different across treatments (p=0.839).

Taken together, these results show that across treatments, participants received the same proportion of good and bad news. Once again thus, one cannot interpret results as a consequence of a systematic difference across subjects and treatments of the quality of information and/or individual guesses. Table 2.3 shows the proportion of good and bad news received in each treatment.

Table 2.3: Proportion of good/bad news received by treatment

	Good News	**Bad News**	**Total**
T0	60 (74.07%)	21 (25.93%)	81 (100%)
T1	64 (76.19%)	20 (23.81%)	84 (100%)
T2	102 (77.27%)	30 (22.73%)	132 (100%)

Note: proportion of good and bad news by treatment

---

[30]Clearly, choosing for instance three charities that have very bad efficiency scores may impact behavior very differently than choosing three highly efficient charities. Similarly, keeping the value of efficiency constant for one charity, one may react to that specific information very differently depending on the relative rank of that charity with respect to the other two.

2.4.1.3 Indicating a favorite

We finally turn to decisions about the favorite charity.

Looking at these decisions, we see that overall most of participants indeed choose to indicate a favorite in both phase 1 (71% of subjects) and phase 2 (75.7%). Decisions regarding the favorite in phase 1 do not differ across treatments when we compare all treatments (p=0.921), as well as when we compare treatment T0 with the other two pooled together (p=0.672). The latter result is important as it excludes that the visibility of donation amounts (known from phase 1) alters individual preferences.

While looking at choices in phase 2 instead, we observe that more participants decide to indicate a favorite charity in phase 2 when we move from T0 to T1 to T2 (p=0.073). It is not surprising thus to observe that when more information is made available to other participants, people also start to switch favorite from one phase to the other more often. When participants choose to select a favorite in phase 2 and/or switch favorite, it is always to switch in favor of a more efficient charity (for treatments T0, T1, T2, p=0.012, p=0.0001, p=0.002 respectively)[31]. The fact that the latter results hold also for treatment T0, suggests that people took seriously the efficiency of their charities[32].

---

[31]We obtain the same result comparing the real efficiency of all charities with the real efficiency of charities selected as favorite in round 2.

[32]Because in treatment T0 the experimenter is unaware of all the decisions of participants, experimenter demand effect cannot be called to explain results from that treatment. This represents stronger evidence than considering all treatments together, as the experimenter in the other two (T1 and T2) witnesses participants declaring their final decision.

2.4.2   The good/bad news effect on donations

In the previous section we have shown that participants did donate a significant portion of their endowment, that they faced comparable efficiency levels across treatments, and did not form guess in a significantly different way from one treatment to another. These general results allow us to analyze more in depth the relationship between the subjective quality of the news and the response in terms of giving under our treatments.

As all participants form incentivized guesses about the efficiency of their charities, some guesses may be lower than the true value of efficiency, some accurate, and others higher than the true efficiency. As we have shown that people do care about how the beneficiaries of their gifts spend their resources, and that people do form meaningful guesses, which we guarantee by monetarily incentivizing correct guesses, we assume that discovering that one's own charity is more efficient than expected represents good news for the donor, while realizing that a charity is not as efficient as one thought represents bad news. Figure 2.1 provides a scatter plot of how variations in donations from phase 1 to phase 2 do relate with the type (and intensity) of the efficiency news.
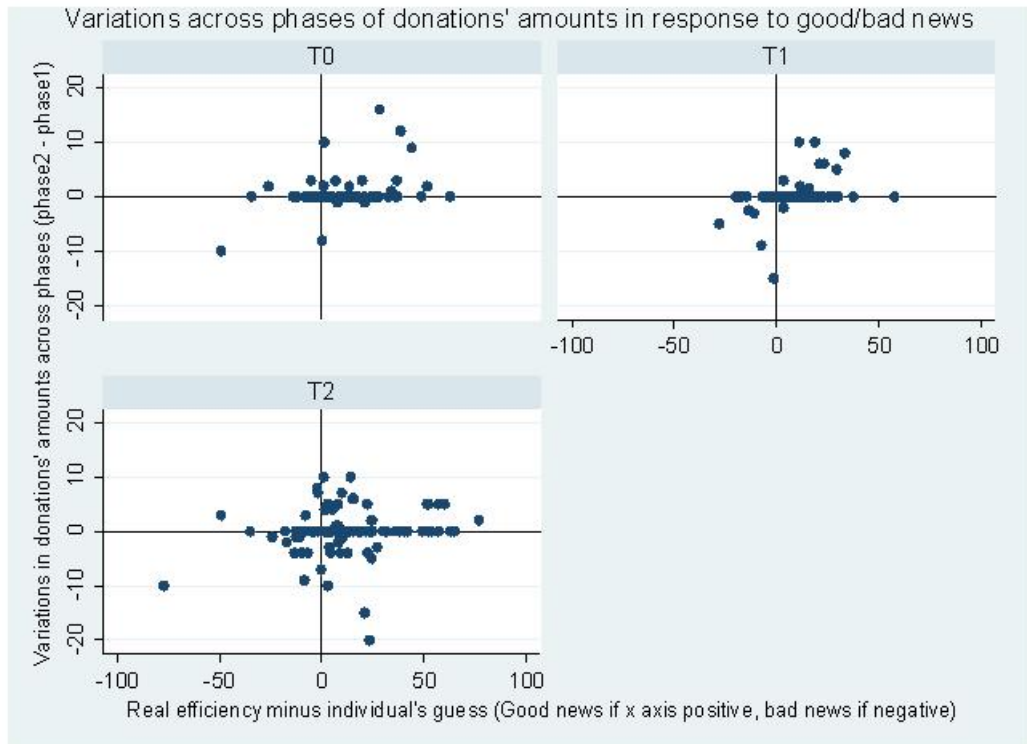
Figure 2.1: Changes in donations across phases by type of news

One can immediately notice that the public visibility of efficiency has a dramatic effect on donors' behavior: treatments where efficiency is private information share the same pattern of response to good news: in both, good news increases individuals' contributions. Differently, when efficiency is made visible to other participants, some subjects start to reduce their contributions in response to good news.

As the only difference is the visibility of final charity's efficiency, the emergence of this new behavioral pattern can be attributed to individuals that do care about the effect efficiency has on their social image. In addition, it can be noticed that when information is fully private (treatment T0) individuals do not adjust their donations in response to efficiency levels below their expectations.

Table 2.4 provides summary statistics about how subjects varied their donations across phases, computed by the type of the news they have received.

Table 2.4: Number of increases and decreases in donations by type of news

	Good News			Bad News		
	n	mean (E$)	s.d.	n	mean (E$)	s.d.
T0						
Increased Donation	11	1.1	3.13	2	0.25	0.78
Decreased Donation	3	3.3	4.04	1	10	
No Change	46			18		
Total	60			21		
% Changed	23%			14%		
T1						
Increased Donation	10	0.83	2.31	0		
Decreased Donation	1	2		5	6.9	5.2
No Change	53			15		
Total	64			20		
% Changed	17%			25%		
T2						
Increased Donation	21	1.11	2.32	4	1.05	2.39
Decreased Donation	12	6.25	5.78	10	4.3	3.33
No Change	69			16		
Total	102			30		
% Changed	32%			46%		

Result 1: when both donation amounts and efficiency are private information (treatment T0), individuals do largely ignore bad news about their charities but reward good news by increasing contributions.

In each treatment, about 25% of the information about charities' efficiency represents bad news for subjects. While in treatments T1 and T2 people did vary their contributions in response to bad news[33] , in treatment T0 we see virtually no variations from phase 1 to phase 2 in terms of donation amounts (p=0.632)[34].

---

[33]For treatment T1 people reduce donations in response to bad news (p=0.020). In treatment T2 some individuals do reduce their donations in response to bad news and some, as discussed further, do increase.

[34]In treatment T0, the average donation in phase 1 before a bad news is 9.3 E$, and 9.1 E$ in phase 2 after bad news is received. In treatment T1 is 9.8 E$ and 8 E$ respectively; in treatment T2, 12.1 E$ in phase 1, and 11.4 E$ in phase 2.

The same result holds when we consider reductions in donation between treatments: reductions after a bad news are significantly lower in T0 compared to treatments T1 and T2 (p=0.021). As in treatment T0 no social image component is involved, this result may be explained by aversion to negative information or image value of ones' own charities. If in fact the absence of response to bad news was driven by a low elasticity of donations to information (i.e. people simply don't care about information), we should observe no variations in donation amounts also in response to good news. We do see instead that when news is good, participants do indeed modify their donation behavior by increasing their contributions to better-than-expected charities (p=0.023), suggesting that efficiency is treated as a signal of *quality* or *deservedness*.

This asymmetry in information processing suggests that the same mechanism of imperfect updating found in other areas of decision making[35] may be in place when it comes to evaluate negative (private) information about charities people care of: in this sense, (private) good and bad news would not only be a proxy for deservedness or quality of the charity, but would directly affect the utility individuals get from holding a positive opinion of their charities. Without a direct "self-image" utility for the subject in fact, we should observe a reduction of donations in response to bad news.

Another argument could be that subjects do not think being "fair" to reduce their support to a good cause, even in face of new negative information; if that was the case however, we should observe the same behavior at least in the closest treatment T1 which, as explained further, is not the case[36].

The unresponsiveness to bad news thus suggests that a mechanism of self-reward may accompany the evaluation of charities' efficiency, which is consistent with the anecdotal evidence that people attach an identity value (or self-image value) to the charities they donate to. This result has an important consequence for the overall market for giving: if people do not punish charities that are worse-than-expected but do increase

---

[35]see i.e. Eil and Rao (2011); Mobius, Niederle, Niehaus and Rosenblat (2012).

[36]People in fact may follow a rule but believe that it would not be socially acceptable. Even with this type of private rule however, response to bad news should be the same in treatment T1 because information about efficiency is still private information.

donations to better-than-expected ones, then the diffusion of efficiency measures will increase the total volume for those donations made under anonymity, no matter the proportion of good and bad news in the population[37].

Result 2: *When efficiency remains private information but the final donation amount is disclosed to others (treatment T1), participants start "punishing" worse-than expected ones by reducing contributions, and continue to "reward" better-than-expected charities by increasing their gift.*

In treatment T1, when people respond to new information, good news is always rewarded with increased donations (p=0.006), while response to bad news is associated with reductions in donation levels (p=0.027)[38]. This result suggests that when information about efficiency is private but donation amounts are visible, individuals treat efficiency as a proxy for *deservedness* or *quality*, punishing thus inefficient charities and rewarding good ones.

Although in treatment T1 the donation amount to the final charity is revealed to other participants, the value information about efficiency has in treatment T0 and T1 should nevertheless be the same, as efficiency has no social image component in both treatments. It is surprising thus to observe that making the donation amount public pushes some participants to vary their donation amounts in response to bad news.

A possible explanation is that the visibility of the final donation amount raises the salience of information about efficiency, making subjects more sensitive to information that they would otherwise disregard. This is still coherent with the idea that taking into account negative information has a cost, which is avoided as far as the cost or effort needed to disregard information is low.

Alternatively, one may argue that individuals use donation amounts to indirectly signal that they do care about efficiency: suppose individuals hold correct priors about what would be the average contribution from a donor that cares about his charities;

---

[37]By pooling all decisions from treatment T0 in fact, we see that donations in phase 2 are significantly higher than donations in phase 1, no matter the quality of information (p=0.029).

[38]Average reduction in response to bad news is 1.8E\$, and average increase in response to good news is 1E\$.

as high donations may signal that the subject chose a very efficient charity, a low donation may signal that the individual is responding to information that he believes being important for others. The latter interpretation has however some shortfalls: first, participants know that others do not observe their own initial decision; by consequence, any declared amount says nothing about whether the individual changed or not his previous decision (and in which direction). Moreover, making a small contribution represents a noisy signal: a low donation may either signal that the person is punishing a charity, or that the individual is simply greedy. The only explanation we feel to exclude is that bad news is used as an excuse to raise one's own payoff: that being the case in fact, we should observe the same behavior in treatment T0.

Both interpretations however suggest that when information about efficiency is private but individual generosity public, individuals treat efficiency as a proxy for *deservedness* or *quality*.

Result 3: *When efficiency is revealed to other participants, a significant fraction of subjects start to reduce donations to better-than-expected charities and to increase contributions to worse-than expected-ones.*

As the efficiency of the final charity becomes public information (T2), we observe two major differences with respect to treatments T0 and T1: first, the percentage of subjects that change their donations between phases in response to new information is significantly higher in T2. Second, while in treatments T0 and T1 almost all of the variation comes from subjects that increased donations in response to good news and decreased donations in response to bad news, in treatment T2 this relationship breaks down, as a significant portion of subjects modify their donations in the opposite way: 12 out of 33 (36%) variations after good news is received are negative variations, and 4 out of 14 (28.5%) variations after bad news represent increases in donation amounts.

As the only difference in treatment T2 is the visibility of efficiency of the final charity, deviations from behavior observed in previous treatments can be only attributed to the social image effect of efficiency visibility.

To formally test this intuition and analyze why the public visibility of charity efficiency pushes a significant fraction of subjects to reduce donations in response to good news and to increase donations in response to bad ones, we divide subjects from T2 in two groups: in group 1 we place all subjects that do have at least one decreased donation after good news (or increased donation after a bad news); this represents the group of subjects that show a behavior that is virtually absent from treatment T0 and T1. For sake of simplicity we call these decisions "deviant" observations. In group 2 we place all the other subjects. Because the minimum requirement to be in group 1 is to have only one "deviant" observation, if results from T2 treatment are due to errors made by subjects or noise, the two groups constructed in this way should show no significant difference one from another, no matter what we compare across the two.

As one may argue that those "errors" can be big relatively to the other observations, it is important to notice that all the following results obtained comparing the two groups do hold even when we categorize in group 1 subjects that had at least two "deviant" observations instead of one. Group 1 includes 12 subjects (36 observations), and group 2 32 subjects (96 observations).

We thus proceed to investigate whether there are systematic differences across the two groups in terms of news received, beliefs, and donation decisions. A first explanation for our results from treatment T2 could be that individuals from group 1 systematically received only good news or only bad news: this being the case, a subject receiving only good news may choose to increase his contribution to the charity that did relatively better than expected, and reduce the donations to the other two charities (whose efficiency would still be a good news).

The same logic would apply for a subject receiving only bad news. By comparing the proportion of good and bad news received in the two groups however, we cannot reject the hypothesis that the two groups indeed received the same composition of good and bad news (p= 0.179)[39]. The same conclusion can be drawn using a two-groups proportion test (p=0.518).

A second possibility is that subjects from group 1 form guesses in a significantly different way; being this the case, the metrics we use to define good and bad news would be unsuitable to compare the two groups. We find no evidence that subjects from the two groups did guess differently (p=0.376).

If news received and beliefs are the same across the two groups, why thus the visibility of charities' efficiency induces group 1 subjects to decrease their gift to better-than-expected charities and increase it in response to bad news?

We do find that the two groups show significant differences in terms of donations' patterns, choices over the favorite mechanism, and the level of concentration of their donations among the three charities.

First, we find that in phase 1, individuals in group 1 donate on average more than subjects in group 2 (p=0.024); the difference between the two groups however disappear when we compare average donations in phase 2 (p=0.341). Group 2 participants donate less in phase 1 but increase their average donations in phase 2 (p=0.011); group 1 participants instead, start with high donations in phase 1 but on average do not increase in phase 2 (p=0.319). Together, these results rule out the possibility that subjects in group 1 are reducing donations in face of good news because they are intrinsically greedier[40].

Second, group 1 subjects switch favorite charity significantly more than group 2: while looking at decisions to either change charity between phase 1 and phase 2, or moving from not having a favorite in phase 1 and indicating one in phase 2, subjects from group 1 use the switch mechanism significantly more than group 2 subjects (p=0.000).

---

[39]The same results are obtained if we confront subjects from group (i) with all other subjects from all three treatments.

[40]Even more convincingly, we can notice that the average donation in phase 1 for group 1 subjects is 10.52 E5 (s.d 5.88), which is very close to the average donation in phase 1 for subjects from treatment T1 (10.29 E$, s.d. 7.8).

Finally, we look at the relative concentration of donations among the three charities between phases. We have seen that group 1 subjects donate more than group 2 in phase 1, but their average donations do not increase across phases; as group 1 also relies significantly more on the switch mechanism, looking at the concentration of donations among their charities is important to understand whether these subjects do simply change their favorite charity without altering the relative proportion of money they send to each of them, or if in phase 2 they put relatively more money on fewer charities.

To do so we use an individual measure of donations' relative concentration in phase 1 and phase 2, namely the Herfindahl index[41]: as the number of charities is fixed to three in both phases, if subjects are indeed concentrating their donations after information is received, the Herfindahl index should be bigger for phase 2 than for phase 1. We find indeed that for group 1 subjects, the total relative amount made available to the three charities is significantly more concentrated in phase 2 than in phase 1 ($p=0.000$). On the other hand Herfindahl indexes from phase 1 and phase 2 are not statistically different when we consider group 2 subjects, and participants in treatments T1 (respectively, $p=0.158$; $p= 0.941$)[42].

To explain these differences we advance two complementary hypotheses: (1) that subjects from group 1 are those that do care the most about the social image effect of efficiency visibility, and (2) that for these donors, efficiency and donation amounts are *substitutes* in terms of their social image payoff function.

The argument that group 1 indeed represents the fraction of subjects that appreciate the most the social image component of efficiency is simple: if group 1 subjects are image motivated, they should donate more than subjects in group 2 in phase 1, as in phase 1 the amount they donate to their charities is the only factor they know it can affect their

---

[41]The Herfindahl index is a measure of the size of firms in relation to the industry and an indicator of the amount of competition among them. It is defined as the sum of the squares of the market shares of the 50 largest firms (or summed over all the firms if there are fewer than 50) within the industry, where the market shares are expressed as fractions. The result is proportional to the average market share, weighted by market share. Increases in the Herfindahl index generally indicate a decrease in competition and an increase of market power, whereas decreases indicate the opposite.

[42]For our baseline treatment, T0, the relative concentration of donations appears higher in phase 2 than in phase 1 ($p=0.007$). This however is not surprising given that in T0 all differences in donations between phases are positives, making thus by definition the index bigger in phase 2.

social image: in phase 1, being generous for them involves an additional image benefit, absent for group 2 subjects. When phase 2 begins however, participants learn that also efficiency will be revealed to others: this means that in phase 2, participants in group 1 have an additional factor that enters positively in their social image payoff function, namely the final charity efficiency. Given that about 75% of the times this new visible information affects positively the image payoff function (i.e. good news), then the social image of group 1 subjects is in phase 2 at least as good as it was in phase 1.

As a consequence, on average they face no need to put more money on the table. On the contrary, some resources now can be taken away, as for them the relative *cost of looking good* has decreased; when it comes to better-than-expected charities thus, appearing good in phase 2 costs relatively less because now, by also signaling that their money is well spent, group 1 subjects can attain the same image utility of phase 1 with a smaller contribution.

On the other hand group 2 subjects, as it happens in T0 and T1, would reward good news, increasing thus their contributions from phase 1 to phase 2.

Finally, if subjects in group 1 care more about what others see, they should rely more on the "favorite mechanism" to raise the probability of a specific charity being selected; they should also concentrate more their gifts in those charities whose donations and efficiency they would be happier to see announced; this means that if we assume that the distribution of charities' efficiency and individual preferences are not correlated, and that the distribution and intensity of preferences over the three charities are not different across group 1 and group 2, then group 1 subjects should switch favorite charity relatively more often than group 2 subjects, and should concentrate more their donations in phase 2, as for them the social image gains from choosing a specific charity are relatively higher. Our previous results suggest thus that the two groups indeed represent different types of donors, and that group 1 subjects are those that value more the image value of charities' efficiency.

We conclude by looking at how efficiency and generosity are related in terms of social image payoffs. To see why subjects from group 1 do trade off efficiency with generosity,

one can notice that the essential difference between treatment T1 and treatment T2 is that in the latter, efficiency has an instrumental value: in T1, the only way image motivated donors can signal their pro-social attitude is to make large donations. Differently, in treatment T2 two pieces of information are made available to others: how *generous* the donor is, and how *efficient* his charity is. It appears reasonable to assume that both efficiency and donations' amounts are positive arguments in terms of social image payoffs (i.e. both are normal superior goods). In treatment T2 thus, an image motivated donor has two means to increase his social image payoff: by signaling that he cares about others (high donations), and by signaling that he cares about how well others serve their stated mission (high efficiency). As a consequence however, when efficiency becomes common knowledge, an image motivated individual may attain the same level of image utility he was expecting before learning that efficiency would be visible by donating slightly less but showing that his generosity is effective.

The strategy we use to test this hypothesis is to construct a measure of ex-ante expected efficacy of gift and an ex-post objective measure of efficacy, and confront them. Before receiving true values of efficiency, in fact, subjects were asked to guess the efficiency of their charities. As at that point subjects had already made their decisions in phase 1 bud did not know the true efficiency, the way in which individuals form guesses could only be conditional upon the way they decided to donate, and their priors about the charities' ability to deliver what they promise. We thus compute the ex-ante expected efficacy of the gift by multiplying the guess about efficiency by the donation made in phase 1. Similarly, once subjects receive the true value of efficiency, the final ex-post objective efficacy of gift is represented by the true value of efficiency multiplied by the donation amount selected in phase 2.

The intuition behind the comparison between these two values is the following: an individual that appreciates the social-image value of efficiency who wants to keep constant the marginal benefit from social image, would choose his donation in phase 2 such that the image benefit he was expecting to receive in phase 1 by declaring only the donation amount, equals the image benefit that can now be obtained by having to declare

both the efficiency and the donation amount. Assuming that his guess rationalizes the donation he made in phase 1, by doing so he would maintain the same image payoff, and increase his personal monetary payoff.

While we compare ex-ante expected efficacy and ex-post objective efficacy, we do find indeed that these are not statistically different for only one group of subjects: group 1 subjects from treatment T2 (p=0.74). For group 2 subjects instead, the ex-post objective efficacy is always significantly higher than the ex-ante expected efficacy (p=0.0001); similar results are found for treatments T0 and T1(p=0.0001 and p=0.0002 respectively).

Taken together, these results show that the visibility of both charities' efficiency and donations pushes image motivated donors to concentrate their donations in fewer charities and to trade off donations' amounts with efficiency.

2.5 Discussion

Our experiment was designed to maximize the meaningfulness of giving, to assess the effect real information on real charities has on donors' behavior, and to evaluate how its impact depends on whether information can be hidden or not from others.

In doing so, we opted for a procedure that necessarily implies some degree of endogeneity in the selection process of charities, and uses an objective measure of financial health which is not completely immune from critiques by practitioners. In this section we wish to address these points in detail.

A first argument concerns the comparability of efficiency across sectors. As different charities serve different missions, the same nominal value of financial efficiency may not mean the same thing when we move from one activity to another. For instance, an efficiency of 80% would be largely recognized as a sign of good health for a charity serving education purposes. However, the same 80% efficiency would look more troublesome for a charity whose main activity would normally involve low overhead and fundraising costs as, for example, providing meals for the homeless.

The size of the association represents another sensitive problem, as smaller charities may not be able to benefice from large economies of scale, which in turn would hamper indicators of financial and organizational efficiency without necessarily mean they stand in poor organizational and financial health.

Finally, it has been argued that larger overhead and fundraising costs may at some point of a charity's life be necessary to deepen investments, and more generally that using only measures of program expenses to evaluate charities can be highly misleading[43].

Indeed our participants were able to select different charities from different sectors, and the pool of charities we presented them includes charities of different sizes. Although we are aware that this may generate some noise, we believe that the benefits of this approach outcast the shortfalls.

A first alternative may have been to restrict our sample to only a subset of homogenous charities belonging to the same cohort. This approach however would have generated in our opinion two layers of much more serious problems: first, by restricting the sample one would truncate observations from subjects that, although willing to give, would not find in the pool charities they like. Second, given the nature of the second phase of the experiment, forcing subjects to choose charities similar one to another would increase the risk of an experimenter demand effect, as with similar charities subjects may feel "forced" to consider them as commodities and reward the most efficient. A second alternative could be to randomly assign subjects to charities, which however would still make hard to disentangle between experimenter demand effect and true generosity. If in fact subjects cannot choose their charity but are randomly assigned to it, they may have an incentive to either send zero or to show that they are *unconditionally* generous.

Overall, we believe that if on the one hand our procedure makes hard the control for the underlying selection process, on the other it allows for a much richer and externally valid environment: donors in fact may have a rather heterogeneous set of

---

[43]E.g. Nancy Lublin, CEO of DoSomething.org, a non-profit that targets youth involvement in social change projects, writes that "[...]low overhead doesn't necessarily mean an organization is awesome at fighting poverty, or that its turnover is low and its people productive. And it certainly doesn't guarantee that the group is spending wisely." (source: http://www.fastcompany.com/1297922/)

causes they care about, which translates into donations to charities very different one from another. As our study precisely aims to understand the effect of a *homogenous* measure on a *heterogeneous* market as charitable giving, trading off some controls over selection processes with wider choice sets appears to us a reasonable compromise.

Second, our design uses a single measure of financial efficiency, the percentage of the total budget spent on the program. Indeed all charities' evaluators and watchdogs, including Charity Navigator, do provide donors with a much richer set of information and indicators. However, as individuals' opinions about what matters in charities' activity can be very different, and individuals may face different costs of searching for information, we did prefer to give subjects a single measure to limit problems of identification. On the other hand, it is true that most watchdogs do indeed weight and synthesize the vast amount of information provided by charities in simple indexes: while "program expenses" may not be the most "information intensive" index, it does represent nevertheless an important variable of financial health, which in turns accounts for more than half of the criteria used by Charity Navigator.

Finally, although we acknowledge that simply relying on synthetic measures of efficiency may be sometimes misleading, we do believe that the same observation may be made for prices in any commodity markets, and does not undermine in any way the importance of understanding how individuals react to signals (and their social value) in a market, would the signals be prices or efficiency measures.

2.6 Conclusion

Our study investigates the effect on charitable giving of the public visibility of charities' efficiency and donors' gifts.

With treatments that progressively increase the visibility of donation amounts and charities' efficiency, we assess the effect bad and good news on ones' own charities have on donors' behavior. First, we find that individuals tend to disregard bad news about their own charities when giving happens under full anonymity, but do increase

their contributions to charities that turn out to be more efficient than expected. Overall thus, to the extent to which giving is made in private, our results suggest that the diffusion of homogenous efficiency measures would increase the overall size of the market for giving. Second, we find that individuals start to react to bad news as soon as donations amounts are made public. Taken together, these results suggest that charities' efficiency is treated as a proxy for *quality* or *deservedness* whenever this information remains private. Differently, we show that when also charities' efficiency becomes public, a significant fraction of individuals starts to reduce their contributions to better-than-expected charities and increase their gifts to worse-than-expected ones. We suggest that when information about efficiency has a social image value, some individuals trade off donations with efficiency, paradoxically reducing revenues of those charities that perform relatively better than others. With information about efficiency becoming part of the core information people have about charities, our experiment shows that donors motivated by their social image may reduce their generosity towards better than expected charities, as the visibility of charities' quality makes appearing generous relatively less expensive.

Appendix A

INSTRUCTION PHASE 1

Thank you for participating in today's experiment. You've earned a $5 show-up bonus for participating. In reading and following the instructions below, you have the potential to earn significantly more. In the experiment you will receive Experimental Dollars (E$s), which will be converted into cash (US Dollars) at the end of the experiment.

The conversion rate from Experimental Dollars (E$) to cash (US Dollars) is the following: 3 Experimental Dollars (E$s) are worth 2 US Dollar in cash.

In today's experiment you will be randomly assigned an ID number: this ID number will identify your choices in today's experiment. Please keep it at hand, as you will be asked to write it on various sheets.

Please note that other participants will not be told your name, and you will not be told their names.

The next section gives a summary description about today's experiment and your task.

Experiment Description

You have been given today 25E$ (excluding show up fee).

On your desk there are 2 envelopes. Please write your ID number on both envelopes now, and then put them aside. You will use them later, when instructed.

In today's experiment you will be asked to select three charities from a large list of charities.

For each of the three charities you choose, you will decide how to split 25E$ between yourself and that specific charitable association. You can send to the charity any amount between 0E$ and 25E$: the amount you indicate will be sent to the charity, and you will keep the remaining E$.

At the end of the experiment, one of the three charities will be randomly selected. You and the randomly selected charity will be paid according to how you decided to split the 25E$. The other two decisions, on the other hand, will not be implemented.

This means that the three individual decisions you will make are independent from one another.

We will send the portion allocated to the randomly selected charity on your behalf. We will use only your ID number as the donor name, and we will make the donation online, whenever

---

[1] Parts of the instructions specific to treatment T0 are red-colored, parts for T1 in green, and parts for T2 in blue.

possible. Later, you will be able to check that the proper amount has been sent to the charity by checking the ID number and the amount on the donation receipt.

A copy of the receipt of the donation will be made available in the ICES Lab within 60 days. We will notify you when the receipts will be available.

The details of your decision (charity chosen, donation amount, etc.) will remain private, and will not be revealed to other participants. You will be paid today by a third party who is not associated with this experiment. Neither the purpose of nor the details of this experiment will be made available to this person.
(T1: In today's experiment, <u>the name</u> of the charities you choose <u>will not</u> be revealed to anyone. At the end of the experiment, however, <u>the amount</u> you choose to donate to the final charity will be revealed to the other participants.)
(T2: In today's experiment, <u>the name</u> of the charities you choose <u>will not</u> be revealed to anyone. At the end of the experiment, however, <u>the amount</u> you choose to donate to the final charity will be revealed to the other participants.)

The next sections detail how the experiment will proceed.

Experiment Procedure

In today's experiment you will be asked to select three charities from a large database of charitable associations.

This list of charities comes from a large database prepared by Charity Navigator (www.charitynavigator.org), a non-profit organization that gathers information about national and international charitable organizations.

The list will be shown on a website preloaded on your computer. Please use the search box to search for your charities by name. If there are several results, you can try narrowing your search by adding more search terms, or you can use the links on the results table to page through the results. Please note, for charities known by their acronyms or abbreviations, you may need to type the full   name of the charity to locate it in your search.

If a charity is not in the database, you cannot choose it. You will have up to 20 minutes to choose your associations.

The database includes both national and international charitable organizations. Many national organizations operate both at a national and local level. You may choose to donate either to a local branch or to the national organization. You **may not** list the same charity twice; that is, if you choose a local branch, you must choose **only** one local branch, and you **may not** choose both a local branch and the national organization.

Along with these instructions, you have been provided with a worksheet for listing the three charities you choose. Please take a moment now to write your ID number at the top of this sheet.

We will now see how the experiment proceeds once you have chosen your 3 charities.

As soon as you have finished completing your list of three charities, we will provide you three stapled sheets, one per charitable organization. The sheets are identical, and contain a few questions about your knowledge and opinion of the charitable organization. Please write your ID number on <u>each sheet</u>. Use a different sheet for each charitable organization and write the name of the charity on it. Take your time to complete the questions accordingly.

Once you have completed the three short questionnaires, we will provide you an additional worksheet. On this sheet, you will be able to write down for each charity how much of the 25E$ you want to send. On this sheet, please be sure to print the charity names on each line, in the same order as you wrote them on your first worksheet. How much you wish to send to each charity should be written on the same line as that charity.

You may also indicate one charity as your favorite among the three by checking the box on that charity's line. As explained further, whether you choose to indicate a favorite charity or not, may affect how the final charity is randomly selected for final implementation.

Remember that the three decisions are independent, as only one of the three decisions will be implemented. This means that, for each and every charity, you should choose how to split the 25E$ between that charity and yourself.

As an example, suppose you have chosen charities "A","B", and "C".
When you choose how to split the 25E$ between you and charity "A", you should choose as if this is the only decision you have to make: in fact, if charity "<u>A"</u> is randomly selected for implementation at the end of the experiment, this will be the only decision relevant for payments.

Similarly, when you choose how to split the 25E$ between you and charity "<u>B"</u>, you should decide as if this is the only decision you have to make: again, if charity "B" is randomly selected at the end of the experiment, this will be the only decision relevant for payments.

The same reasoning applies to charity "C".

Please note that you are <u>not allowed</u> to cross off, rub out, delete or change a decision once is written on the answer sheet, so think carefully before writing. You may use scratch paper before writing your answers on the answer sheet.

Once you have made your decisions, we will collect the sheet with your list of charities (but <u>not</u> the one with your decisions) and provide you additional instructions.

If you have any questions, please raise your hand.

Final random selection of one decision for final payoffs

As explained above, at the end of the experiment only one decision among the three will be relevant for your final payoff; thus only one charity will receive the amount you decided to send.

While making your decisions, you will have been given the opportunity to indicate one charity as your favorite.

Your answer will determine how the decisions are implemented.

We will flip one coin in front of everyone.

1. If you indicated a favorite, and if the outcome of the coin-flip is HEAD, then the decision associated with your favorite charity will be used to calculate final payment.
2. Regardless of whether you indicated a favorite, if the outcome of the coin-flip is TAIL, then the decision implemented will be determined by the result of a die roll (we will divide the roll by two and round up to determine which decision will be implemented, by number). Note that with the die roll your favorite charity may still be randomly selected for implementation.

If you did not indicate a favorite, you may ignore the coin flip, but we will still roll the die. In this case, all three decisions will have the same probability of being implemented.

This means that when you indicate a favorite, with 50% of probability the favorite charity will be selected by the coin flip. If it's not selected by the coin flip, there is still 33.3% of probability that the favorite will be selected by the die roll. The overall probability for the favorite charity is thus 66.5%.

Please note that you will not be allowed to cross off, rub out, delete or change a selection once is written on the paper, so think carefully before writing.

Please raise your hand if you have any questions.

CHARITIES' LIST SHEET

Write your ID NUMBER: _____

My Charities List

Please write below the list of 3 charities of your choice (PRINT your answer in a clear and legible way)

IMPORTANT: Please always use the same order for the following sheets.

1. _____
2. _____
3. _____

QUESTIONNAIRE ABOUT CHARITIES (3 identical sheets, here only 1st one shown)

Please write your ID NUMBER:

_____

PRINT the name of the **1st charity** on your list:

_____

_____
_____

Please answer the following questions by circling the corresponding number:

1) Think about the mission of this charity. From your perspective, how important are the issues that this charity addresses?

1	2	3	4	5	6	7
Not at all Important			Important			Extremely Important

2) Think about the mission of this charity. In your opinion, how important are, for your family and friends, the issues that this charity addresses?

1	2	3	4	5	6	7
Not at all Important			Important			Extremely Important

3) Have any of your family members ever donated time or money to this charity?

YES   NO   DON'T KNOW

4) For the following statement, please indicate whether you agree or not: "This charity serves its cause with efficiency."

1	2	3	4	5
Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

FIRST COMPREHENSION QUIZ

Write your ID number : _____

QUIZ 1

Please answer the following questions

1) You have received 25E$:

   TRUE   FALSE

2) For each of the three decisions, you can send any amount between 0E$ and 25E$

   TRUE   FALSE

3) The decision you take for one of the charities affects your choice for the other two decisions:

   TRUE   FALSE

4) Suppose you choose to send 25E$ to the **first** charity: how many E$ you will have at your disposal to split between you and the **second** charity? _____

5) Suppose you choose to send 5E$ to the **second** charity: how many E$ will you have at your disposal to split between you and the **third** charity?_____

6) Other participants will be told the name of the organizations you sent money to.

   TRUE   FALSE

7) Other participants will be told the amount you decided to send to the organization randomly selected for implementation.

   TRUE   FALSE

8) If you choose to not select a favorite charity, which is the probability that, for example, the decision associated with the second charity will be implemented for final payments?

_____

DECISION SHEET PHASE 1

Write your ID NUMBER: _____

Print below the name of the charities you have chosen (in the same order as you wrote them on the previous sheet. Write down how much you want to send to the charity on the same line as that charity. Remember that the 3 decisions are independent.

 Among the charities you have listed, do you have a charity you want to indicate as your favorite? If so, please check the box corresponding to the charity you want to indicate as favorite. Please note that you won't be allowed to change your selection once done, so think carefully.

Charity	Amount to the Charity	Favorite?
1._____	_____	☐
2._____	_____	☐
3_____	_____	☐

INSTRUCTIONS PHASE 2

The first part of the experiment is over.

In this second (and last) part we will provide you additional information about the charities you have selected.

At the end of this phase you will have the possibility to decide again how to split 25E\$ between you and each of the charities you have already selected. You may rewrite the same amount you have chosen before, or change it. These last three decisions will be final, and will be the ones relevant for payments.

Please note that, as before, all the details of your decisions will be kept private, including the names of the charities you have selected, as well as your donation amounts, which was your favorite, etc.
(T1: Please note that, as before, the only information that will be publicly available is the amount you choose to donate to the final, randomly selected charity. All other information, such as the charity's name, efficiency rating, your favorite, etc. will be kept private.)
(T2 *(This part is at the end of phase 2 instructions)* : Please note that the exact efficiency of the final charity that will be implemented will be made public, together, as said before, with the amount you choose to donate to the final charity. This means that you will be asked to declare how many E\$ you choose to give to the final charity, and what is the exact efficiency score for that charity. All other information, such as the charity's name, your favorite, other two decisions etc. will be kept private.)

Decision 2

You will now receive instructions about a homogenized measure commonly used to evaluate charitable organizations' efficiency. The instructions will include an explanation on how to interpret those measures.

Once you finish reading, we will distribute 3 sheets. You will be asked to guess (on one sheet per charity) the **actual efficiency for each charity you have chosen**. If your guess is within +/- 5% of the true value of the charity, you will receive an additional 6E\$. The decision implemented at the end of the experiment determines which guess will be checked and, if correct, paid.

You will also be asked to give an estimate of the probability that your charity's actual efficiency is within five ranges, or quintiles. These quintiles begin at 0 and increase in 20 percent increments to 100 percent (0 to 20 percent, 20 to 40 percent, etc.). Quintiles are ranges over the program expenses variable. The probability you assign to each quintile represents how likely you think it is that your charity's actual program expenses falls into that range. The quintiles should not be confused with your probability that the charity's actual efficiency falls within that quintile. The sum of all the assigned probabilities must be 100 percent.

Once you write your guesses we will give you back your summary decision sheet and we will reveal to you the exact efficiency of all the charities you have listed.

These ratings are publically available from CharityNavigator at www.charitynavigator.org.

You will then have the opportunity to decide again how to split the 25E$ between you and each charity. Again, those three decisions are independent.

<u>These last three decisions you are about to make are the only ones that will be used to calculate your payments.</u>

EXPLANATION OF THE EFFICIENCY MEASURE

Charity Efficiency

CharityNavigator (CN) is an independent, non-profit charity evaluator. CN uses a variety of data to rank charities based on their finances and transparency.

Today, we will show you a single, homogenous measure that reports how much a charity spends on the programs and services that charity provides. The measure is called "Program Expenses", and you can find a detailed definition at the CN website.

Let's illustrate what we mean by "Program Expenses". Say a non-profit named Charity Blue had a budget last year of $5 million. Charity Blue's mission is to provide low-cost blue pencils to needy school children. If Charity Blue spent $1 million on fundraising and administrative expenses, and $4 million on providing blue-colored pencils to needy schools, Charity Blue spent 80% on program expenses. In other words, 80% of the charity's resources were spent pursuing its stated mission.

We will ask you to guess the efficiency (percentage spent on program expenses) of the charities you have chosen. We will also ask you to guess the average efficiency of related charities.

GUESSING PHASE (3 identical sheets, here only 1st shown)

Guess Phase

Please write your ID NUMBER: _____

PRINT the name of your **1ˢᵗ charity**: _____

_____

1) What is your best guess for the exact efficiency score of your charity?

Efficiency Score: _____%

2) Now, what is your best guess for how likely you think it is that the exact efficiency score of your charity falls in each of the five intervals of efficiency written below. Please note that the sum of the 5 probabilities should add up to 100%.

						Total
Likelihood						
Interval of efficiency	0%-20%	21%-40%	41%-60%	60%-80%	81%-100%	-

3) Which label do you think best describes the sector where your charity operates?

Label that best describes the sector: _____

REMINDER

Final random selection of one decision for final payoffs

As explained above, at the end of the experiment only one decision among the three will be relevant for your final payoff; thus only one charity will receive the amount you decided to send.

You have been asked the following question:

> *Among the charities you have listed, do you have a charity you want to indicate as your favorite?*

Your answer will determine how the decisions are implemented.

We will flip one coin in front of everyone.

1. If you answered **"yes"***,* and if the outcome of the coin-flip is HEAD, then the decision associated with your favorite charity will be used to calculate final payment.
2. Regardless of your answer, if the outcome of the coin-flip is TAIL, then the decision implemented will be determined by the result of a die roll (we will divide the roll by two and round up to determine which decision will be implemented, by number). Note that with the die roll your favorite charity may still be randomly selected for implementation.

If your answer is **"no"** (meaning that you do not want to indicate a favorite), you may ignore the coin flip, but we will still roll the die. In this case, all three decisions will have the same probability of being implemented.

This means that when you indicate a favorite, with 50% of probability the favorite charity will be selected by the coin flip. If it's not selected by the coin flip, there is still 33.3% of probability that the favorite will be selected by the die roll. The overall probability for the favorite charity is thus 66.5%.

Please note that you will not be allowed to cross off, rub out, delete or change a selection once is written on the paper, so think carefully before writing.

We will first toss a coin. If you have indicated a favorite charity this coin flip is relevant for you. If the outcome is HEAD, please circle the name of your favorite charity on the answer sheet.
If the outcome is TAIL, please wait for the die roll.

We will then roll the die.

If you have indicated a favorite charity and the outcome of the coin toss was TAIL, please circle the name of the charity corresponding to the number of the die roll.

If you did not indicate a favorite charity, please circle the name of the charity corresponding to the number of the die roll.

Afterward, please place three documents into the large envelope at your workstation. Both decision sheet (your initial donation and your final donation) as well as the document labeled "Guess Phase" go into this envelope. Please write "CHARITY" on the front of the envelope, near your ID number.

Print your ID number and write "ME" on the other envelope. You may use this envelope to collect your payment. You are advised to keep this envelope for future reference, since you will need your subject ID to check donation receipts in the future.

We will now distribute two short questionnaires. Once you have completed them, you may exit the room and receive payment.

A third person, who is unaware of the content of the experiment, will pay you in private.

Essay 3

Hidden Costs of Control under Aligned Monetary Interests[1]

3.1   Introduction

Traditional agency theory predicts that when agents' and principals' monetary interests are not aligned, agents will try to act in their own interest, thus the exercise of control represents a possible way for principals to limit agents' opportunistic decisions. Recent empirical findings however have pointed out that for principals, controlling subordinates may come with hidden costs (see i.e. Frey 1993, Fehr and List 2003, Falk and Kosfeld 2006, Charness and Levine 2007, Dickinson and Villeval 2008, Schneidler 2010).

In the present study we depart from the classic framework of conflicting monetary interests, and explore whether exercising control over agents also involves hidden costs in the presence of *aligned* monetary interests. In particular we investigate whether *personal rules* (implemented case by case through direct monitoring) and *impersonal rules* (imposed ex-ante by principals to every type of agent) impact agents' decisions in the same way, the underlying hypothesis being that the signaling value of principals' authority matters for agents.

---

[1]This chapter is a joint work with Jason Aimone.

Indeed in many naturally occurring environments agents' and principals' monetary interests are aligned. Within companies and firms, often both parties share a vested interest in extracting surplus from a third party (i.e. a client). In a consulting or law firm, for instance, both partners and shareholders get a benefit from extracting the maximum payoff from a client. Similarly, high quality restaurants profit both waiters' and owner's revenues. A franchiser and a franchisee, again, may share in preset percentages the benefits of a successful business. In these environments monetary interests are aligned, but relative authority, accountability, and social distance to the customers are not [2].

We hypothesize that these factors are sufficient to generate failure of authority. Borrowing terminology from Aghion and Tirole (1997), principals often hold formal authority (the right to decide), while agents have real authority (the effective control over decisions). Agents typically deal directly with customers, providing the required services (e.g. exercising real authority over firms' decisions), while principals design internal institutions and enforce formal constraints on the way the service has to be provided (i.e. they exercise formal authority)[3]. In some firms, restrictive rules imposed upon subordinates are *general* and *impersonal*. In others, principals monitor their subordinates directly, imposing *personal* or *ad hoc* restrictions whenever deemed necessary.

If monetary interests drive the interaction between firms (principals and agents) and clients, *whether* and *how* a principal exercises control over agents should not influence how the latter produces value for the customer, as the exercise of formal authority does not conflict with agent's payoffs. However, if agents care about whether a principal's restriction represents a *personal critique* of[4] or a *voluntary interference* with their own actions, or an *impersonal rule* that they know is set independently from their behavior, then agents may in the former case be willing to sacrifice resources to punish a principal, generating hidden costs of control. The intuition is simple: under impersonal rules

---

[2]Nor necessarily are factors like altruism, inequity aversion, or alternative other-regarding preferences.

[3]Minimum quality standards, for instance, can be imposed to an agent in order to guarantee a lower bound of service's quality. Similarly, principals may want to set maximum quality standards; this may come in useful to avoid congestion or to extract the maximum value from every economic exchange (i.e. imposing a time limit/goal to customer service agents for calls in a call center).

[4]And thus an implicit acknowledgment of agent's "type" (i.e. generous or greedy).

there is no direct connection between agents' preferences for reciprocity (and therefore choices), and principals' restrictive behavior. Being restricted thus does not imply any direct conflict between the two actors. When formal authority is instead exercised through monitoring and personal restrictions, the agent receives two signals: that the principal disagrees with the agent's choice, and that the principal is directly taking actions to limit agent's real authority. If agents have preferences over the signaling value of restrictions, then they should oppose restrictions more fiercely whenever they are personal.

We predict thus that within a firm, an agent would reciprocate clients' trust (exercise real authority) very differently, depending on whether principal's restrictions (the exercise of formal authority) represent a sign of *distrust* or *disagreement* (personal rule) or a mere signal of principal's preferences (impersonal rule). To do so we use a modified trust game that allows for the presence of a third player, a principal, in addition to the standard, investor (a client), and trustee (an agent). More specifically, the client (a.k.a. investor) chooses to trust an agent (a.k.a. trustee) by sending him part of his initial endowment. Agents make two reciprocation decisions in the game: a first "free" decision (in the sense that he can choose to send back any amount to the client), and a second decision, which may be constrained by the principal. After agent's first choice, the principal decides whether or not to exercise formal authority by restricting the maximum amount the agent can send back to the client in his second reciprocation decision. The agent then decides again how much to send back to the client under this restriction (or lack of a restriction)[5].

This simple framework is implemented under two different treatments: (i) PERSONAL Treatment: principal's exercise of control is *informed*, in the sense that principals, before choosing to control or delegate, do observe agents' type (e.g. observe the first reciprocation decision made by the agent therefore imposing personal restrictions) ;(ii) IMPERSONAL Treatment: principals' choice is *uninformed*, meaning that their choice

---

[5]Agent's second decision thus can be either "free", in the sense that he again can send back any amount he wants to the client, or "restricted", meaning that this time the maximum amount he can send back is limited. A random draw determines whether the first or second agent decision is used for payments.

only expresses their own preferences for reciprocation (e.g. formal rules are imposed under the "veil of ignorance", that is independently from agents' type).

We hypothesized that, controlling for agents' initial trustworthiness (e.g. 1st reciprocation decision), principals' restrictions would be more effective in reducing clients' payoffs when that exercise of formal authority is made using an *impersonal rule*. Our results suggest that indeed the way formal authority is implemented, either personally or impersonally, matters. Controlling for agents' initial reciprocation to clients, we find that restricted agents are relatively more generous with the client when the restriction is *personal* (principal informed) compared to when the restriction is *impersonal* (principal uninformed). The latter result provides evidence that, even with aligned monetary interests, hidden costs of control may exist whenever restrictions represent a direct opposition to agents' actions, and that these costs can be mitigated by exercising control via impersonal rules. More explicitly, agents appear to be willing to costly punish the principal (by sending more to the client) whenever the restriction imposed by the latter represents an ad hoc or intentional restriction (e.g. I restrict you because *you were* doing wrong); when restrictions instead only signal principal's preferences (e.g. I restrict you because *I am* greedy), agents significantly reduce their trustworthiness towards their client.

Surprisingly however, despite ex-post relatively higher payoffs, we find that principals tend to recur to formal authority relatively less often under an impersonal rule system. We speculate that principals may be reluctant to be "first movers" on unfair decisions, and prefer to signal their greed only when strictly necessary: if in fact principals have preferences for appearing fair and trustworthy (although not necessarily for being fair and trustworthy)(see i.e. Dana, Weber, Kuang 2007; Hao and Houser 2011), then personal restrictions allow them to signal their greed only when necessary (e.g. when the agent is too generous); on the other hand imposing a restricting impersonal rule signals their type (e.g. untrustworthy) also to agents that "didn't need to know", as they are already not overly generous. The remainder of the paper is organized as follows:

section 3.2 presents related literature; section 3.3 details our design, section 3.4 presents results and section 3.5 concludes.

3.2 Literature

The literature on delegation and authority within organizations is rich and vast[6].From a theory perspective, several models have been proposed to explore agent-principal relationships[7]. Closely related to our work, Aghion and Tirole (1997), from which we borrow the notion of formal and real authority, develop a theory of the allocation of formal authority (the right to decide) and real authority (the effective control over decisions) within organizations, showing that factors like overload, lenient rules, urgency of decisions, and reputation do increase subordinates' real authority. Ellingsen and Johannesson (2008) investigated the link between signaling value of restrictions and agent's pride, showing that control systems erode agents' morale by signaling them that the principal is not worth impressing: as a consequence, principals' restrictions lead agents to shade and shirk for reasons other that pure monetary interests.

On the empirical side, several studies have confirmed theoretic intuitions about the importance of non-monetary factors in hierarchies. Most findings converge to the idea that even when the exercise of control has zero costs and asymmetric information is not an issue, limiting agents' behavior might be a suboptimal strategy for principals if this is perceived by agents as a sign of distrust (Falk and Kosfeld 2006). In particular, it has been shown both in the lab and in the field that principals avoid imposing a minimum (binding) threshold of effort to workers, and that this is ex-post efficient. These hidden costs of distrust seem mitigated by whether authority is perceived as legitimate (Schnedler 2010). More generally, the effect of authority on workers' performances appears to

---

[6]See i.e. Miller (1993) for a general discussion on approaches to agency problems in economics and sociology.

[7]Armstrong, Vickers (2010), and Zhao (2007) explored optimal delegation with agent's discretionary search possibilities. Bernheim (1994) studied the relationship between authority acceptance and status maintenance. Aghion, Dewatripont and Rey (2004) study situations where a principal can transfer control to an agent but cannot commit herself to do so. Hart and Holmstrom (2010) explored the role of decision rights on ex-post conflicts and scope of the firm.

be mediated by complementary forces: social distance, the link between employer's payoffs and employees' performances, and sense of agency, are all important factors that concur to define the extent to which authority constraints may have detrimental or positive effects (see i.e. Charness and Levine 2000; Fehr and List 2003; Dickinson and Villeval 2008). As recent papers have pointed out, there are reasons other than agents' reaction to authority that make delegation optimal: Bartling and Fischbacher (2012), and Oexl and Grossman (2012) have shown that delegation can be an effective instrument principals can use to shift the blame (and punishment) on their employees for unfair outcomes received by third parties[8]. These results appear in line with more general evidence on people desire to appear fair (Dana, Weber, Kuang 2007) or honest (Akerlof 1983; Hao and Houser 2011). Despite restrictions often backfire, individuals appear to dislike delegating subordinates: Fehr, Herz and Wilkening (2012), for instance, show that individuals tend to retain authority, even when the opposite would be in their best interest. An explanation for why this may be the case, is betrayal aversion. Recent studies in fact (see Aimone and Houser 2011,2012a, 2012b) have demonstrated that individuals' distrust largely depends on how betrayal is perpetrated, showing that betrayal aversion depends on the intentionality behind betrayal. Although our paper investigate a complementary question (e.g. what determines willingness to betray), we share with these papers the attention on personal versus impersonal hierarchies' organization.

Building on this literature, our paper provides two important contributions: we explore the hidden costs of control in environments where principals' and agents' monetary interests are aligned, and we do so in the context of an investment game where in all treatments clients are unaware of the active role of the principal. This allows us to understand how, in life-likely settings, agents respond to different mechanisms of control enforcement when signaling outside the firm is not possible, avoiding confounding

---

[8]These studies differ from ours in the followings: (i) principal's delegation towards agents, and fairness towards third parties is studied using a dictator game framework ; (ii) third parties are always aware of who shall be blamed.

effects as principals' image concerns in face of clients, and agents' strategic hedging on principals' decisions (i.e. "it was my boss, not me!") .

3.3   Experimental Design

3.3.1   Overview

90 subjects were recruited at George Mason University from March to May 2012 to participate in a modified investment game. The experiment was programmed and conducted using the software z-tree (Fischbacher 2007). Mean age was 22.4, with 43% females, 57% males. Upon arrival at the lab, subjects are randomly assigned to a role: agent, principal, or client, which they know they will keep for the duration of the experiment . Subjects read instructions (appendix A) and are told that for every 3 experimental currency units (ECU) they earn during the experiment they receive 2 US Dollars in cash at the end of the experiment. The experiment lasts four rounds, known in advance to each subject. Each participant receives 10 ECU as initial endowment at the beginning of a round. There are two main phases of the experiment, which we will call the standard phase and the delegation phase, and a belief elicitation phase. The delegation phase is a private interaction between the agent and principal and is thus unknown to the client. Agents and principals know in advance that there is a 400% chance that decisions from the delegation phase will not be implemented. All other information is common information to all three subject types.

3.3.2   Standard Phase

First, as in a standard investment game, the client decides whether to invest 2, 6, or 10 ECU to the agent. This amount is multiplied by six during the transfer, and is observed by both the agent and the principal. The agent then chooses how many ECU to transfer back to the client: the agent can in this phase choose to return any amount between zero and six times the amount transferred by the client . This amount is not multiplied again. Henceforth we refer to this unrestricted decision as decision "Free1".

3.3.3 Delegation Phase

After the agent makes his standard phase decision, Free1, the computer randomly draws a restriction value uniformly distributed between half the amount initially sent by the client (non-multiplied) and two times the amount initially sent by the client. Both agents and principals observe this realization. The principal then decides whether to impose this restriction on the agent's decision. If the principal chooses to not impose the restriction, the agent makes a new (unrestricted) return decision between zero ECU and six times the amount sent by client. If the principal instead chooses to impose the restriction, the agent's delegation phase decision is constrained and they can only choose to send an amount between zero ECU and the value of the randomly chosen restriction. Henceforth we refer to an unrestricted agent's delegation phase decision as "Free2" and a restricted agent's delegation phase decision as "Restricted2".

3.3.4 Beliefs' elicitation phase

After the agent makes their delegation phase decision (but before they learn what others did), we elicit all participants' beliefs about other players' expectations. Clients are asked what they expect back as a return. Both agents and principals are asked what they think the client expects back. Principals are additionally asked what they expect the agent chose to return to the client in the delegation phase. Agents are also asked what they think the principal expects the agent to return in the delegation phase. All decisions are monetarily incentivized[9]. At the end of the period, a random draw determines whether agents' standard phase decision (Free1) or delegation phase decision (Free2 or Restricted2) will be used for payments: with 40% of agents' standard phase decisions used for payments, and with 60% of agents' delegation phase decision used. Whatever is not returned to the client is split evenly between the agent and the principal. This simple setup is repeated four times, with random perfect strangers' matching, so that no subject ever plays with the same counterpart (in any role) twice.

---

[9]Each player knows that if his guess is within +/- 3ECU from the true value, he will receive additional money. See appendix A for detailed questions.

3.3.5   Treatments: PERSONAL vs IMPERSONAL restrictions

Our experiment consists of two treatments, which differ only in what the principal observes. The PERSONAL treatment is conducted just as described above, and the principal perfectly observes their own agent's Free1 decision prior to the delegation phase. This is common knowledge between agents and principals. This means that before choosing whether to impose or not the restriction on the agent, the principal observes how much the agent chose to send to the client when unrestricted, e.g. Free1.

In the IMPERSONAL treatment instead, while the principal still observes the client's choice they do not observe the agent's Free1 decision. This is common knowledge between agents and principals. This means that when choosing whether to impose the restriction, the principal does not know whether the agent chose a high or low return amount. The principal only observes the final return along with the client after the belief elicitation phase. 60 subjects participated in the PERSONAL treatment and 30 in IMPERSONAL. Our results are organized as follows: in section 3.4.1 we provide summary results on players' decisions. Section 3.4.1.1 analyzes in details agents' behavior, and section 3.4.1.2 presents results on principals' decisions.

3.3.6   Discussion of design

Before proceeding to results, we wish to highlight few features of our design. First, in our experiment clients have no knowledge of the role the principal has in the game. This restriction on client's information set has two advantages: (i) the accountability of agents with respect to clients' payoffs is always fix and common knowledge, ruling out agents' image concerns as an explanation for their decisions[10]; (ii) it mimics a common feature in real firms: internal hierarchical choices are rarely observed by customers. Second, in our experiment the effective value of the restriction is randomly drawn from a uniform distribution between half and two times the amount sent by the client (and this is known

---

[10]If in fact clients were aware of principals' role, selfish agents who dislike being seen as such, may systematically betray clients' trust whenever principals impose restrictions (e.g. delegating accountability), and send back high amounts whenever is clear that they have been delegated.

by both agents and principals). This procedure increases variability in our data by limiting reference point effects and therefore systematic selection of certain values from the principals (i.e. principals always choosing a certain number or percentage of clients' investment). Finally, we chose to have nature selecting agents' final decision (either Free1 or Free2/Restricted2). By doing so we make sure agents' Free1 decisions are not taken randomly (as may be the case if agents knew Free1 being always an "interim" decision), and we capture a realistic feature of monitoring (e.g. PERSONAL treatment) and general rules (e.g. IMPERSONAL treatment): although hard, the probability for agents to get round them is always positive.

3.4   Results

3.4.1   General Results

Table 1 summarizes our main descriptive statistics.

Treatments	N	mean	sd	As % of clients' investment
PERSONAL				
Client Decision	80	5.90	3.12	-
Client Expectations	80	6.66	4.72	-
Agents' FREE1	80	6.46	9.43	109.4%
Agents' FREE2	28	2.89	4.63	48.9%
Agents' Restricted2	52	3.31	4.63	56.1%
Principals' Restrictions (%Yes)	80	0.65	0.48	-
IMPERSONAL				
Client Decision	40	7.90	2.72	-
Client Expectations	40	11.13	8.98	-
Agents' FREE1	40	7.03	8.00	88.9%
Agents' FREE2	22	7.27	8.45	92.0%
Agents' Restricted2	18	1.94	3.13	24.5%
Principals' Restrictions (%Yes)	40	0.45	0.50	-

Table 3.1: Summary statistics

As in many investment and trust games, trust in this environment does not pay for clients (see i.e. Berg, Dickhaut and McCabe 1995). Clients sent on average 5.9 ECU in treatment PERSONAL and 7.9 ECU in treatment IMPERSONAL (the difference between treatment is significant, p=0.000), and they received back 4.5 ECU and 5.8 ECU respectively[11]. Clients left the lab with average earnings of 3.02 US $ (excluding 5$ show up fee) in PERSONAL, and 4.03 US $ in IMPERSONAL. The difference across treatment is not significant (p=0.433).). In terms of expectations, in both treatments clients expected to receive more than what they sent. For instance, in PERSONAL they were expecting

---

[11]Unless mentioned otherwise, all reported p-values come from two-tailed Mann-Whitney tests.

back on average 6.66ECU (115% of what they sent), and in IMPERSONAL 11.1 ECU (164% of what they sent)(p=0.003). Agents and principals, who split evenly what was not transferred to the client, earned, respectively, in PERSONAL and IMPERSONAL 12.12 US\$ and 15.7 US \$ on average[12]. Note, that in the IMPERSONAL treatment, all players received relatively higher payoffs. This can be attributed to the significantly higher investments by clients which increase the total size of the pie. Since both treatments were identical from the clients' perspective the increased investments must be attributed to random chance.

Overall, in principals decisions, we observed that principals restrict significantly less under IMPERSONAL than PERSONAL (respectively 45% and 65% of the times, p=0.037). Across treatments, agents' initial decisions (Free1) are not statistically different (p=0.379) nor are agents' Restricted2 decisions (p=0.173). However, agents' unrestricted transfers to clients (Free2) are significantly higher under IMPERSONAL than PERSONAL (p=0.030). While the results regarding FREE1 decisions are fully comparable, one has to be careful in interpreting differences regarding agents' behavior in the second delegation phase as agents left free and agents left restricted in the two treatments are in fact not immediately comparable, because what the principal observes in the standard phase determines selection in PERSONAL but not in IMPERSONAL. This means that while agents in PERSONAL are sorted by principals' decision according to their type (e.g. "generous" or "greedy"), in IMPERSONAL different types of agents are pooled together[13]. As discussed in the next section, we control for these selection differences by restricting observations under IMPERSONAL treatment to comparable groups. In other-words we compare decisions of generous types to generous types, and greedy types to greedy types.

---

[12]As participants were each randomly rematched four times , though with perfect strangers procedure, one may be concerned by time trends. We find that this is not the case. Clients' investments do not decrease or increase over periods (for PERSONAL and IMPERSONAL, p=0.243 and p=0.781 respectively; Jonkeree-Terpsa tests). Agents' FREE1 decisions do not differ across periods (p=0.662 and p=0.982), nor do agents' FREE2 (p=0.600 and p=0.249) or Restricted2 (p=0.651 and p=0.288). Finally, principals' choice between restricting also shows no significant round effect (p=0.530 and p=0.579).

[13]This raises thus the average transfer Free2 (some generous agents are left free), and reduces Restricted2 (some greedy agents are restricted) under IMPERSONAL treatment.

The next section analyzes develops on this intuition further and analyzes in more detail the behavior of agents and principals.

3.4.1.1 Agents' behavior

In what follows, we compare agents' decisions across phases and treatments as a percentage of what the client sent. As in fact all agents' and principals' decisions are conditional upon what clients sent to them, we use agents' and principals' decision variables expressed as a ratio with respect to client's decision[14]. This is particularly important given the significant difference in what clients send between treatments (on average 2 ECU more in our IMPERSONAL treatment, p=0.012).

We first look at differences, in each treatment, between the FREE1 decision and Restricted2. Not surprisingly we find a significant reduction in transfers in both treatments for "generous agents": average reduction of the return rate of client's investment after restriction is 129% for agents in PERSONAL (p=0.000), and 110% for agents in IMPERSONAL (p=0.001)[15]. As the value of the restriction is always randomly drawn between half and two times the amount sent by the client, it is not surprising that in both treatments generous agents are forced to reduce their return to the client. Of interests though is, controlling for the generosity of the agent in phase 1, whether generosity after a restriction is the same or different depending upon whether the restriction was personal in nature or impersonal.

To appropriately compare like subjects, we use a simple strategy: we look at a subsample of restricted agents in IMPERSONAL treatment that behave in the first phase (Free1) as those that are restricted in PERSONAL treatment. This means that we include only "generous agents" from IMPERSONAL treatment who in Free1 transferred back at least what the client sent them (non-multiplied). In other words, we do not consider from

---

[14]Results, however, are not qualitatively different than using absolute values.

[15]Even by considering the whole pool from treatment HIDDEN however, the difference in IMPERSONAL treatment is still significant (41% decrease in client's return rate from Free1 to Restricted2; p=0.01), as it is in IMPERSONAL.

the restricted IMPERSONAL pool those subjects that would have not been restricted if their Free1 had been observed by their principal (as was the case in PERSONAL).

Table 3.2 reports a simple OLS with clustered standard errors showing the effects of treatments, previous decisions and beliefs on agents' Restricted2 return rate ($\frac{Restricted2}{CLIENTDECISION}$). Decisions taken in Free1 positively affect restricted decisions in subsequent phase (p=0.083); moving from personal to impersonal restrictions reduces the return rate for clients from restricted agents (p=0.059). Finally, restricted agents' beliefs over clients' expectations reduce their Restricted2 transfers (p=0.038), while beliefs over what the principal expects them to do increases it (p=0.019). The value of the randomly drawn restriction has instead no significant effect on Restricted2 decisions (p=0.238)[16].

	(1)
VARIABLES	Restricted2 return Rate
Free1 return rate	0.197*
	(0.105)
Treatment (0=personal;1=impersonal)	-0.346*
	(0.168)
Agent belief on client expectation	-0.0128**
	(0.00559)
Agent belief on princ. expectation	0.0702**
	(0.0265)
Restriction's value	0.0200
	(0.0162)
Constant	-0.0324
	(0.231)
Observations	34
R-squared	0.647

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.2: Main effects on Restricted2/Client Decision

---

[16]It must be kept in mind however that restriction values are uniformly drawn from half to two times the amounts sent by clients. It is likely thus that the effect this variable has is captured by the effect of clients' decisions. Both effects however, go in the same direction.

Result 1: *Controlling for initial decisions (Free1), when restricted by their principals agents return to clients more when restrictions are personal than when are impersonal.*

Figure 1 shows the distribution of Restricted2 decisions for agents that sent back to the client in their decision Free1 at least what they received (once again, not multiplied).

Figure 3.1: Restricted2 return rates of "generous" agents (Free1 return rate at least equal 1) by treatments

The treatment effect is evident in the histogram of Restricted2 decisions of generous agents in the PERSONAL and IMPERSONAL treatments. Given similar levels of initial reciprocity, agents that face personal restrictions send relatively more to their clients in their second decision (Restricted2) than counterparts who face impersonal restrictions (p=0.032)[17] . It is important to notice that even without restricting observations to comparable samples, the difference is still significant (on all observations, p=0.080).

We interpret this result as evidence for hidden costs of controlling agents, even under aligned monetary interests. In treatment PERSONAL, the very fact an agent receives

---

[17]Mean value of Restricted2 for PERSONAL treatment is 6.07(s.d. 4.96), for IMPERSONAL 3.42(s.d. 2.93)

a restriction implies that the principal explicitly believes the agent was too generous in their decision in the standard phase. If agents dislike being directly criticized, then sending back to the client a relatively higher amount represents a form of effective costly punishment to the principal. Note that the cost of the punishment is not only on the side of the principal, but on the part of the agent as well, since returning more to the client reduces earnings to both the principal and the agent.

In both treatments some agents were, of course, not restricted. We next check differences in behavior of agents who received delegated authority across treatments. To do so we apply the same idea we used for restricted agents: we consider only agents in the IMPERSONAL treatment whose Free1 decisions would have not been restricted by principals had they been made in the PERSONAL treatment. We therefore keep only "greedy" agents, those that sent back to the client in Free1 less than what the client originally sent.Opposite to what we saw with generous agent types, "greedy" types became relatively more greedy when delegation (not being restricted) was personal than when it was impersonal (p=0.064). This result represents additional indirect evidence that agents care about the signaling value of the restriction, and are therefore willing to sacrifice resources to punish principals that impose ad hoc restrictions (generating thus hidden cost of control). When agents are left free, in fact, the situation is symmetrical: delegation under PERSONAL treatment here takes the form of an implicit acknowledgement that the agent's actions are not overly generous. In IMPERSONAL, delegation is not acknowledgement of the greediness of the agent. If agents give value to informed principals' decision, we should expect then that a personal delegation would lead to lower reciprocity towards the client than an impersonal delegation, as in the former case agent's pride and principal's decision are aligned[18].

---

[18]It must be noticed that in our design, principals cannot restrict agents with minimum amounts to be sent back to clients. This means that from an agent's perspective, being delegated under PERSONAL treatment does not represent an endorsement with certainty: it may be the case, in fact, that the principal wanted the agent to be more generous, but he had no means to signal it. For the sake of our analysis however, what matters is the comparison between treatments of what a delegation signals: in our IMPERSONAL treatment, a delegating principal is only signaling his type, meaning that under PERSONAL, the (subjective) likelihood of being endorsed (via delegation) is, although not 100%, still higher than under IMPERSONAL.

Before turning to principals' behavior, we report agents' second-order beliefs over other players' expectations. As mentioned, beliefs' elicitation took place after principals' decisions, before agents' second choice. On average, agents in PERSONAL treatments believe clients expect more than agents in IMPERSONAL do (p=0.050). If we divide restricted and delegated agents, we discover that this result is driven by agents who are restricted. The beliefs of agents to whom authority is delegated do not differ across treatments (p=0.858). Restricted agents, however, believe clients are expecting more when restrictions are personal rather than impersonal (p=0.026). When we look at what agents believe the principal is expecting them to do in the second decision, we find essentially the same results. Overall, agents in PERSONAL believe principals are expecting them to be relatively more generous compared to agents beliefs about principals in IMPERSONAL (p=0.044). This result is driven, as before, by agents who end up being restricted (p= 0.036). When authority is delegated to agents, beliefs are not statistically different across treatments (p=0.372). The difference in beliefs that appears in the second order expectations of restricted generous agents provides a glimpse at the internal rationale for why these agents are seen returning relatively more to clients. Agents are fully aware that clients are unaware of the ability of principals to restrict return rates in both the PERSONAL and IMPERSONAL treatments, therefore knowledge of how such restrictions are decided upon, personally or impersonally, should play no role in the formation of agent's second order beliefs as we see is the case. The thought that "this is what is expected of me" may play a crucial role in helping the restricted agents overcome their own monetary incentives and decide to return more when restrictions are PERSONAL.

3.4.1.2 Principals' behavior

Table 3 summarizes summary statistics for principals' decisions and beliefs.

	Treatments			
	Visible		**Hidden**	
	No Restr.	**Restr.**	**No Restr.**	**Restr.**
Principals' Choice (%)	35%	65%	55%	45%
		(47)		(50)
Princ. belief over clients' expectations	154%	120%	229%	116%
	(184)	(101)	(190)	(91)
Princ. expected rate returned by agent	111%	71%	208%	0.58
	(137)	(93)	(201)	(47)

Note: s.d. in parenthesis

Table 3.3: Principals' summary statistics

Result 2: *Principals are more likely to enforce personal restrictions than impersonal rules, even though expectations on agents' behavior do not justify this difference.*

By looking at table 3 one can notice two things: principals tend to restrict less under IMPERSONAL than PERSONAL (45% to 65% respectively, p=0.037) despite expectations that agents are more generous in phase 2 under IMPERSONAL than under PERSONAL (p=0.048)[19].

Taken together, these statistics come as a puzzle. Across treatments, agents' average returns to the client are the same in phase 1 (p=0.548). Principals indeed do observe agents' phase 1 decisions in PERSONAL, but do not under IMPERSONAL. This means that in PERSONAL what principals expect agents to send back to clients in phase 2 is conditional upon priors, agents' observed actions, and observed clients' investments. Conversely, in IMPERSONAL, expectations are conditional only upon priors and the observed investments of clients. If, as seen, principals in IMPERSONAL incorrectly

---

[19]Regardless of whether they restrict or not

believe that agents left free in phase 2 will be relatively too generous (compared to observed agents' generosity under PERSONAL), why would they restrict them less?

A possible explanation is reciprocity towards clients. As discussed in section X.1, clients do transfer on average 2ECU more in IMPERSONAL than PERSONAL. If principals appreciate that, and expect that agents would appreciate it as well, they then may reciprocate by imposing restrictions less frequently. This interpretation however would imply a great degree of responsiveness of principal's behavior to clients' trust. We test this in a twofold way. First, we look, across treatments, at principals' decisions only for cases where the client sent the maximum possible amount. Then, we test the hypothesis that, within each treatment, principals' restrictive decision depend son clients' level of trust. We do not find evidence for the hypotheses in either test. Principals' level of restriction in response to the subset of clients sending the maximum (10) is not statistically different between treatments (p=0.38). Similarly, for both PERSONAL and IMPERSONAL treatments, a jonkeree test cannot reject the hypothesis that principals restrict in the same way regardless of the amount sent by the client (respectively p=0.82 and p=0.39). Moreover, the value of the randomly drawn restriction is significantly higher in IMPERSONAL than PERSONAL (respectively 10.12 and 7.27, p=0.000)[20], which should make easier for reciprocal principals to restrict in IMPERSONAL treatment, differently from what we find[21].

We suggest, alternatively, that principals may be reluctant to be "first movers" on unfair decisions, and prefer to signal their greed only when strictly necessary. If principals have preferences for appearing fair and trustworthy[22] (see i.e. Dana, Weber, Kuang 2007; Hao and Houser 2011), then in a PERSONAL restriction environment, they must signal their greed only when necessary, when the agent is obviously too generous. Delegation

---

[20]This is no surprise as the restriction is randomly drawn between half and two times the client transfer, which as well is higher under IMPERSONAL.

[21]An alternative explanation may be that under IMPERSONAL, principals do not restrict because they believe agents would decide as they would do if they had real authority. This however implies that principals from IMPERSONAL treatment should be systematically more generous than their counterparts from PERSONAL treatment, which finds no evidence in our data and no rationale in our experimental design

[22]Although not necessarily for being fair and trustworthy.

of authority by the principal lets the known greedy agent shoulder responsibility in other cases. In the impersonal rule environment, imposing a restricting impersonal rule signals their type to both generous and greedy agents (e.g. that they are untrustworthy). In this environment even agents that "didn't need to know", discover the type of the principal. This rationale provides one possible explanation for the reduced willingness to restrict in the IMPERSONAL treatment. This result appears to be in line with recent studies on delegation and blame aversion under principal-agent aligned monetary interests. Recent papers have in fact explored motives behind principals' delegation under aligned principal-agent monetary interests, and shown how delegation can be an effective instrument principals can use to shift the blame (and punishment) on their employees for unfair outcomes received by third parties (Bartling and Fischbacher 2012; Grossman 2012)[23]. Our paper nicely complements this literature by showing how delegation of unfair choices may be important for principals even when they do not face the threat of being punished by clients, and when their accountability in face of latter cannot be manipulated, as it is in the aforementioned works.

3.5   Conclusion

Our paper investigates the effects of control in principal-agents' relationships where monetary interests are aligned. We compare environments where formal authority is enforced either via direct monitoring or impersonal rules. Our results show that different hierarchical institutions produce different outcomes for firms, even when signaling outside the firm is not possible: given similar types of agents, personal restrictions generate significantly more costly punishment (hidden cost of control) from the agent side than impersonal rules. At the same time however, principals tend to exercise their authority less when restrictions are impersonal, meaning they refrain from restricting when unaware of their agent's type.

---

[23]These studies differ from ours in the followings: (i) principal's delegation towards agents, and fairness towards third parties is studied using a dictator game framework ; (ii) third parties are always aware of who shall be blamed.

Our paper points out an important trade off relevant for institution design within organizations. When choosing how to organize the internal hierarchy, our experimental results show that personal rules restricting the behavior of agents will result in generous agent deciding relatively more favorably for clients rather than principals, and the opposite behavioral switches for greedy agents. These results further indicate that the internal make up of distribution of agent types, e.g. generous versus greedy, is also expected to influence the overall impact of personal and impersonal rules. The nature of how rules are enacted has a large effect, not only upon the behavior of principals and agents within an environment, but also upon their beliefs. Our results demonstrated that the internal structure of rules bias agents' beliefs about the expectations of clients, in particular the beliefs of generous types, leading them to think that clients and principals are more generous than when rules come about in an impersonal manner.

Further our results demonstrate that the way the rule makers go about instituting rules also influences the expectations and behavior of the rule makers themselves as well. Firms should be cognoscent of the fact that those in charge of adopting and imposing new rules, typically managers and directors, will be systematically less willing to impose new rules when those rules are impersonal, as by doing so they would signal their low reciprocity towards clients more often than necessary. we estimate however that the net effect of impersonal rules is positive, suggesting thus that *impersonal rules* should be preferred to *personal restrictions*, even when both do not harm agents' monetary payoffs.

As monitoring subordinates directly typically involves a cost, our results suggest that principals who want their agents to keep as much surplus as possible within the firm, may doubly benefit from general and impersonal rules, rather than direct monitoring: (i) reduced monitoring costs and (ii) increased revenues. We show this is particularly true in environments where firms expect their employees will be particularly over-generous, such as in organizations where tips are a major source of agent income, as with bartenders or wait-staff. Future studies may find incorporating "tipping" into an environment such as ours valuable.

Appendix A

Appendix A : instructions[1]

<u>INSTRUCTIONS (Role A)</u>

Thank you for participating in today's experiment. You've earned a $5 show-up bonus for participating. In reading and following the instructions below, you have the potential to earn significantly more. In the experiment you will earn Experimental Dollars (E$s) which will be converted into cash (US Dollars) at the end of the experiment.
For every 3 E$s you have at the end of the experiment you will be paid 2 US Dollar in cash.

Your Role: A

The card you drew upon entering has randomly assigned you to a role: your role for this experiment is role A**.**
You will be randomly matched with two other participants: one participant randomly assigned to role B, and the other participant randomly assigned to role C.
Note that your counterparts will not be told your name, and you will not be told their names. The same applies for counterparts B and C.

The experiment consists in four independent periods. You will keep your role in every period, and you be randomly rematched with other participants in every period, as explained below.

There will be <u>four phases</u> in each period, and you will be asked to take <u>two decisions</u> in phase 1 and 3.

Below is a description of how a period proceeds.

Phase 1- Participant C's Decision and <u>your</u> Decision 1

You, participant B, and participant C have each been given an initial 10E$s.

As you read in participant C instructions, your counterpart in role C will decide among these three alternatives:
1) Send 2E$ out of their 10E$ to participant A (YOU)
2) Send 6E$ out of their 10E$ to participant A (YOU)
3) Send 10E$ out of their 10E$ to participant A (YOU)

You and participant B will observe participant C decision.

This amount is then multiplied by 6 and given to you.
Your <u>decision 1</u> consists in choosing how many of these E$ you want to send back to participant C.

Any amount you do not send back to participant C will be divided equally between you and participant B.

For example:

If participant C chooses to send 2E$, you will choose an amount between 0E$ and 12E$ (2E$ x 6) to send back to C. Participant B receives 50% of the amount not returned to C. You receive the remaining 50%. This amount will not be multiplied again.

If participant C chooses to send 6E$, you will choose an amount between 0E$ and 36E$ (6E$ x 6) to send back to C. Participant B receives 50% of the amount not returned to C. You receive the remaining 50%. This amount will not be multiplied again.

If participant C choose to send 10E$, you will choose an amount between 0E$ and 60E$ (10E$ x 6) to send back to C. Participant B receives 50% of the amount not returned to C. You receive the remaining 50%. This amount will not be multiplied again.

Please note that what you read so far represents <u>all</u> the information that is available to participant C.

The remaining sections of the instructions are private information between participants A and B and are <u>not known</u> by participant C. Participant C does not have any private information.

You and participant B observe exactly the same information unless otherwise noted below.

Phase 2 and 3 - Participant B decision and your decision 2

The decision 1 you made in phase 1 will <u>not be observed</u> by participant B until the end of the period, and could not be observed at all, as detailed below.

Once you make your decision 1, in phase 1, participant B <u>will observe</u> how much you chose to send back to participant C.

In phase 2, participant B will take a decision that can affect how you can take decision 2 in phase 3.

In phase 3 you will have to take decision 2. This Decision 2 is similar to decision 1. Whether your decision 1 or decision 2 will be used to calculate your group's payoffs will be decided randomly, as will be explained further below.

After you take decision 1, the computer will randomly draw **a number between** <u>half</u> of the amount participant C sent to you, <u>and twice</u> the amount participant C sent to you.

Both you and participant B will observe the realization of this value *before* participant B takes his decision.

<u>Participant B decision</u>: participant B has to decide if he wants to restrict your decision 2 to be less than the value of the random number. Again, when participant B takes his decision, he does <u>not</u> know what your decision 1 was.

If participant B decides to <u>restrict</u> your decision 2, the random value will represent the maximum amount <u>you</u> will be able to send back to participant C in your decision 2. This means that your decision 2 can be only between 0E$ and the value of the random draw that participant B decided to impose you.

If he decides to <u>not restrict</u> your decision 2, your decision 2 can be between 0E$ and 6 times the amount C sent, as for decision 1 you took in phase 1.

Once participant B takes his decision, you will observe his decision, and phase 3 starts. In phase 3 you will take decision 2: you will decide how many E$ you want to send back to participant C; as said before, your decision 2 may be affected by participant B decision and may be randomly selected for implementation.

For example, if participant C chose to send to you 6E$, you receive 36E$. After you take decision 1, the computer will randomly draw a number between 3E$ (half of 6E$, the amount participant C sent to you) and 12E$ (two times 6E$, the amount participant C sent to you).
If the random number is 4, then <u>if participant B chooses to restrict your decision 2</u> then your decision 2 <u>cannot exceed 4</u>. <u>If instead participant B chooses to not restrict</u>, then your decision 2 could exceed 4.

If the random number is 11, then <u>if participant B chooses to restrict your decision 2</u> then your decision 2 cannot exceed 11.
<u>If instead he chooses to not restrict you</u>, then your decision 2 could exceed 11.

Phase 4 - Random selection between participant A Decision 1 and Decision 2

Whether your decision 1 or decision 2 will be used to calculate your group's payoffs depends on a second random computer draw. More specifically:
- With a 40% chance, Decision 1 made in Phase 1 will be relevant for the period's payoffs of the group. If this is the case, participant B will observe your decision 1.
- With a 60% chance, Decision 2 made in Phase 3 will be relevant for period's payoffs of the group. If this is the case, participant B will observe your decision 2.

Participant B thus only observes either decision 1 or decision 2.

Please note that you and participant B will each receive the initial 10E$ independently from the decisions taken during the experiment.

Payment

This simple setup is repeated four times. Every participant in today's experiment (including you) will keep the same role in each period.

You will be placed randomly in a new group each period, and asked to choose take decisions within the new group.

Please note that your previous group of three people will never be in a group together again.

Please note again that participant C will not observe the outcome of a period until the end of the game, after the 4th period. You and your counterpart B will observe payoffs at the end of each period.

At the end of the 4th period, one of the four periods will be randomly chosen as the period relevant for final payments. You will be paid accordingly to decisions made by you and participants in your group in that specific period.

Please note that during each period, you will be asked to make guesses about the expectations of other participants. If your guesses for the randomly selected period are correct, you will earn additional E\$. The details will be explained during the experiment.

INSTRUCTIONS (Role B)

Thank you for participating in today's experiment. You've earned a $5 show-up bonus for participating. In reading and following the instructions below, you have the potential to earn significantly more. In the experiment you will earn Experimental Dollars (E$s) which will be converted into cash (US Dollars) at the end of the experiment.
For every 3 E$s you have at the end of the experiment you will be paid 2 US Dollar in cash.

Your Role: B

The card you drew upon entering has randomly assigned you to a role: your role for this experiment is role B**.**
You will be randomly matched with two other participants: one participant randomly assigned to role A, and the other participant randomly assigned to role C.
Note that your counterparts will not be told your name, and you will not be told their names. The same applies for counterparts A and C.

The experiment consists in four independent periods. You will keep your role in every period, and you be randomly rematched with other participants in every period, as explained below.

In each period there will be <u>four phases</u>, and you will be asked to make <u>one decision</u> in phase 2.

Below is a description of how a period proceeds.

Phase 1- Participant C's Decision and Participant A's Decision 1

You, participant A, and participant C have each been given an initial 10E$s.

As you read in participant C instructions, your counterpart in role C will decide among these three alternatives:
1) Send 2E$ out of their 10E$ to participant A
2) Send 6E$ out of their 10E$ to participant A
3) Send 10E$ out of their 10E$ to participant A

This amount is multiplied by 6 and given to Participant A. Participant A then chooses how many of these E$ he wants to send back to participant C.

Any amount participant A does not send back to participant C, will be divided equally between you and participant A.

For example:

If participant C chooses to send 2E$, your counterpart A will choose an amount between 0E$ and 12E$ (2E$ x 6) to send back to C. You receive 50% of the amount not returned to C. Participant A receives the remaining 50%. This amount will not be multiplied again.

If participant C chooses to send 6E$, your counterpart A will choose an amount between 0E$ and 36E$ (6E$ x 6) to send back to C. You receive 50% of the amount not returned to C. Participant A receives the remaining 50%. This amount will not be multiplied again.

If participant C choose to send 10E$, your counterpart A will choose an amount between 0E$ and 60E$ (10E$ x 6) to send back to C. You receive 50% of the amount not returned to C. Participant A receives the remaining 50%. This amount will not be multiplied again.

Please note that what you read so far represents all the information that is available to participant C.

The remaining sections of the instructions are private information between participants A and B and are <u>not known</u> by participant C. This means that participant C does not observe your role (participant B's role) in the experiment. Participant C does not have any private information.

You and participant A observe exactly the same information, unless otherwise noted below.

Phase 2 and Phase 3 - Your decision and Participant A's decision 2

Once participant A makes his/her decision 1, in phase 1, you <u>will not</u> observe how much they chose to send back to participant C at this point.

Once participant A makes his/her decision 1, in phase 1, you will observe how much they chose to send back to participant C.

Participant A will have to make another decision, "Decision 2", (very similar to decision 1) in phase 3. Whether participant A decision 1 or decision 2 will be used to calculate your group's payoffs will be decided randomly, as will be explained further below. You will only observe their decision (either 1 or 2) that is relevant for payoffs. Your decision in phase 2 may affect A's decision 2.

After participant A makes his decision 1 (After you observe participant A's decision 1), the computer will randomly draw a number between <u>half</u> of the amount participant C sent to participant A, and <u>twice</u> the amount participant C sent to participant A.

Both you and participant A will observe the realization of this value *before* you make your decision.

Your decision is to decide if you want restrict participant A's decision 2 to be less than the value of the random number.

If you decide to restrict A's decision, the randomly value will represent the maximum amount participant A will be able to send back to participant C in his decision 2. This means that participant A decision 2 can be between 0E$ and the value of the random draw.

If you decide to not restrict A's decision, A's decision 2 can be between 0E$ and 6 times the amount C sent.

For example, if participant C chose to send to participant A 6E$, participant A receives 36E$. The computer will randomly draw a number between 3E$ (half of 6E$, the amount participant C sent to A) and 12E$ (two times 6E$, the amount participant C sent to A). If the random number is 4, then <u>if you choose to restrict participant A's decision 2</u> then Participant A's decision 2 cannot exceed 4. <u>If instead you choose to not restrict</u>, then participant A's decision 2 could exceed 4. If the random number is 11, then <u>if you choose to restrict participant A's decision 2</u> then Participant A's decision 2 cannot exceed 11. <u>If instead you choose to not restrict</u>, then participant A's decision 2 could exceed 11.

Phase 4 - Random selection between participant A Decision 1 and Decision 2

Whether participant A decision 1 or decision 2 will be used to calculate your group's payoffs, depends on a second random computer draw. More specifically:
- With a 40% chance, Decision 1 made in Phase 1 will be relevant for the period's payoffs of the group. If this is the case you will observe their decision 1.
- With a 60% chance, Decision 2 made in Phase 3 will be relevant for period's payoffs of the group. If this is the case you will observe their decision 2.

Payment

This simple setup is repeated four times. Every participant in today's experiment (including you) will keep their same role in each period.

You will be placed randomly in a new group each period, and asked to choose how much to send to your role A counterpart.

Please note that your previous group of three people will never be in a group together again.

Please note again that participant C will not observe the outcome of a period until the end of the game, after the 4th period. You and your counterpart A will observe payoffs the end of each period.

At the end of the 4th period, one of the four periods will be randomly chosen as the period relevant for final payments. You will be paid accordingly to decisions made by you and participants in your group in that specific period.

Please note that during each period, you will be asked to make guesses about the expectations of other participants. If your guesses for the randomly selected period are correct,

you will earn additional E$. The details will be explained during the experiment.

INSTRUCTIONS (Role C)

Thank you for participating in today's experiment. You've earned a $5 show-up bonus for participating. In reading and following the instructions below, you have the potential to earn significantly more. In the experiment you will earn Experimental Dollars (E$s) which will be converted into cash (US Dollars) at the end of the experiment.
For every 3 E$s you have at the end of the experiment you will be paid 2 US Dollar in cash.

Your Role: C

The card you drew upon entering has randomly assigned you to a role: your role for this experiment is role C.
You will be randomly matched with two other participants: one participant randomly assigned to role A, and the other participant randomly assigned to role B.
Note that your counterparts will not be told your name, and you will not be told their names. The same applies for counterparts A and B.

The experiment consists in four independent periods. You will keep your role in every period, and you be randomly rematched with other participants in every period, as explained below.

Below is a description of how a period proceeds.

How to make your decision

You have been given an initial 10 E$s.

Participants B & A have also been given an initial 10E$s.

You have to decide how many of your E$ you want to send to your counterpart A. You have three alternative choices:
1) Send 2E$ out of your 10E$ to counterpart A;
2) Send 6E$ out of your 10E$ to counterpart A;
3) Send10E$ out of your 10E$ to counterpart A.

Whatever amount you send will be multiplied by six by the experimenter and given to your counterpart A. Participant A will observe your choice, and then will decide how many E$ to send back to you. Whatever A does not send back to you is divided between participants A and B at preset percentages, known to both A and B.

For example:

If you choose to send 2E$, your counterpart A will choose an amount between 0E$ and 12E$ (2E$ x 6) to send back to you. A and B split the remaining by the preset percentages. This amount will not be multiplied again.

If you choose to send 6E$, your counterpart A will choose an amount between 0E$ and 36E$ (6E$ x 6) to send back to you. A and B split the remaining by the preset percentages. This amount will not be multiplied again.

If you choose to send 10E$, your counterpart A will choose an amount between 0E$ and 60E$ (10E$ x 6) to send back to you. A and B split the remaining by the preset percentages. This amount will not be multiplied again.

Your payoff will be the sum of E$ participant A sends back to you plus (in case) the E$ you did not send.

Payment

This simple setup is repeated four times. Every participant in today's experiment (including you) will keep their same role in each period.

You will be placed randomly in a new group each period, and asked to choose how much to send to your counterpart A.

Please note that your previous group of three people will never be in a group together again.

You will see the outcome of each period only at the end of the game (after the 4th period.) At that point, one of the four periods will be randomly chosen as the period relevant for final payments.
You will be paid accordingly to decisions made by you and your counterparts in that specific period only.

Quiz (Participant C)

1) All participants have been each endowed with 10E$

   TRUE   FALSE

2) If you send 2E$, how many E$ does participant A receive?_____

3) If you send 6E$, how many E$ does participant A receive?_____

4) If you send 10E$, how many E$ does participant A receive?_____

5) If you send 10E$, how many E$ can participant A send back to you?

   Minimum: _____E$

   Maximum:_____E$

6) The experiment consists of 6 periods:

   TRUE   FALSE

7) You observe the outcome of each round only after the last period:

   TRUE   FALSE

8) Only one period will be randomly selected for payments: TRUE   FALSE

Quiz (As and Bs)

9) All participants have been each endowed with 10E$.

TRUE   FALSE

1) If participant C sends 6E$ to participant A, how many E$ does participant A receive? _____E$

2) If participant C sends 10E$ to participant A, how many E$ does participant A receive? _____E$

3) Participants A and B always receive their 10E$ at the end of the experiment, no matter what happens in the experiment

TRUE   FALSE

4) Participants C always receives his/her 10E$ at the end of the experiment, no matter what happens in the experiment

TRUE   FALSE

5) Participants A and B receive participant C's instructions.

TRUE   FALSE

6) Participant C receives the instructions of participants A and B. TRUE   FALSE

7) Participant B decision is to choose whether or not to impose the randomly drawn restriction on participant A:

TRUE   FALSE

8) Participant B chooses whether participant's A decision 1 or decision 2 will be implemented:

TRUE   FALSE

9) Suppose the random value of the restriction in a period is 4E$, and participant B decides to impose the restriction, can participant A send 7E$ to participant C?

YES   NO

10) In the latter scenario, what is the maximum amount of E$ participant C can earn if decision 2 is selected?_____ E$

11) What is the probability that decision 2 is selected?_____%

General Conclusion

This thesis contains two essays that explore the role of social image concerns in charitable giving decisions, and an essay that investigates how self-image, or pride, affects principal-agent relationships characterized by converging monetary interests, but misalignment of other non-monetary factors.

Understanding the role image concerns have on economic decision-making is of critical importance for both public policy , and non profit or firm practices. All individuals living in civil societies are often called to sacrifice resources to improve the well-being or wealth of others, squaring with the temptation of indulging over ones' own private benefit. Costly actions that benefit others may come essentially in two forms: people can spend resources to directly increase the well-being of others, or can renounce to something that benefits themselves but generates a negative externality. Public policies are thus often implemented with the objective of either encouraging private pro-social behaviors, or to dissuade behaviors that may generate negative externalities. Not infrequently, social exposure and peer pressure are used as an incentive mechanism. The first question raised thus is whether social exposure has the same efficacy in pushing people to do good, and to keep them from doing bad. As moral and monetary costs involved by pro-social "dilemmas" vary depending on the context, we further test whether the response to different combinations of these costs is similar or different for the two types of dilemmas. To answer these questions we look at both behavioral and neural responses to choices involving doing good or avoiding doing bad. While social image matters when it comes to choices involving *how generous* peo-

ple should be (and how costly being generous is), it may also be an important factor determining whether and how people take into account *how efficient* the beneficiary of their gifts is. The second question investigated in this dissertation is how donors respond to new information about the efficiency of their charities, and how the response depends on that information being visible or not to others. As in fact information can be ignored, it can't be hidden from others when it is of public domain. The essay thus concentrates on donors motivated by prestige, and explores whether good or bad news are treated as substitutes or complements with respect to the size of the gift. Finally, we concentrate our attention on within-firms hierarchical relationships, and explore how different mechanisms to implement hierarchies affect subordinates' willingness to comply with principals' authority, and principals' willingness to exercise that authority. As individuals give importance to the signaling value of coworkers' actions, we confront the effect of personal monitoring (or ad-hoc exercise of authority on subordinates) with impersonal rules, that is, restrictions imposed ex-ante to all subordinates regardless of their behavior. To focus on the signaling value of delegation or restrictions for agents' pride and principals' "shame" , the third essay investigates principal-agent relationships characterized by alignment of monetary interests and asks which type of hierarchical institution is more likely to push subordinates to comply with principal's authority, and to push the latter to enforce it.

The first essay uses fMRI imagery to compare two types of pro-social decisions where monetary and moral payoffs are at odds: incurring a monetary cost to allow a donation to a positively evaluated charity, and refusing to earn money to avoid a monetary transfer to an anti-charity. These decisions are taken under two conditions: full anonymity, and social exposure (individuals' decisions are observed by others). We systematically vary the cost of giving by proposing participants variable amounts for themselves (either cost or reward) and variable amounts for the association that would be transferred if the offer is accepted. Our results show that the response to public exposure is behaviorally different for two types of decisions. For decisions regarding the positive charity, the difference between private and public trials is significant only when

the dilemma is hard, that is, when losses for subjects are high but gains for the charity are even higher. Differently, for decisions concerning the negative association, we observe difference between private and public treatments only when the perspective gains for the negative association are high, regardless of the perspective gains for the subjects. This means that when gains for the anti-charity are below 16 Euro, subjects behave in public and private in the same way, regardless of how remunerative it is to accept. Differences between the two decisions are found also at a neural level, particularly when it comes to monetary gains deriving from selfish private decisions, which are dissociated in the Orbitofrontal Cortex.

The second essay studies how new information about real charities efficiency affects small donors contributions and whether the public visibility of that information matters for donors. Our study investigates the effect on charitable giving of the public visibility of charities efficiency and donors gifts. With treatments that progressively increase the visibility of donation amounts and charities efficiency, we assess the effect bad and good news on ones own charities have on donors behavior. First, we find that individuals tend to disregard bad news about their own charities when giving happens under full anonymity, but do increase their contributions to charities that turn out to be more efficient than expected. Overall thus, to the extent to which giving is made in private, our results suggest that the diffusion of homogeneous efficiency measures would increase the overall size of the market for giving. Second, we find that individuals start to react to bad news as soon as donations amounts are made public. Taken together, these results suggest that charities efficiency is treated as a proxy for quality or deservedness whenever this information remains private. Differently, we show that when also charities efficiency becomes public, a significant fraction of individuals starts to reduce their contributions to better-than-expected charities and increase their gifts to worse-than-expected ones.

The third essay investigates the effects of control in principal-agents' relationships where monetary interests are aligned. We compare environments where formal authority is enforced either via direct monitoring or impersonal rules. Our results show that differ-

ent hierarchical institutions produce different outcomes for firms, even when signaling outside the firm is not possible: given similar types of agents, personal restrictions generate significantly more costly punishment (hidden cost of control) from the agent side than impersonal rules. At the same time however, principals tend to exercise their authority less when restrictions are impersonal, meaning they refrain from restricting when unaware of their agent's type. This essay points out an important trade off relevant for institution design within organizations. When choosing how to organize the internal hierarchy, our experimental results show that personal rules restricting the behavior of agents will result in generous agent deciding relatively more favorably for clients rather than principals, and the opposite behavioral switches for greedy agents. These results further indicate that the internal make up of distribution of agent types, e.g. generous versus greedy, is also expected to influence the overall impact of personal and impersonal rules. The nature of how rules are enacted has a large effect, not only upon the behavior of principals and agents within an environment, but also upon their beliefs. Our results demonstrated that the internal structure of rules bias agents' beliefs about the expectations of clients, in particular the beliefs of generous types, leading them to think that clients and principals are more generous than when rules come about in an impersonal manner. Further we show that those in charge of adopting and imposing new rules, typically managers and directors, will be systematically less willing to impose new rules when those rules are impersonal, as by doing so they would signal their low reciprocity towards clients more often than necessary.

These three essays contain some limitations that constitute a baseline for future research.

In the first essay subjects making decisions over the positive charity always face a cost of giving strictly lower than one. This means that in our environment there is no possibility to make "inefficient" contributions. While our experiment allows us to investigate the neural correlates of different combinations of monetary losses and "moral" gains, it remains unclear whether the reward deriving from public pro-social behavior is driven by outcomes or actions itself. Otherwise said, we do not know

whether the sense of reward comes from the positive outcome generated or from the very fact of having sacrificed personal resources, no matter the consequences. A design that may help to shed light on that may involve a variable cost of giving greater than one, that is, part of the money donated may not reach the selected charity, or may reach it with variable probability. A neuro-experiment that involves "burning money" thus, would help to further our understanding of what drives the sense of reward associated with looking good. On the other hand, we found differences between the two decisions in terms of how money-maximizing actions are processed in the brain when social image is not a concern: while not making a donation to a positive charity correlates with reward-related regions, accepting money and allowing a transfer to a negative association recruits punishment and disgust-related regions. Further research would be needed to disentangle between two possible explanations: subjects may perceive that the social norms embedded in these two decisions are different (e.g. not being generous does not violate a social norm, while letting monarchic groups get funding does), or subjects may treat those decisions differently because of the different perceived relationship between actions and consequences (e.g. "by not being charitable I am not doing good, but I am not doing bad either" vs "by accepting money I am doing bad").

In the second essay, subjects are surprised in the second phase with information about the efficiency of their charities. In our design there is no possibility for subjects to not receive that information. This means that despite information can be ignored (as happens in one of our treatments), it can't be avoided. Indeed in real life situations acquiring information typically involves an active decision: do I want to know it or not? It would thus be interesting to extend our design by allowing subjects to actively choose whether to receive or not information about their charities. Such experiment would help to understand which types of people are more likely to withdraw the chance to get informed, and which types of charitable causes are more likely to induce "blessed ignorance". Such evidence would give us a glimpse in how ideologies and identity values affect people's willingness to search for information and update their priors.

141

Finally, in the third essay we allow principals to impose restrictions to the maximum amount their agent can send back to the client. This means that principals have a chance to limit agents' generosity towards clients, but have no means to control untrustworthy agents. A first natural extension thus would be to analyze the symmetric condition, that is, allowing principals to limit agents' greed. Second, in our design the effective value of the restriction is randomly drawn from a uniform distribution bounded between half and two times the amount sent by the client. As we did so to avoid principals' choosing focal values and to ease the interpretation of agents' second order beliefs on principals' actions, it would nevertheless be interesting to let principals freely choose the value of the restriction. This could be carried out by letting them choose simultaneously an upper and a lower restriction, allowing to further understand how agents' compliance with principal authority depends on the span of action left after restrictions are imposed.

Bibliography

Aghion, P., M. Dewatripont, and P. Rey. 2004. "Transferable control." *Journal of the European Economic Association* 2:115–138.

Aghion, P., and J. Tirole. 1997. "Formal and real authority in organizations." *Journal of political economy*, pp. 1–29.

Aimone, J., and D. Houser. 2012a. "Harnessing the Benefits of Betrayal Aversion." Working paper, George Mason University, Interdisciplinary Center for Economic Science.

—. 2012b. "What you don't know won't hurt you: a laboratory analysis of betrayal aversion." *Experimental Economics* 15:571–588.

Aimone, J.A., and D. Houser. 2011. "Beneficial betrayal aversion." *PloS one* 6:e17725.

Akerlof, G.A. 1983. "Loyalty filters." *The American economic review*, pp. 54–63.

Andreoni, J. 1989. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *The Journal of Political Economy*, pp. 1447–1458.

—. 1990. "Impure altruism and donations to public goods: a theory of warm-glow giving." *The Economic Journal* 100:464–477.

—. 1988. "Why free ride?: Strategies and learning in public goods experiments." *Journal of Public Economics* 37:291–304.

Andreoni, J., and J. Miller. 2003. "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica* 70:737–753.

143

Andreoni, J., and R. Petrie. 2004. "Public goods experiments without confidentiality: a glimpse into fund-raising." *Journal of Public Economics* 88:1605–1623.

Ariely, D., A. Bracha, and S. Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99:544–555.

Armstrong, M., and J. Vickers. 2010. "A model of delegated project choice." *Econometrica* 78:213–244.

Arrow, K.J., and F.H. Hahn. 1971. *General competitive analysis*. Holden-Day San Francisco.

Babcock, L., and G. Loewenstein. 1997. "Explaining bargaining impasse: The role of self-serving biases." *The Journal of Economic Perspectives* 11:109–126.

Bartling, B., and U. Fischbacher. 2012. "Shifting the blame: On delegation and responsibility." *The Review of Economic Studies* 79:67–87.

Bateson, M., D. Nettle, and G. Roberts. 2006. "Cues of being watched enhance cooperation in a real-world setting." *Biol Lett* 2:412â€"414.

Beck, A.T., and R.W. Beck. 1972. "Screening depressed patients in family practice. A rapid technic." *Postgraduate Medicine* 52:81–85, PMID: 4635613.

Becker, G.S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82:1063–93.

Benabou, R., and J. Tirole. 2006a. "Incentives and Prosocial Behavior." *The American Economic Review* 96:1652–1678, ArticleType: primary_article / Full publication date: Dec., 2006 / Copyright Â© 2006 American Economic Association.

—. 2006b. "Incentives and prosocial behavior." *The American economic review* 96:1652–1678.

Berg, J., J. Dickhaut, and K. McCabe. 1995. "Trust, reciprocity, and social history." *Games and economic behavior* 10:122–142.

144

Bernheim, B.D. 1994. "A theory of conformity." *Journal of political Economy*, pp. 841–877.

Berns, G.S., E. Bell, C.M. Capra, M.J. Prietula, S. Moore, B. Anderson, J. Ginges, and S. Atran. 2012. "The price of your soul: neural evidence for the non-utilitarian representation of sacred values." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367:754–762.

Blakemore, S.J., and U. Frith. 2004. "How does the brain deal with the social world?" *Neuroreport* 15:119–128, PMID: 15106843.

Botvinick, M.M., T.S. Braver, D.M. Barch, C.S. Carter, and J.D. Cohen. 2001. "Conflict monitoring and cognitive control." *Psychol Rev* 108:624â€"652.

Bracha, A., O. Heffetz, and L. Vesterlund. 2009. "Charitable giving: The effects of exogenous and endogenous status." Working paper, University of Pittsburgh Working Paper.

Bracha, A., M. Menietti, and L. Vesterlund. 2011. "Seeds to succeed?: Sequential giving to public projects." *Journal of Public Economics* 95:416–427.

Bradley, G.W. 1978. "Self-serving biases in the attribution process: A reexamination of the fact or fiction question." *Journal of Personality and Social Psychology* 36:56.

Cabeza, R., and L. Nyberg. 2000. "Imaging cognition II: An empirical review of 275 PET and fMRI studies." *Journal of Cognitive Neuroscience* 12:1–47, PMID: 10769304.

Cabeza, R., S.E. Prince, S.M. Daselaar, D.L. Greenberg, M. Budde, F. Dolcos, K.S. LaBar, and D.C. Rubin. 2004. "Brain activity during episodic retrieval of autobiographical and laboratory events: an fMRI study using a novel photo paradigm." *Journal of Cognitive Neuroscience* 16:1583–1594, PMID: 15622612.

Call, J., and M. Tomasello. 1998. "Distinguishing intentional from accidental actions in orangutans (Pongo pygmaeus), chimpanzees (Pan troglodytes) and human children (Homo sapiens)." *Journal of Comparative Psychology* 112:192.

Camerer, C., G. Loewenstein, and D. Prelec. 2005. "Neuroeconomics: How neuroscience can inform economics." *Journal of economic Literature* 43:9–64.

Cauda, F., F. D'Agata, K. Sacco, S. Duca, G. Geminiani, and A. Vercelli. 2011. "Functional connectivity of the insula in the resting brain." *NeuroImage* 55:8–23, PMID: 21111053.

Charness, G., and D.I. Levine. 2007. "Intention and stochastic outcomes: An experimental study." *The Economic Journal* 117:1051–1072.

Chib, V.S., A. Rangel, S. Shimojo, and J.P. O'Doherty. 2009. "Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex." *J. Neurosci.* 29:12315–12320.

Chumbley, J.R., and K.J. Friston. 2009. "False discovery rate revisited: FDR and topological inference using Gaussian random fields." *NeuroImage* 44:62–70.

Craig, A.D.B. 2009. "How do you feel–now? The anterior insula and human awareness." *Nature Reviews. Neuroscience* 10:59–70, PMID: 19096369.

Critchley, H.D., P. Rotshtein, Y. Nagai, J. O'Doherty, C.J. Mathias, and R.J. Dolan. 2005. "Activity in the human brain predicting differential heart rate responses to emotional facial expressions." *NeuroImage* 24:751–762, PMID: 15652310.

Curtis, C.E., and M. D'Esposito. 2004. "The effects of prefrontal lesions on working memory performance and theory." *Cognitive, Affective, & Behavioral Neuroscience* 4:528–539, PMID: 15849895.

Dana, J., R.A. Weber, and J.X. Kuang. 2007. "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness." *Economic Theory* 33:67–80.

D'Argembeau, A., D. Feyers, S. Majerus, F. Collette, M. Van der Linden, P. Maquet, and E. Salmon. 2008. "Self-reflection across time: cortical midline structures differentiate between present and past selves." *Social Cognitive and Affective Neuroscience* 3:244–252, PMID: 19015116.

D'Argembeau, A., D. Stawarczyk, S. Majerus, F. Collette, M. Van der Linden, and E. Salmon. 2010. "Modulation of medial prefrontal and inferior parietal cortices when thinking about past, present, and future selves." *Social Neuroscience* 5:187.

Decety, J., and J.A. Sommerville. 2003. "Shared representations between self and other: a social cognitive neuroscience view." *Trends in Cognitive Sciences* 7:527–533, PMID: 14643368.

Decety, J., J.A. Sommerville, et al. 2003. "Shared representations between self and other: a social cognitive neuroscience view." *Trends in cognitive sciences* 7:527–533.

Delgado, M.R., M.M. Gillis, and E.A. Phelps. 2008. "Regulating the expectation of reward via cognitive strategies." *Nat Neurosci* 11:880â€"881.

Dickinson, D., and M.C. Villeval. 2008. "Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories." *Games and Economic Behavior* 63:56–76.

Eil, D., and J.M. Rao. 2011. "The good news-bad news effect: Asymmetric processing of objective information about yourself." *American Economic Journal: Microeconomics* 3:114–138.

Ellingsen, T., and M. Johannesson. 2008. "Pride and prejudice: The human side of incentive theory." *The American Economic Review* 98:990–1008.

Falk, A., and E. Fehr. 2003. "Why labour market experiments?" *Labour Economics* 10:399–406.

Falk, A., and M. Kosfeld. 2006. "The hidden costs of control." *The American economic review*, pp. 1611–1630.

Fehr, E., and C.F. Camerer. 2007. "Social neuroeconomics: the neural circuitry of social preferences." *Trends Cogn Sci* 11:419â€"427.

Fehr, E., and L. Goette. 2007. "Do workers work more if wages are high? Evidence from a randomized field experiment." *The American Economic Review*, pp. 298–317.

Fehr, E., H. Herz, and T. Wilkening. forthcoming. "The lure of authority: Motivation and incentive effects of power." *American Economic Review*, pp. .

Fehr, E., G. Kirchsteiger, and A. Riedl. 1993. "Does fairness prevent market clearing? An experimental investigation." *The Quarterly Journal of Economics* 108:437–459.

Fehr, E., and J.A. List. 2004. "The hidden costs and returns of incentivesâ€"trust and trustworthiness among CEOs." *Journal of the European Economic Association* 2:743–771.

Fehr, E., B. Rockenbach, et al. 2003. "Detrimental effects of sanctions on human altruism." *Nature* 422:137–140.

Fischbacher, U. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10:171–178.

Fischbacher, U., S. Gächter, and E. Fehr. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters* 71:397–404.

Fong, C.M. 2007. "Evidence from an Experiment on Charity to Welfare Recipients: Reciprocity, Altruism and the Empathic Responsiveness Hypothesis*." *The Economic Journal* 117:1008–1024.

Frey, B.S. 1993. "Motivation as a limit to pricing." *Journal of Economic Psychology* 14:635–664.

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

Frey, B.S., and F. Oberholzer-Gee. 1997. "The cost of price incentives: An empirical analysis of motivation crowding-out." *The American economic review*, pp. 746–755.

Frith, C.D., and T. Singer. 2008. "The role of social cognition in decision making." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363:3875–3886, PMID: 18829429.

Frith, U., and C.D. Frith. 2003. "Development and neurophysiology of mentalizing." *Philosophical Transactions of the Royal Society B: Biological Sciences* 358:459–473, PMID: 12689373 PMCID: 1693139.

Gneezy, U., and A. Rustichini. 2000. "Pay enough or don't pay at all." *The Quarterly Journal of Economics* 115:791–810.

Goette, L., D. Huffman, S. Meier, and M. Sutter. 2012. "Competition Between Organizational Groups: Its Impact on Altruistic and Antisocial Motivations." *Management Science* 58:948–960.

Greene, J.D. 2007. *The secret joke of Kant's soul*. MIT Press Cambridge, MA.

Greene, J.D., L.E. Nystrom, A.D. Engell, J.M. Darley, and J.D. Cohen. 2004. "The neural bases of cognitive conflict and control in moral judgment." *Neuron* 44:389–400, PMID: 15473975.

Greene, J.D., R.B. Sommerville, L.E. Nystrom, J.M. Darley, and J.D. Cohen. 2001. "An fMRI investigation of emotional engagement in moral judgment." *Science (New York, N.Y.)* 293:2105–2108, PMID: 11557895.

Grossman, Z. 2010. "Self-signaling versus social-signaling in giving." *working paper*, pp. .

Gusnard, D.A., and M.E. Raichle. 2001. "Searching for a baseline: functional imaging and the resting human brain." *Nature Reviews. Neuroscience* 2:685–694, PMID: 11584306.

Haley, K.J., and D.M. Fessler. 2005. "Nobody's watching. Subtle cues affect generosity in an anonymous economic game." *Evol Hum Behav* 26:245 – 256.

Hammers, A., R. Allom, M.J. Koepp, S.L. Free, R. Myers, L. Lemieux, T.N. Mitchell, D.J. Brooks, and J.S. Duncan. 2003. "Three-dimensional maximum probability atlas of the

human brain, with particular reference to the temporal lobe." *Human Brain Mapping* 19:224–247, PMID: 12874777.

Hampton, A.N., P. Bossaerts, and J.P. O'Doherty. 2008. "Neural correlates of mentalizing-related computations during strategic interactions in humans." *Proceedings of the National Academy of Sciences of the United States of America* 105:6741–6746, PMID: 18427116.

Hao, L., and D. Houser. 2010. "Honest lies.", pp. .

Harbaugh, W.T. 1998a. "The prestige motive for making charitable transfers." *The American Economic Review* 88:277–282.

—. 1998b. "What do donations buy?: A model of philanthropy based on prestige and warm glow." *Journal of Public Economics* 67:269–284.

Harbaugh, W.T., U. Mayr, and D.R. Burghart. 2007a. "Neural responses to taxation and voluntary giving reveal motives for charitable donations." *Science* 316:1622â€"1625.

—. 2007b. "Neural responses to taxation and voluntary giving reveal motives for charitable donations." *Science* 316:1622–1625.

Hare, T.A., C.F. Camerer, D.T. Knoepfle, and A. Rangel. 2010. "Value Computations in Ventral Medial Prefrontal Cortex during Charitable Decision Making Incorporate Input from Regions Involved in Social Cognition." *J. Neurosci.* 30:583–590.

Hare, T.A., C.F. Camerer, and A. Rangel. 2009. "Self-control in decision-making involves modulation of the vmPFC valuation system." *Science* 324:646â€"648.

Hare, T.A., J. O'Doherty, C.F. Camerer, W. Schultz, and A. Rangel. 2008. "Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors." *J Neurosci* 28:5623â€"5630.

Hart, O., and B. Holmstrom. 2010. "A theory of firm scope." *The Quarterly Journal of Economics* 125:483–513.

Hedden, T., and J. Zhang. 2002. "What do you think I think you think?: Strategic reasoning in matrix games." *Cognition* 85:1–36.

Heekeren, H.R., I. Wartenburger, H. Schmidt, K. Prehn, H.P. Schwintowski, and A. Villringer. 2005. "Influence of bodily harm on neural correlates of semantic and moral decision-making." *NeuroImage* 24:887–897, PMID: 15652323.

Heekeren, H.R., I. Wartenburger, H. Schmidt, H.P. Schwintowski, and A. Villringer. 2003. "An fMRI study of simple ethical decision-making." *Neuroreport* 14:1215–1219, PMID: 12824762.

Holländer, H. 1990. "A social exchange approach to voluntary cooperation." *The American Economic Review*, pp. 1157–1167.

Houser, D., S. Vetter, and J. Winter. 2012. "Fairness and cheating." *European Economic Review*, pp. .

Izuma, K., D.N. Saito, and N. Sadato. 2010. "Processing of the incentive for social approval in the ventral striatum during charitable donation." *Journal of Cognitive Neuroscience* 22:621–631, PMID: 19320552.

Kable, J.W., and P.W. Glimcher. 2007. "The neural correlates of subjective value during intertemporal choice." *Nat Neurosci* 10:1625â€"1633.

Karlan, D., and J.A. List. 2007. "Does price matter in charitable giving? Evidence from a large-scale natural field experiment." *The American Economic Review*, pp. 1774–1793.

Latane, B. 1970. "Field studies of altruistic compliance." *Representative Research in Social Psychology*, pp. .

Lee, D., M.F.S. Rushworth, M.E. Walton, M. Watanabe, and M. Sakagami. 2007. "Functional specialization of the primate frontal cortex during decision making." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27:8170–8173, PMID: 17670961.

151

Lévy-Garboua, L., C. Meidinger, and B. Rapoport. 2006. "The formation of social preferences: Some lessons from psychology and biology." *Handbook of the economics of giving, altruism and reciprocity* 1:545–613.

List, J.A., and D. Lucking-Reiley. 2002. "The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign." *Journal of Political Economy* 110:215–233.

Luo, Q., M. Nakic, T. Wheatley, R. Richell, A. Martin, and R.J.R. Blair. 2006. "The neural basis of implicit moral attitude–An IAT study using event-related fMRI." *NeuroImage* 30:1449–1457.

M Fong, C., and F. Oberholzer-Gee. 2011. "Truth in giving: Experimental evidence on the welfare effects of informed giving to the poor." *Journal of Public Economics* 95:436–444.

Mazaika, P., F. Hoeft, G. Glover, and A.L. Reiss. 2009. "Methods and Software for fMRI Analysis for Clinical Subjects."

Medford, N., and H.D. Critchley. 2010. "Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response." *Brain Structure & Function* 214:535–549, PMID: 20512367.

Meltzoff, A.N., and W. Prinz. 2002. *The imitative mind: Development, evolution and brain bases*, vol. 6. Cambridge University Press.

Milham, M.P., M.T. Banich, and V. Barad. 2003. "Competition for priority in processing increases prefrontal cortex's involvement in top-down control: an event-related fMRI study of the stroop task." *Brain Research. Cognitive Brain Research* 17:212–222, PMID: 12880892.

Miller, E.K. 2000. "The prefrontal cortex and cognitive control." *Nature Reviews. Neuroscience* 1:59–65, PMID: 11252769.

Miller, E.K., and J.D. Cohen. 2001. "An integrative theory of prefrontal cortex function." *Annu Rev Neurosci* 24:167–202.

Miller, G.J. 1993. *Managerial dilemmas: The political economy of hierarchy*. Cambridge University Press.

Mobius, M.M., M. Niederle, P. Niehaus, and T.S. Rosenblat. 2011. "Managing self-confidence: Theory and experimental evidence." Working paper, National Bureau of Economic Research.

Moll, J., R. de Oliveira-Souza, I.E. Bramati, and J. Grafman. 2002a. "Functional networks in emotional moral and nonmoral social judgments." *NeuroImage* 16:696–703, PMID: 12169253.

Moll, J., R. de Oliveira-Souza, P.J. Eslinger, I.E. Bramati, J. Mourão-Miranda, P.A. Andreiuolo, and L. Pessoa. 2002b. "The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 22:2730–2736, PMID: 11923438.

Moll, J., F. Krueger, R. Zahn, M. Pardini, R.d. Oliveira-Souza, and J. Grafman. 2006. "Human fronto-mesolimbic networks guide decisions about charitable donation." *Proc Natl Acad Sci U S A* 103:15623–15628.

Nicolle, A., M.C. Klein-Flügge, L.T. Hunt, I. Vlaev, R.J. Dolan, and T.E. Behrens. 2012. "An agent independent axis for executed and modeled choice in medial prefrontal cortex." *Neuron* 75:1114–1121.

North, D. 1981. *Growth and structural change*.

Northoff, G., and F. Bermpohl. 2004. "Cortical midline structures and the self." *Trends in Cognitive Sciences* 8:102–107, PMID: 15301749.

Ochsner, K.N., and J.J. Gross. 2005. "The cognitive control of emotion." *Trends in Cognitive Sciences* 9:242–249, PMID: 15866151.

Oexl, R., and Z.J. Grossman. 2012. "Shifting the blame to a powerless intermediary." *Experimental Economics*, pp. 1–7.

153

Oldfield, R.C. 1971. "The assessment and analysis of handedness: the Edinburgh inventory." *Neuropsychologia* 9:97–113, PMID: 5146491.

Padoa-Schioppa, C., and J.A. Assad. 2006. "Neurons in the orbitofrontal cortex encode economic value." *Nature* 441:223â€"226.

Petrides, M. 2000. "The role of the mid-dorsolateral prefrontal cortex in working memory." *Experimental Brain Research. Experimentelle Hirnforschung.* 133:44–54, PMID: 10933209.

Plassmann, H., J. O'Doherty, and A. Rangel. 2007. "Orbitofrontal cortex encodes willingness to pay in everyday economic transactions." *J Neurosci* 27:9984â€"9988.

Pollatos, O., K. Gramann, and R. Schandry. 2007. "Neural systems connecting interoceptive awareness and feelings." *Human Brain Mapping* 28:9–18, PMID: 16729289.

Potters, J., M. Sefton, and L. Vesterlund. 2005. "After you: endogenous sequencing in voluntary contribution games." *Journal of Public Economics* 89:1399–1419.

—. 2007. "Leading-by-example and signaling in voluntary contribution games: an experimental study." *Economic Theory* 33:169–182.

Rege, M., and K. Telle. 2004a. "The impact of social approval and framing on cooperation in public good situations." *Journal of Public Economics* 88:1625–1644.

—. 2004b. "The impact of social approval and framing on cooperation in public good situations." *Journal of Public Economics* 88:1625–1644.

Ribar, D.C., and M.O. Wilhelm. 2002. "Altruistic and Joy-of-Giving Motivations in Charitable Behavior." *Journal of Political Economy* 110:425–457.

Saxe, R., and N. Kanwisher. 2003. "People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind"." *NeuroImage* 19:1835–1842, PMID: 12948738.

Saxe, R., and A. Wexler. 2005. "Making sense of another mind: The role of the right temporo-parietal junction." *Neuropsychologia* 43:1391–1399.

Schnedler, W., and R. Vadovic. 2011. "Legitimacy of control." *Journal of Economics & Management Strategy* 20:985–1009.

Seo, H., D.J. Barraclough, and D. Lee. 2007. "Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex." *Cerebral Cortex* 17:110–117, PMID: 17548802.

Shang, J., and R. Croson. 2009. "A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods." *The Economic Journal* 119:1422–1439.

Sharot, T., R. Kanai, D. Marston, C.W. Korn, G. Rees, and R.J. Dolan. 2012. "Selectively altering belief formation in the human brain." *Proceedings of the National Academy of Sciences* 109:17058–17062.

Sharot, T., C.W. Korn, and R.J. Dolan. 2011. "How unrealistic optimism is maintained in the face of reality." *Nature neuroscience* 14:1475–1479.

Sirigu, A., E. Daprati, P. Pradat-Diehl, N. Franck, and M. Jeannerod. 1999. "Perception of self-generated movement following left parietal lesion." *Brain: A Journal of Neurology* 122 (Pt 10):1867–1874, PMID: 10506089.

Small, D.A., and G. Loewenstein. 2003. "Helping a victim or helping the victim: Altruism and identifiability." *Journal of Risk and Uncertainty* 26:5–16.

Smith, A. 1759. *The theory of moral sentiments*.

Straube, T., and W.H.R. Miltner. 2011. "Attention to aversive emotion and specific activation of the right insula and right somatosensory cortex." *NeuroImage* 54:2534–2538, PMID: 20946962.

Stutzer, A., L. Goette, and M. Zehnder. 2011. "Active Decisions and Prosocial Behaviour: a Field Experiment on Blood Donation*." *The Economic Journal* 121:F476–F493.

Sugden, R. 1984. "Reciprocity: the supply of public goods through voluntary contributions." *The Economic Journal* 94:772–787.

Svenson, O. 1981. "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica* 47:143–148.

Tamir, D.I., and J.P. Mitchell. 2010. "Neural correlates of anchoring-and-adjustment during mentalizing." *Proceedings of the National Academy of Sciences of the United States of America* 107:10827–10832, PMID: 20534459.

Tom, S.M., C.R. Fox, C. Trepel, and R.A. Poldrack. 2007. "The neural basis of loss aversion in decision-making under risk." *Science* 315:515â€"518.

Vesterlund, L. 2003. "The informational value of sequential fundraising." *Journal of Public Economics* 87:627–657.

Wallis, J.D. 2007. "Orbitofrontal cortex and its contribution to decision-making." *Annu Rev Neurosci* 30:31â€"56.

Wallis, J.D., and E.K. Miller. 2003. "Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task." *Eur J Neurosci* 18:2069â€"2081.

Watanabe, M., and M. Sakagami. 2007. "Integration of cognitive and motivational context information in the primate prefrontal cortex." *Cerebral Cortex* 17:101–109, PMID: 17725993.

Zajonc, R.B. 1965. "Social facilitation." *Science (New York, N.Y.)* 149:269–274, PMID: 14300526.

THREE ESSAYS ON THE ECONOMICS OF IMAGE MOTIVATION

Luigi Butera

Abstract

This thesis contains two essays that assess the role of social image concerns in charitable giving decisions and one essay that explores the role of authority in hierarchical relationships characterized by aligned monetary interests.

The first essay uses a neuro-economic experiment to compare the effect of social exposure on two types of pro-social behaviors: doing good and avoiding doing bad. We find that image gains deriving from visible acts of generosity are computed by the brain as rewards for both decisions. Differently, selfish decisions with no negative image consequences are computed differently: not doing good (and thus saving money) correlates with reward-related regions, while doing bad (and thus earning money) correlates with regions related to punishment's anticipation and moral disgust.

The second essay studies how information about real charities' efficiency (and its social visibility) affects small donors contributions. We find that individuals disregard bad news about their own charities when giving happens under full anonymity, but do increase their contributions to charities that perform better than expected. Differently, when both the amount donated and the efficiency of the recipient are public knowledge, donors motivated by social image concerns treat the quantity and the quality of their donations as substitutes.

The third investigates the effects of control in principal-agents' relationships where monetary interests are aligned. By comparing direct control and general impersonal rules, we show that direct monitoring generates significantly more hidden costs of control from the agent side than impersonal rules. At the same time however, principals tend to exercise their authority less when restrictions are impersonal, as these force them to signal their greed also when unnecessary.

Keywords: Charitable giving; Social image; Information; Authority; Institutions

Résumé

Cette thèse contient deux essais qui évaluent le rôle de l'image sociale dans les décisions de dons, et un essai qui explore le rôle de l'autorité dans les relations hiérarchiques caractérisées par des intérêts monétaires alignés.

Le premier essai utilise une expérience neuro-économique pour comparer l'effet de l'exposition sociale sur deux types de comportements pro-sociaux : faire le bien et éviter de faire le mal. Nous constatons que les gains d'image découlant des actes visibles de générosité sont calculés comme des récompenses dans les deux cas. Les décisions égoïstes sans conséquences d'images négatives sont calculées différemment : ne pas faire le bien (et économiser ainsi de l'argent) est corrélé avec les régions du cerveau liées à la récompense, alors que faire le mal (et donc gagner de l'argent) est corrélé avec les régions du cerveau liées à l'anticipation de la punition et au dégoût moral.

Le deuxième essai étudie comment l'information sur la véritable efficacité des organismes de bienfaisance (et sa visibilité sociale) affecte les contributions des petits donateurs. Nous constatons que les individus ignorent les mauvaises performances des organismes de bienfaisance quand le don est couvert par l ?anonymat mais augmentent leurs contributions aux organismes de bienfaisance qui sont plus performants que prévu. Toutefois, lorsque le montant donné et l'efficacité du destinataire sont de notoriété publique, les donateurs motivés par des préoccupations d'image sociale traitent la quantité et la qualité de leurs dons comme des substituts.

Le troisième essai étudie les effets du contrôle dans les relations de principal-agent où les intérêts monétaires sont alignés. En comparant le contrôle direct avec des règles impersonnelles générales, nous montrons que le contrôle direct génère plus de coûts cachés du côté de l'agent que les règles impersonnelles. En même temps, les directeurs ont tendance à moins exercer leur autorité lorsque les règles sont impersonnelles, car cela les obligent à signaler leur cupidité même lorsque ce n ?est pas nécessaire.

Mots Clés : Generosité ; Image sociale ; Information ; Autorité ; Institutions