JOB A. HARMS

# Essays on the Behavioral Economics of Social Preferences and Bounded Rationality

# Essays on the Behavioral Economics of Social Preferences and Bounded Rationality

# Essays on the Behavioral Economics of Social Preferences and Bounded Rationality

## Essays over de Gedragseconomie van Sociale Preferenties en Beperkte Rationaliteit

## Thesis

to obtain the degree of Doctor from the

Erasmus University Rotterdam

by command of the

rector magnificus

Prof. dr. H.A.P. Pols

And in accordance with the decision of the Doctorate Board

The public defense shall be held on

14 June 2018 at 09:30 hrs

by

Job August Harms

born in Amsterdam, the Netherlands

## Doctoral committee

Promotor: Prof.dr. H.R. Commandeur

Other members: Prof.dr. S. Rosenkranz

Prof.dr. H.D. Webbink

Dr. J.T.R. Stoop

Co-promotor: Dr. K.E.H. Maas

**Erasmus Research Institute of Management – ERIM**
The joint research institute of the Rotterdam School of Management (RSM)
and the Erasmus School of Economics (ESE) at the Erasmus University Rotterdam
Internet: http://www.erim.eur.nl

**ERIM Electronic Series Portal:** http://repub.eur.nl/

**ERIM PhD Series in Research in Management**, 457
ERIM reference number: EPS-2018-457-S&E
ISBN 978-90-5892-519-0


Design: PanArt, www.panart.nl
Coverphoto:Harm van de Poel
This publication (cover and interior) is printed by Tuijtel on recycled paper, BalanceSilk®
The ink used is produced from renewable resources and alcohol free fountain solution.
Certifications for the paper and the printing production process: Recycle, EU Ecolabel, FSC®, ISO14001.
More info: www.tuijtel.com




TO THE MEMORY OF MY FATHER

# Acknowledgements

I am grateful to everyone that played a role in the creation of this dissertation. Your support provided me with much anecdotal evidence for the existence of altruism, and your advice helped me in dealing with my own bounded rationality.

In particular, I am greatly thankful to Harry Commandeur, who entrusted me with the opportunity to conduct this research and who provided invaluable guidance in my explorations. You have helped me in finding a balance between pursuing my broad interests and converging towards a coherent dissertation. Most of all, I greatly appreciate your encouragement to reflect on the broader implications of my research. Being sceptical about economics as a "dismal science", you have inspired me to pursue economic research that contributes to a better world, based on a realistic view of human nature.

Furthermore, the writing of this dissertation would not have been possible without my co-promotor Karen Maas. Your energy and positive attitude made has made the work a lot of fun, and you too have inspired me do research for the greater good. Moreover, you have greatly helped me in setting up research projects with many different organizations.. Most of all, I am thankful for your warmth and empathy, which has only strengthened my view of a prosocial human nature.

Next, I express great appreciation to Dinand Webbink. After I left my previous job, you took me on as a research assistant, which formed the springboard for my later PhD position. Furthermore, through our joint work I learned many valuable things about statistics and policy evaluation. While causal effects are often hard to isolate, I am confident you had a significant effect on making this work possible.

To all my co-authors, it has been a great pleasure working with you. Ernst Fehr, Michel Maréchal, Alain Cohn and Helen Toxopeus: I have learned many things from you about rigorous experimentation and hope our project will form the basis of interesting future studies. Furthermore, I am particularly thankful to Ernst Fehr and Michel Maréchal for hosting in Zürich for a research visit in the spring of 2017, it was a lovely experience. Next, I am thankful to John Protzko, Kellie Liket and Vera Schölmerich for the work on our free will disbelief paper; to Haki Pamuk for your contribution to the ambiguity aversion experiment and finally to Mark Sanders and Stephanie Rosenkranz for inviting me to your research project on complex choices; the decision to join was easily made.

I am thankful to many colleagues in Rotterdam and elsewhere for stimulating and fruitful discussions, including but not limited to the following people: Frank Hubers, Martijn Hendriks, Dana Sisak, Jan Stoop, Robert Dur, Carly Relou, Marjelle Vermeulen, Renate Kersten, Fedes van Rijn, Giel Ton and David Reinstein. Furthermore, many thanks to my former ECSP colleagues, including Lucas Meijs, Michiel de Wilde, Lonneke van der Waa and Charles Erkelens. May the altruistic force be with you! And for those I forget to mention here, please blame it on my bounded rationality.

Finally, I thank my loved ones. In particular, I thank my parents for fostering my curiosity, my sisters for being there for me, and my dear friends Matthias Havenaar, Laure Heilbron, Rik van Lieshout and Rutger Beekman for critical feedback and support. And to all the others, you are also in my thoughts. Most of all, I thank my sweet Anne-Marie. You cheer me up. I'm a fool for you.

# Contents

**Chapter 1: Introduction**

Summary ..... 1
The puzzle of human social behavior ..... 2
Social nature in philosophy and economics ..... 3
The evolution of human sociality ..... 5
Social preferences and bounded rationality ..... 7
The behavioral synthesis ..... 8
Overview ..... 9
Outline of chapters ..... 13

**Chapter 2: Better Bankers**

Abstract ..... 19
2.1 Introduction ..... 20
Rationale of study ..... 20
Overview of experiment ..... 23
Main contributions ..... 255
Outline ..... 266
2.1 Experimental design ..... 266
Ethics program ..... 266
Randomization ..... 288
Audit study ..... 299
2.3 Results ..... 33
Descriptive statistics ..... 33
Baseline ..... 34
Compliance ..... 36
Average treatment effect ..... 39
Heterogeneous treatment effects ..... 41
Survey ..... 44
2.4 Discussion ..... 47
Summary of main results ..... 47
Interpretation ..... 48
Limitations ..... 49
Policy relevance and future research ..... 50
Appendix 2.1 Mystery Shopping Instrument ..... 52
Appendix 2.2 Online Survey Instrument ..... 53
Appendix 2.3 Case Pilot Test mTurk ..... 555

**Chapter 3: Professional Expertise and the Reference Group Bias**


Abstract ..... 57
3.1 Introduction ..... 58
3.2 Background ..... 60
Reference group bias ..... 6060
Expert decision making ..... 622
3.3 Study 1 ..... 644
Participants ..... 644
Materials and procedures ..... 666
Empirical strategy ..... 688
Results ..... 699
3.4 Study 2 ..... 722
Participants ..... 744
Materials and procedures ..... 755
Empirical strategy ..... 766
Results ..... 777
3.5 Conclusion ..... 8181
Appendix 3.1 Results and Instructions (study 1) ..... 822
Appendix 3.2 Results and Instructions (study 2) ..... 877


**Chapter 4: Free to Help?**


Abstract ..... 93
4.1 Introduction ..... 944
Free will belief and social behaviors ..... 944
4.2 Methods ..... 999
Procedure ..... 999
Treatment ..... 999
Decision task ..... 100100
Questionnaire ..... 1022
Subjects and randomization ..... 1023
Ethics statement ..... 1033
4.3 Results ..... 1043
Manipulation check ..... 1044
Treatment effects ..... 1045
Thou Shalt Help? ..... 1088
4.4 Discussion ..... 1099
Limitations and strengths ..... 110

Practical implications and future research ..... 11010
Appendix 4.1 ..... 1122

**Chapter 5: Choosing for Colleagues**
Abstract ..... 119
5.1 Background ..... 120
Ambiguous returns to business training ..... 120
Deciding for others ..... 122
Research question ..... 124
5.2 Methods ..... 124
Setting ..... 124
Procedure ..... 125
Decision task ..... 126
Treatments ..... 128
Identification ..... 130
5.3 Results ..... 132
Self-other decisions ..... 132
Social Distance ..... 134
5.4 Discussion ..... 135
Limitations and strengths ..... 136
Policy implications and future research ..... 138
Appendix 5.1 Instructions ..... 140

**Chapter 6: Choice Complexity, Benchmarks and Costly Information**
Abstract ..... 149
6.1 Introduction ..... 150
6.2 Experimental design ..... 157
Design ..... 157
Experimental procedures ..... 161
6.3 Hypotheses and estimation strategy ..... 162
Hypotheses ..... 162
Estimation strategy ..... 164
6.4 Results ..... 165
Descriptive statistics ..... 165
Main estimation results ..... 167
6.5 Discussion ..... 171
Appendix 6.I Experiment screens ..... 175

Appendix 6.2 Instructions ..... 178
Appendix 6.3 Survey Instruments ..... 182

**Chapter 7: Conclusion**
7.1 Overview ..... 185
7.2 Main conclusions..... 186
7.3 Suggestions for future research ..... 189
7.4 Policy implications for private and public sector ..... 192
7.5 Epilogue..... 197

**Nederlandse samenvatting**.....201
**English summary**.....211
**References** .....213
**Curriculum Vitae**.....233
**Portfolio.**.....234
**The ERIM PhD Series...**.....235

# List of figures

Figure 3-1 WTC and absolute number of lives ..... 69
Figure 3-2 WTC and reference group size ..... 70
Figure 3-3 Project rating and number of lives saved, study 2 ..... 78
Figure 3-4 Project rating and reference group, study 2 ..... 79
Figure 3-5 Evaluation mode, expertise and bias..... 80
Figure 4-1 Treatment effect, by religiosity..... 105

# List of tables

Table 2.1 Description of cases ..... 31
Table 2.2: Data Structure ..... 33
Table 2.3 Fraction of ethical responses per case, baseline ..... 34
Table 2.4 Randomization balance ..... 35
Table 2.5 Compliance with program ..... 39
Table 2.6 Treatment effects on ethical response ..... 40
Table 2.7 Treatment effects on all outcome variables ..... 41
Table 2.8 Treatment effect, by baseline ethicality ..... 42
Table 2.9 Treatment effect, by baseline ethicality ..... 44
Table 2.10 Survey responses, treatment effects ..... 45
Table 2.11 Ethicality ratings, per case ..... 55
Table 3.1 Summary statistics per session ..... 66
Table 3.2 Characteristics of cases for subsidy applications ..... 67
Table 3.3 Self-reported decision criterion ..... 71
Table 3.4 Subject characteristics, study 2 ..... 74
Table 3.5 Overview of cases, study 2 ..... 76
Table 3.6 OLS-estimates of reference group effect, pooled data ..... 82
Table 3.7 Fixed effects estimates of reference group effect, pooled data ..... 83
Table 3.8 Fixed effect estimates of reference group effect, per session ..... 84
Table 3.9 OLS-estimates of reference group effect, pooled data ..... 87
Table 3.10 Fixed effects estimates of reference group effect, pooled data ..... 88
Table 4.1 Overview of decision-tasks ..... 101
Table 4.2 Summary statistics ..... 103
Table 4.3 Manipulation check ..... 104
Table 4.4 OLS regression of treatment effects ..... 105
Table 4.5 Probit regression of treatment effects ..... 107
Table 4.6 Probit analysis of determinants of helping norm ..... 108
Table 5.5.1 Respondent characteristics ..... 125
Table 5.2 Decision tasks ..... 127
Table 5.3 Randomization balance ..... 129
Table 5.4 Self-other differences in risk and ambiguity attitudes ..... 133
Table 5.5 Social distance and risk and ambiguity attitudes ..... 134
Table 6.1 Example Payoff Matrix ..... 158
Table 6.2 Payoffs by choice ..... 160
Table 6.3 Demographics by treatment ..... 166
Table 6.4 Decisions and payoffs by treatment ..... 167
Table 6.5 Treatment effects on payoffs ..... 168
Table 6.6 Treatment effects on decision quality ..... 169
Table 6.7 Treatment effects relative vs. absolute benchmarks ..... 170
Table 6.8 Treatment effects relative vs. absolute benchmarks ..... 170
Table 6.9 Treatment effects relative vs. absolute benchmarks on advice demand ..... 171

# Glossary of key terms

| Term | Definition |
|---|---|
| Altruism | Behavior that is costly to oneself and beneficial to another person. The term was coined by French philosopher Auguste Comte (1798-1857) as antonym of egoism. |
| Ambiguity | State in which probabilities and/or payoffs associated with outcomes are unknown. Also referred to as "Knightean uncertainty" after economist Frank Knight (1885-1972). Contrasted with risk, in which probabilities and payoffs are precisely defined. |
| Bounded rationality | The notion that bounds to cognitive capacities, decision time and the tractability of the problem lead to limited rationality |
| Cognitive bias | Systematic deviation of rationality in cognition and subsequent decision making (for definition of rationality see below) |
| Complex choice | Choice between multiple options which in turn each consist of multiple attributes |
| Cooperation | Process whereby multiple being work together for mutual benefit. Typically contrasted with competition. |
| Cultural evolution | Development of culture and its constituent behaviors and norms through the process of natural selection |
| Evolution | Change in characteristics of populations over successive generations |
| Field experiment | Application of the experimental method to test theories outside of the laboratory in a "natural" context that more closely resembles the context in which the phenomenon on interest occurs |

| | |
|---|---|
| Free will belief | Degree to which people believe they have free will. Does not specify what is meant by free will. |
| Genotype | The genetic constitution of an organism |
| Indirect altruism | Altruistic behavior which is reciprocated by others later in time, often through the mechanism of reputation |
| Kin altruism | Altruistic behavior directed to genetically related others, for example offspring |
| Multi-level selection | Proposed mechanism of evolution whereby selection pressure on the level of both genes, organism and group |
| Natural selection | Key mechanisms driving evolution through differential survival and reproduction of genes, organisms and groups through differences in phenotype |
| Other-decision | Choice made on behalf of another person |
| Phenotype | Observable characteristics, behaviors and behavioral products of an organism |
| Proportion dominance | Tendency to place more decision weight on relative values than absolute values |
| Rationality | The state of being rational: making judgments and decisions that are logically consistent with one's beliefs and preferences |
| Reciprocal altruism | Altruistic behavior implemented with the expectation that the beneficiary will later act similarly |
| Self-serving bias | The tendency to interpret information in a way that contributes to an overly positive self-image |
| Social preferences | Preferences related to the wellbeing or payoffs of others, also referred to as "other-regarding preferences" |
| Social norms | Informal rules governing the behavior of members of a group |

Chapter 1

# Introduction


## *Summary*

What drives us to help complete strangers? How do cognitive biases shape our prosociality? And what can policymakers and organizations do to promote smart and social decisions? This dissertation presents a series of experiments that aim to provide new insights related to these questions. Chapter 1 sketches how these experiments fit into the existing body of research. Chapter 2 presents a field experiment about the effect of an ethics program on bank employee behavior towards clients. Chapter 3 presents an online experiment about the role of expertise in susceptibility to cognitive biases in the allocation charitable funds. Chapter 4 presents a lab experiment about the influence of beliefs about free will on donation behavior. Chapter 5 presents a lab-in-the-field experiment about attitudes towards ambiguity in decisions made on behalf of others. Chapter 6 presents a lab experiment about the effect of choice architecture on decision quality in complex choices. Finally, chapter 7 concludes.

## *The puzzle of human social behavior*

Humans can be very selfish, *and* very selfless. Consider for example the case of the Indian librarian Paalam Kalyanasundaram who, for a period of 30 years, donated all his salary to the poor (Pareek, 2014). At the same, there are people such as Bernard Madoff, whose firm, according to prosecutors, engaged in fraud amounting to nearly $65 billion, including fraud with funds from charitable organizations (Arvedlund, 2009). These anecdotes suggest that people differ strongly in what economist call "social preferences", the propensity to cooperate and help others. Whereas economic theory traditionally assumed that humans are narrowly self-interested, it is now widely accepted that this assumption is often violated (Fehr and Schmidt, 1999; Charness and Rabin, 2002; Gintis et al., 2003; Falk and Fischbacher, 2006; Thaler, 2015a; Bowles, 2016). Similarly, the assumption that people pursue their goals in a rational manner is often violated because of deeply rooted cognitive biases and limits to human cognitive capacities, so-called bounded rationality (Simon, 1982; Gigerenzer and Selten, 2002; Kahneman, 2003a).

This thesis revolves around these two central topics in the field of behavioral economics; (i) social preferences and (ii) bounded rationality and cognitive biases. It also considers their interplay. Cognitive bias and bounded rationality can result in people making sub-optimal choices for themselves, for example when gamblers think they are more likely to beat the casino after a streak of losses (Croson and Sundali, 2005). However, biases and bound rationality can also stand in the way of people's helping of others. Consider someone who wants to donate money to charity to help people in developing countries. Due to the inclination to focus more on proportions than absolute quantities, people often focus more on overhead cost ratios than on the absolute number of people that are helped. Because of this "overhead aversion", charitable donations are often not allocated in a manner that achieves the most good (Gneezy et al., 2014).

Social preferences and bounded rationality play a role in many economic phenomena, ranging from charitable donations to bank fraud. Before presenting the chapters of the thesis, I now sketch the literature on these topics, spanning from philosophy and early economics to biology and psychology and ultimately back to the field of behavioral economics.

## *Social nature in philosophy and economics*

For many centuries philosophers have been fascinated by the topic of the social nature of mankind. For example, Aristotle (384–322 BC) wrote in his "Rhetoric" that *"Most man are rather bad than good and the slaves of gain…as a rule men do wrong whenever they can"*.

In contrast, Aristotle's contemporaries in China such as the Confucian philosopher Mencius (372 – 289 BC) posited that humans are good natured but require society to cultivate their innate goodness: "*If you let people follow their feelings, they will be able to do good. This is what is meant by saying that human nature is good"*

Similarly, thinkers in the emerging field of political economy in the 17th and 18th century held widely differing views about human social nature and its role in society. For example, Bernard Mandeville (1670-1733) wrote in his "Fable of the Bees" that virtue or altruism was not needed for a functioning society but rather that: "…*The great Taskmasters Necessity, Avarice, Envy and Ambition… keep the Members of the Society to their labour, and make them submit, most them cheerfully, to the Drudgery of their Station, Kings and Princes not excepted*" (Mandeville, 1714).

In contrast, Adam Smith (1723-1790) noted in his "Theory of Moral Sentiments" that humans have a sympathetic nature and derive pleasure from the fortune of their fellow men, and share the sorrows of their misery: *"How selfish soever man may be supposed, there are evidently some principles in his nature,*

*which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it."*[1] However, whilst recognizing the social nature of humans, Adam Smith also posited in this same book that humans are by nature primarily focused on satisfying their personal desires: *"Nature has directed us to the greater part of these by original and immediate instincts. Hunger, thirst, the passion which unites the two sexes, and the dread of pain, prompt us to apply those means for their own sakes, and without any consideration of their tendency to those beneficent ends which the great Director of nature intended to produce by them"* (Smith, 1759).

Smith's contemporary David Hume (1711-1776) wrote in his "Essays: Moral, Political and Literary" that policymakers should assume that humans only care about their own personal interest: *"In contriving any system of government, and fixing the several checks and controls of the constitution, every man ought to be supposed a knave, and to have no other end, in all his actions, than private interest"* (Hume, 1758).

For many years, philosophers and early political economists continued to disagree about the nature of human sociality. Then, in 1859 Charles Darwin published "The Origin of Species". This book presented a ground-breaking theory of evolution through natural selection, which strongly challenged the prevailing view of humans as being "intelligently designed" and fundamentally distinct from all other species. Furthermore, Darwin's theory provided a framework by which to explain the social nature of animals and humans alike.

[1] Smith further wrote: "…*Of this kind is pity or compassion, the emotion we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. That we often derive sorrow from the sorrows of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous or the humane, though they perhaps may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it."*

## *The evolution of human sociality*

Approximately half a century after Hume and Smith published their seminal theories on human nature and sentiments, the political economist and demographer Thomas Robert Malthus published the book "An Essay on the Principle of Population" in which he observed that given the human tendency to reproduce, the limited availability of resources will result in famine and starvation, unless births are controlled: "*…Yet in all societies, even those that are most vicious, the tendency to a virtuous attachment is so strong, that there is a constant effort towards an increase of population. This constant effort as constantly tends to subject the lower classes of the society to distress and to prevent any great permanent amelioration of their condition"* (Malthus, 1798).

It was Malthus' theory on the relation between resource availability, population dynamics and the consequent "struggle for survival" that inspired Charles Darwin to formulate his theory of evolution through natural selection. Darwin's key insight was that the struggle for survival would drive nature to select certain organisms with traits adaptive to this competitive environment to be more successful in reproducing, hence driving this variant of the organism to be more numerous in subsequent generations (Darwin, 1859). Although at the time Darwin was still unsure about the mechanisms by which such variation arises and transmits, which only became clear with the discovery of the structure of DNA in the 1950's (Watson and Crick, 1953).

This so-called "modern synthesis" of evolutionary theory and genetics proved fruitful not only for understanding the origins of species, but also to understand the origins of social behavior. Firstly, from an evolutionary standpoint it can make good sense for an organism to be altruistic towards other organisms that are genetically related. When the British geneticist David Haldane was asked whether he would give his life to save his drowning brother he famously replied "*no, but I would to save two brothers or eight cousins*". The rationale here is that

siblings typically share half of their genes, and cousins one eighth, so that by saving two brothers you save – statistically speaking – the same number of genes that you would if you would preserve yourself. This type of altruism is referred to as kin selection and is described by Hamilton's rule "r>c/b" which states that acts of kin altruism will occur if the degree of relatedness between the altruist and the recipient exceed the cost/benefit ratio of this act (Hamilton, 1964).

Aside from kin altruism, humans also provide costly help to non-relatives. Robert Trivers proposed that for repeated interactions, it can be advantageous to both parties to behave in a mutually altruistic manner, whereby the roles of altruist and recipient are rotated over time, such that both parties benefit compared to selfish or non-cooperative strategies which have higher short-term payoffs, but will undermine cooperation in the future (Trivers, 1971).

And even in non-repeated interactions between non-related parties altruism may be observed. One proposed explanation for such behaviors is through reputation building. If others know that you were helping someone, they will in turn be more likely to help you in the future (Alexander, 1987; Nowak and Sigmund, 2005). It has been proposed that the evolution of indirect reciprocity was intertwined with the development of cognitive capacities to keep track of other people's behavior and make moral judgments as well as the development of strategies to deceive others about one's reputation, so-called "second-order free-riding" (Fowler, 2005). Various examples of indirect reciprocity have been demonstrated, for example in the domain of charitable donations, where subjects preferentially supported others that had donated to a charity (Milinski et al., 2002).

However, not all behaviors can be explained with the theories of kin-, reciprocal- altruism and indirect altruism. Consider for example a soldier that gives his life to defend his country from invaders. It is argued that this tendency to help group members but behave aggressively towards outsiders, so-called "parochial altruism", evolved as a strategy in an environment where violent between-group conflicts were common (Choi and Bowles, 2007). Even though

human societies have grown much larger and more complex since humanity migrated out of Africa, tendencies towards parochial altruism are still observed among humans both in small and large-scale societies (Bernhard et al., 2006; Dreu et al., 2010). Furthermore, recent experiments suggest this tendency is rather intuitive even in the latter group, suggesting that our evolutionary past still resonates clearly in our present sociality, despite considerable changes in our environment (Dreu et al., 2015).

## *Social preferences and bounded rationality*

Evolution been a key force in shaping not just our social preferences, but also our cognition and behavior more broadly. This is the central notion in evolutionary psychology (Barkow et al., 1995). Consider for example our love for sugary and fatty foods. Whereas in our ancestral environment it was beneficial to eat as much of these foods when they were available – to build up some reserves for future periods of hunger – in our modern "supermarket society" this same preference can drive people to overeating and consequent health problems. In other words, while we are hardwired to optimize evolutionary success our mental machinery is not necessarily calibrated to the conditions of our present-day world.

In addition, there are fundamental limits to the cognitive capacities of humans. Although a human brain is immensely complex - comprising approximately one hundred billion neurons – its computational power is limited nevertheless. In addition to the bounds to our computational power, time constraints make it even harder to make optimal decisions. These limitations imply so-called "bounded rationality", whereby people do not necessarily choose an optimal solution but rather one that is satisfactory (Simon, 1982). As a result of these evolutionary imprints and inherent bounds to our rationality we sometimes make choices that do not maximize our personal long-term wellbeing, as

illustrated for example by the rise of obesity and consequent health problems (Afshin et al., 2017).

Furthermore, irrationality may also occur when humans want to cooperate or help others. Consider for example someone who wants to donate 50 euros to save the lives of children in developing countries. This person may have two options. They can donate to charity A, which sends the money to a poor family with two hungry children; or to charity B, which provides bed-nets in areas with high malaria prevalence. The donation to charity A is expected to save the lives of the two hungry children, while the donation to charity B is expected to save the lives of five children. Given the objective of saving as many children as possible would be thus be rational for this potential donor to select charity B. However, this person may nevertheless donate to charity A because it can show a picture of the specific two children that will be helped, whereas in the malaria prevention case these five potential victims are still "statistical". Since pictures of hungry children typically trigger feelings of empathy (Burt and Strongman, 2005), the donor may select choose to help the "identifiable victims" although more children could been saved otherwise (Jenni and Loewenstein, 1997).

## *The behavioral synthesis*

Throughout most of the 20th century, economics models have been built on the notion that that humans are rational and narrowly self-interested. However, this view has been challenged in recent decades. Through integration of insights from psychology and the application of experiments, the emerging discipline of behavioral economics has identified systematic deviations from rationality in the form of various cognitive biases (Kahneman et al., 1991; Kahneman, 2003a; Mullainathan and Thaler, 2000). Furthermore, behavioral economics has shown that humans systematically deviate from rational self-interest: they donate to strangers (Engel, 2011); they reject offers which they perceive as unfair (Fehr and

Schmidt, 1999; Thaler, 1988); they punish cheaters at a cost to themselves (Fehr and Gächter, 2002); and they care – even when being merely passive observers with no "skin in the game" – that others are treated fairly as well (Fehr and Fischbacher, 2004).

Furthermore, through the increasing use of experiments with diverse populations in different settings, a better understanding is emerging regarding the ways by which they cognitive biases and social preferences are shaped by culture, beliefs, social norms and other local conditions (Charness and Fehr, 2015; Henrich et al., 2005; Levitt and List, 2009). Moreover, the genetic and neural mechanisms underlying these behaviors are increasingly studied in the emerging disciplines of geno- and neuro-economics (Benjamin et al., 2012; Glimcher and Fehr, 2013). Finally, insights in behavioral economics are increasingly obtained and applied outside of academia, with governments, businesses and other organizations using behaviourally informed experiments to investigate how- and how well their activities, interventions and policies work.

## *Overview*

This thesis aims to contribute to the field of behavioral economics by means of a series of lab- and field-experiments focused on the causes, consequences and interplay of social preferences and bounded rationality. Table 1.1 provides an overview of the focus of the various chapters.

Table 1.1 Focus, per chapter

| | | Experimental manipulation | | | | |
|---|---|---|---|---|---|---|
| | | Ethics training | Size of reference group | Text about free will | Choice for self/ other | Bench-mark product |
| Main outcome | Helpfulness towards client | Ch 2. | | | | |
| | Rating of charitable project | | Ch. 3 | | | |
| | Amount donated to charity | | | Ch. 4 | | |
| | Demand for business training | | | | Ch. 5 | |
| | Selection of financial product | | | | | Ch. 6 |

Chapter 2 considers the role of culture and knowledge in ethical behavior in the banking sector. This study builds on the finding that the professional culture in the financial sector contributes to dishonest and unethical behavior of employees (Cohn et al., 2014). Using a nationwide randomized controlled trial with a commercial bank, we test if a so-called "ethics program" whereby employees discuss work-related ethical dilemmas promotes more client-centric behavior. In addition to changing the culture and social norms, this intervention also aims to serve as a platform for knowledge exchange between employees. Using a newly developed version of the audit study methodology this study provides novel evidence about effectiveness of ethics programs.

Next, chapter 3 investigates the role of professional expertise in susceptibility to the psychological bias known as proportion dominance. Given that proportions can typically be more easily evaluated than absolute quantities, many people prefer charitable projects were large fractions of a group are saved,

even if these projects do not save the greatest number of people in absolute terms (Bartels, 2006). Given that previous studies on proportion dominance and this so-called "reference group bias" were conducted only among students, this study compares susceptibility to this bias among professionals from the charitable sector and a group of students.

Subsequently, chapter 4 focuses on the role of personal beliefs about free will on charitable donations. This study builds on previous findings that people become inclined to cheat for personal financial gains when their belief in free will is undermined (Vohs and Schooler, 2008), which in turn builds on the theory self-serving bias whereby people attribute negative behaviors to external factors. Using an online experiment to test the effect of free will belief on donation behavior, this study extends the literature on self-serving bias to free will beliefs and altruistic behaviors.

Chapter 5 considers how ambiguity - the lack of information about payoffs and probabilities - influences how people make choices for themselves and others. This study builds on the theory that decision-making for others involves greater social and psychological distance which results in more abstract and distanced thinking (Trope and Liberman, 2010). Since previous research has shown differences in risk attitudes for self- and other-decisions, this raises the question whether attitudes towards ambiguity also differ in choices for others. Using a lab-in-the-field experiment with entrepreneurs to investigate this question, this study extends the empirical literature on self-other differences to a new domain of decision-making and a new population of subjects outside of the typical population of so-called "western, educated, industrialized, rich and developed" (WEIRD) subjects used in most lab experiments (Henrich et al., 2010a).

Finally, chapter 6 addresses the topic of complex choices, where people face multiple options which each consist of multiple attributes. As such choices are computationally complex and people have bounded rationality, such choices – for example in the context of pension schemes, health insurance and mobile phone

subscriptions, can lead to suboptimal decisions. A growing body of research suggests that decision-making in such contexts can be improved without reducing the choice set, but merely by changing the "choice architecture", the way in which the information is presented[17]. Building on this notion, this study uses a lab experiment where subjects choose between financial products to investigate how decision quality is affected by reference points that provide the market average for each product. The rationale for this approach is that reference points make it easier for consumers to compare costs of product features.

Table 1.2 Research type and contribution, per chapter

| | | **Type** | | |
|---|---|---|---|---|
| | | Replication | Extension | Innovation |
| **Contribution** | Theory | | C2 | |
| | Methodology | | | C2 |
| | Empirical | C4 | C3, C4, C5 | C2, C6 |

As such, this thesis presents five chapters on the drivers of social preferences and bounded rationality. The main contribution of these chapters is empirical, both in terms of replication and extension (chapters 3, 4 and 5) and in terms of novel evidence (chapters 2 and 6). Furthermore, chapter 2 introduces a newly developed method to study unethical behavior in the financial sector. Finally, chapter 2 extends the theory about the role of culture and bounded rationality in social preferences to the design of an ethics program for a large commercial bank. The next section provides an overview of the respective chapters.

## *Outline of chapters*

Table 1.3 Overview of chapters

| Ch. | Title | Research question | Sub-questions | Theoretical perspective | Methodology |
|---|---|---|---|---|---|
| 2 | Better bankers | What is the effect of an ethics program on bank employee client-focus? | How does the program work? | Culture, bounded rationality | Field experiment with employees in bank branches |
| 3 | Professional experience and the reference group bias | Are professionals in charitable organizations less susceptible to the reference group bias than students? | Is the reference group bias reduced under joint evaluation of projects? | Cognitive bias and proportion dominance | Lab experiment with students and professionals from charitable organizations |
| 4 | Free to help? | What is the effect of free will belief on charitable donations? | Does this effect differ for religious and non-religious people? | Motivated reasoning and social preferences | Online experiment with subjects in Amazon mTurk |
| 5 | Choosing for colleagues | How do ambiguity attitudes compare for choices people make for themselves vs. for others? | Could delegated decisions improve business outcomes? | Ambiguity aversion, self-other differences in decision-making | Lab-in-the-field experiment with entrepreneurs in Bangladesh |
| 6 | Choice complexity, benchmarks and costly information | Do consumers facing complex choice problems make better decisions when shown a benchmark product? | Are benchmarks more effective when attributes expressed in relative terms? | Complex choices, decision aids | Lab experiments with students |

Table 1.3 (continued)

| Data | Main results | Co-authors | Status | Outlet |
|---|---|---|---|---|
| (i) Behavioral data from mystery shopping visits before and during program, (ii) post-experiment survey | No effect on average, negative effect for sub-group of branches with high baseline variance in perceived ethicality | Alain Cohn, Ernst Fehr, Karen Maas, Michel Marechal, Helen Toxopeus | Working paper | |
| (i) Reported willingness to support projects with variable absolute and relative life savings (through vignette), (ii) post-experiment survey | Students and experts' both susceptible to reference group bias, joint evaluation reduces bias particularly for students | Dinand Webbink | Submitted, under review (2017) | Judgment and Decision Making |
| (i) Behavioral data from 24 incentives binary dictator games (donations to charity), (ii) post-experiment survey | Free will disbelief reduces charitable giving, but only for non-religious subjects | Kellie Liket, John Protzko, Vera Scholmerich | Published (2017) | PLoS ONE |
| (i) Vignette to measure attitudes towards risk and ambiguity in self- and other-decisions, (ii) post-experiment survey | Subjects display similar degree of ambiguity aversion in self- and other-decisions, higher degree of AA for other-choices conditional on belief that prob.<0.5 | Karen Maas, Haki Pamuk | Revise and resubmit (2018) | Small Business Economics |
| (i) Behavioral data from incentived choice tasks, (ii) post-experiment survey | Decision making is improved with benchmark products with attributes framed in relative terms | Mark Sanders, Stephanie Rosenkranz | Submitted, under review (2017) | Journal of Econ. Behavior and Org. |

Chapter 2: Better bankers: a field experiment

Despite widespread misbehavior in the financial sector, little is known about the causal effects of policies to promote ethical behavior among bank employees. In this field experiment, we test the effect of an ethics program in which bank employees jointly discuss ethical dilemmas they encounter in their interactions with clients. This ethics program was developed jointly with the bank and implemented for a period of circa two months. Ethical behavior is measured using an audit study in which mystery shoppers – actors presenting themselves as potential clients – elicit advice from bank employees about various financial products. The scripts for these mystery visits were constructed such that it was in the interest of employees to pursue a product sale, rather than to inform the client that the financial product may not be suitable. We find that the program promotes ethical behavior in shops with low levels of baseline ethicality, whereas an opposite effect is found among shops with higher baseline ethicality. Survey data suggests this differential effect resulted partly from the fact that employees perceive the program as a signal about the social norm, with conformity preferences driving behavior towards this norm.

Chapter 3: Professional expertise and the reference group bias

People have a tendency to focus more on relative than absolute quantities. For example, most people have greater liking of a filled small cup than a three times larger cup that is half-full. This so-called "proportion dominance" effect has also been shown to play a role in the field of charitable donations, as private donors have tendency to favour projects that help larger fractions of smaller groups rather than projects that help smaller fractions but a larger overall number of people. This effect has been demonstrated in experiments with student populations, but it is unclear whether professionals in charitable organizations behave similarly. Through a combined experiment with students and professionals from the charitable sector in the Netherlands, we find that both groups are

similarly prone to the reference group bias. This finding adds to a large literature on expert decision-making that indicates that many cognitive biases are insensitive to professional expertise. In turn, this study calls for interventions to steer professionals in the charitable sectors towards allocating funds more based on the absolute number of people that can be helped.

Chapter 4: Free will belief and altruism

Do we have free will? This ancient and unresolved philosophical question has surprising implications for the social sciences. Previous studies show that reduced belief in free promotes dishonest behavior, suggesting that people justify their selfish tendencies by not believing in free will. Building on this work, we investigate whether free will belief also shapes altruistic behavior. In an online experiment subjects' belief in free will is primed through a reading task. We find that belief in free will influences charitable giving among non-religious subjects, but not among religious subjects. This could be explained by our finding that religious subjects associate more strongly with social norms that prescribe helping the poor, and might therefore be less sensitive to the effect of reduced belief in free will. These results indicate the effect of free will belief are more nuanced than suggested in previous studies.

Chapter 5: Ambiguity attitudes in self-other decisions

Many business decisions are made on behalf of others, and involve ambiguity. Whereas it is widely known that most people are averse to ambiguity, it remains unclear if they display similar levels of ambiguity aversion when making choices on behalf of others. We conduct a survey experiment among entrepreneurs in Bangladesh to investigate such self-other differences in ambiguity attitudes. Subjects are presented various hypothetical choices between certain payoffs and ambiguous or risky payoffs. The results indicate that entrepreneurs are less ambiguity averse when deciding for colleagues compared to deciding for

themselves. This self-other difference is only found for entrepreneurs that believe the ambiguous outcome to occur with a probability of less than fifty percent. Furthermore, we find that subjects are more risk averse when choosing for others. Finally, we find that social distance between the decision-maker and the other person does not explain ambiguity attitudes.

Chapter 6: Choice complexity, benchmarks and costly information

Modern life confronts us with many complex choices between numerous multi-attribute options, such as in the case of selecting a mortgage or pension scheme. Given bounded rationality, people often make suboptimal decisions in such situations, which can in turn lead to societal problems, as the subprime mortgage crisis illustrated. One approach towards enabling consumers to make better decisions in such situations is to provide so-called 'benchmarks' to which the available options can be compared. We experimentally investigate the effect of information interventions on decision quality in complex choices. Choice options were framed as financial products and could be objectively ranked. In our benchmark treatments one option was revealed as having average values for all attributes, either in relative or absolute terms. In our costly information treatment two options were revealed as being suboptimal. We find that costly information and relative benchmarks improve decision quality. Finally, the provision of benchmarks has limited effects on demand for costly information.

Chapter 2

# Better Bankers

## A Field Experiment*


## Abstract

Despite widespread misbehavior in the financial sector, little is known about the causal effects of policies to promote ethical behavior among bank employees. In this field experiment, we tested the effect of an ethics program in which bank employees jointly discussed ethical dilemmas they encounter in their interactions with clients. This ethics program was developed jointly with the bank and implemented for a period of two months. Ethical behavior was measured using an audit study in which "mystery shoppers" – actors presenting themselves as potential clients – elicited advice from bank employees about various financial products. The scripts for these mystery visits were constructed such that employees faced a trade-off between pursuing a product sale and informing the client that the financial product might not be suitable. Two key findings emerged. First, the program had no significant effect and the majority of employees in both the treatment and control groups did not provide advice that places the client central. Second, a negative treatment effect was observed for bank shops with above-median levels of variance in baseline ethicality. This suggests that the group


*Joint work with Alain Cohn, Ernst Fehr, Karen Maas, Michel Maréchal and Helen Toxopeus

meetings may have inadvertently provided a platform for less ethical employees to undermine the client-focus of their colleagues.

## 2.1 Introduction

### *Rationale of the study*

There is widespread "misbehavior" in the financial sector. This is illustrated by a number of widely publicized fraud cases, such as the LIBOR manipulation scandal (Abrantes-Metz et al., 2012) and the recent Wells Fargo scandal involving cross-selling of products and the creation of fraudulent accounts[2]. Contrary to commonly heard criticisms that such cases are exceptional, it was recently found that approximately seven percent of financial advisors in the United States have misconduct records (Egan et al., 2016), suggesting that such misbehavior is rather commonplace. Similarly, an audit study found that many financial advisors reinforce their clients' biases in order to further their own interest, at the cost of their clients (Mullainathan et al., 2012). Correspondingly, consumer trust in the banking sector is low[3][4], with approximately two-thirds of US citizens indicating they have some, limited or no trust in banks. Another recent global survey revealed that only around one-quarter of bank customers trust their bank to provide them with unbiased advice[5].

This misbehavior in the financial sector is recognized by policymakers as a key risk to financial stability and has been attributed to a number of factors, including financial incentives to sell certain products, irrespective of whether these are

[2] See, for example, the Wall Street Journal article "How Wells Fargo's High-Pressure Sales Culture Spiraled Out of Control" by Emily Glazer (2016), accessed 12-09-2017

[3] According to a 2017 Gallup poll, 67% of the US population had some, limited or no trust in banks. A similar result was found in another 2015 survey, indicating that financial advisors are often perceived as dishonest and untrustworthy (Anna. 2015. _Brokers are Trusted Less than Uber Drivers, Survey Finds._ Wall Street Journal.)

[5]Source: Ernst & Young 2016 Global Consumer Banking Survey, accessed 12-09-2017 via: http://www.ey.com/gl/en/industries/financial-services/banking---capital-markets/ey-global-consumer-banking-survey-2016

optimal for the client (Danilov et al., 2013; Inderst and Ottaviani, 2009). In addition, such misbehavior has been attributed to the business culture in banks and financial institutions. A recent experiment by Cohn et al. (2014) showed that bank employees primed with their professional identity were more likely to cheat for financial gains. This result suggests that it is the professional culture and organizational norms in banks, rather than the nature of employees per se, that is promoting dishonesty and misbehavior.

Several strategies have been proposed to address this issue. Firstly, it has been argued that financial incentives should be changed, for instance by reducing sales incentives and introducing "clawbacks" by which bonuses can be revoked if products do not continue to perform over longer periods of time, so as to curtail short-termism. For example, in 2014 the European Union introduced a bonus cap for bank employees at a maximum of one hundred percent of the fixed salary[6]. In the Netherlands, bonuses were capped even further, at twenty percent of base pay[7]. Although economic theory suggests that such measures will reduce employee motivation to sell products which do not serve the clients' best interest, there is limited empirical evidence about the effect of such reforms on bank employee behavior.

Furthermore, it has been argued that the organizational culture in banks should promote ethical behavior. To this end, various interventions have been proposed including: ethics codes, integrity oaths, ethics trainings and ethics reminders. As for codes of conduct, meta-analytic research suggests that the mere existence of such codes does not significantly increase ethical behavior, and might even reduce ethical behavior in the absence of enforcement, presumably because it signals that the bank implements this ethics code as a form of "window dressing".

[6] Source: Directive 2013/36/EU of the European Parliament and of the Council of 26 June 2013 on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, amending Directive 2002/87/EC and repealing Directives 2006/48/EC and 2006/49/EC Text with EEA relevance § 94

[7] Source: https://www.law.ox.ac.uk/business-law-blog/blog/2017/04/dutch-bonus-cap-and-brexit

In contrast, active enforcement of an ethics code is associated with reduced levels of unethicality (Kish-Gephart et al., 2010).

Regarding the proposal of an ethics oath for the banking sector (Boatright, 2013), similar in spirit to the Hippocratic oath in the medical profession, the Netherlands has recently introduced a so-called "bankers' oath" requiring all employees to swear to place the clients' interests central in their decision-making[8]. Lab experiments suggest that commitments and promises, even when non-binding, can promote cooperation and honesty, presumably because people have a preference for behaving according to their promises (Charness and Dufwenberg, 2010; Vanberg, 2008). In a similar fashion, signing an insurance claim form at the beginning rather than the end was shown to significantly reduce insurance fraud, suggesting the ex-ante signature induced a commitment to ethical behavior (Shu et al., 2012). In sum, these results suggest that an integrity oath for the banking industry could promote client-centric behavior, but field evidence is still lacking.

As for ethics programs and trainings, many different initiatives have been implemented by banks. According to a 2015 survey by the US Ethics Resource Center[9], comprehensive programs are related to reduced misconduct, increased reporting of misconduct and reduced pressure to compromise standards. However, given the observational nature of this research, it remains unclear to what extent this reflects a causal effect. Furthermore, it is not clear which components of these programs are most effective (Treviño et al., 2014). For example, some programs focus more on individualized learning about ethical behavior, whereas others involve more social processes such as speaking up in a group. Lab research suggests that various forms of peer influence such as discussing ethical issues with others can promote ethical behavior (Gino and Pierce, 2009; Gunia et al., 2012) but it remains unclear how effective such approaches are in the context of

[8] Source: Dutch Central Bank, accessed 12-09-2017 via: https://www.dnb.nl/en/publications/dnb-publications/newsletters/nieuwsbrief-banken/nieuwsbrief-banken-januari-2013/dnb284394.jsp
[9] Source: Ethics & Compliance Initiative, accessed 13-09-2017 via: https://connects.ethics.org/viewdocument/the-state-of-ethics-in-large-compan

the banking sector.

Another strategy that has been suggested is to use so-called "ethical reminders". This approach builds on the notion that humans have limited memory and may forget about their good intentions and not have others' interest in mind during the critical moment of decision-making. Reminders have been successfully applied in the field to promote behaviors such as donating blood (Stutzer et al., 2011; Vuletić, 2015) and voting in presidential elections (Dale and Strauss, 2009), as well as in more personal decisions such as timely repayment of loans (Cadena and Schoar, 2011) and adhering to saving plans (Karlan et al., 2016). However, it remains unclear whether such reminders are also effective within the cultural context of commercial banks. It could be the case that moral reminders mainly promote ethical behavior among people who were already inclined towards such behaviors but that they are less effective in restraining unethical behavior, such as cheating a client. Furthermore, in the presence of financial incentives to sell products to clients, it is even more questionable how effective moral reminders and ethics programs will be in changing employee behavior, i.e. in restraining them to sell suboptimal products to clients.

## *Overview of experiment*

In sum, there is limited evidence about the effectiveness of various "soft measures" to promote ethical behavior in the banking sector. To address this knowledge gap, we conducted the first field experimental test of the effect of an ethics program on client-focused behavior among employees in a large commercial bank. To this end, we worked with a bank to develop an ethics program for front-office staff which we then randomly assigned among a total of N=94 bank shops spread across the country where this bank operates. The main component of this intervention was a weekly meeting in which employees discussed an ethical dilemma they encountered in their daily work. Furthermore,

employees were encouraged to implement a new slogan when greeting clients, which aimed to serve as a moral reminder. However, compliance with this component was very limited, so the main component of the treatment was the weekly meetings.

To measure bank employee behavior, we implemented an audit study, whereby bank employees were not aware that they were being studied. This method is particularly suitable for studying behaviors that are socially undesirable and which people may wish to hide if they know they are being studied. The audit study method has been successfully used to study racial and sex discrimination in the labor market (Bertrand and Mullainathan, 2004; Neumark et al., 1996). More recently, this method has also been used to study unethical behavior in the financial sector, for example among financial advisors (Mullainathan et al., 2012) and in the context of setting up shell companies for tax avoidance (Sharman, 2010). In our study, we worked with a professional "mystery shopping" firm which sends professional actors to organizations, often for the purposes of testing the service quality.

We developed standardized case scripts that the mystery shoppers used to elicit advice from front-office bank employees. These case scripts constituted a query where the client's interest did not align with the interest of bank employees, who have financial incentives to sell financial products such as credit and insurance. The main outcome we measured was whether the bank employee provided an advice which served the client's best interest, or whether their advice steered the client towards a product sale, thereby serving their own interest. Every bank shop was visited six times before the ethics program was implemented and six times during the period when the program was implemented in the treatment shops. More details on the audit methodology are provided in section 2. After completion of the ethics program, we sent a survey to all bank employees to gather information about their attitudes towards the program and their perceptions regarding social norms in the bank.

The main results of our study are as follows. First, we find that during the baseline in the majority of audit visits the bank employee did not provide the advice which best served the client's interest. Second, we find that the ethics program had no effect on the bank employee's behavior in the audit study, suggesting it was ineffective in promoting more ethical behavior. Third, we find that the ethics program had a negative effect for shops with high variance in baseline ethicality, while a positive effect was observed for shops with low variance in, and absolute levels of, baseline ethicality. This result suggests "ethical contagion", whereby the ethics program provided a platform for the less ethical employees to exert a negative influence on their colleagues' norms and behavior. The positive effect for the low-variance group suggests that without such "contagion", group discussion can indeed promote more ethical behavior.

## *Main contributions*

To the best of our knowledge, this is the first study that applies a randomized controlled trial to estimate the causal effect of an ethics program within a company. As such, it builds on a growing literature of field experimental evidence on the dynamics of ethical behavior in organizations (Pierce and Balasubramanian, 2015; Pierce and Snyder, 2008). The evidence we present is relevant for academics trying to understand the drivers of unethical behavior as well as for policymakers and organizations who wish to limit such behavior, for example in the financial sector. Our main result, that the ethics program had an insignificant net effect, casts doubts on the notion that soft measures alone will suffice to cut down on unethical behavior in the financial sector. It also suggests that further revisions of the financial incentives, such as bonuses for product sales, might also be needed. In addition, our results warrant caution in designing and implementing ethics programs in an organizational context. Given that such ethics programs might be interpreted by employees as signals about how other

employees behave, preferences for conformity can drive down ethical behavior among workers with higher initial levels of ethical behavior.

Furthermore, we contribute to the growing body of literature that uses audit methodologies to measure ethical behavior in the financial sector. Our design is similar in spirit to Mullainathan et al. (2012), who use a mystery shopping methodology to study how financial advisors fail to de-bias their clients in order to advance their own personal interests, for example by selling investment products with higher agent fees. In a similar fashion, we find that many of the front-office bank employees in our sample fail to inform clients when credit and insurance products might not be suitable. Our results contribute to the literature that is showing how "misbehavior" in the financial sector is still widespread (Egan et al., 2016; Lo, 2015), and calls for increased efforts to promote ethical behavior in the financial sector.

### *Outline*

The paper proceeds as follows. In chapter 2 we provide an overview of the ethics program, the experimental design and the audit study methodology. In chapter 3 we provide the results of the baseline study, the main treatment effects and the results from our ex-post survey. We then discuss these results in chapter 4, with a particular focus on the differential effect of the program on employees in shops with high vs. low baseline levels of ethicality. Finally, in chapter 5 we conclude.

## 2.1 Experimental design

### *Ethics program*

The program was developed jointly with a large commercial retail bank, which provides a range of financial products and services to its clients. During the period 2015-2016, a number of meetings were held between the research team and various managers from the banks to discuss possibilities for ethics programs that

could be randomly assigned to either individual employees or bank shops. In 2016, it was agreed that the program would consist of two main components: (i) a weekly meeting in which employees would discuss ethical dilemmas and (ii) a new client greeting to make the topic of ethics more salient during the critical moment of front-office employees' interactions with clients. Subsequently, the program materials were developed. These included a manual for the managers of the selected bank shops to use during the program kick-off, a list of example dilemmas to be discussed during the weekly meetings and instructions with the wording of the new client greeting. The program was then launched among the selected bank shops in February-March 2017 and ran for a period of six weeks. Two weeks prior to the start of the program, managers received an email and video message from the bank's director, as well a set of documents with instructions on how to implement the program. We will now describe in more detail the two main components of the program.

Managers were instructed to integrate the weekly meetings into their pre-existing weekly team meetings, during which they discuss operational questions with the shop employees. These meetings typically last 15-30 minutes and are often implemented at the beginning of the work week, so specific goals and targets can be discussed with the team. The ethics program started with a kickoff meeting, in which the manager communicated to the shop employees that the bank headquarters promotes a culture in which employees openly discuss ethical dilemmas with each other so they can give each other feedback, learn from each other and promote a "client-first" approach. Furthermore, it was explained exactly what is meant by an ethical dilemma; namely a situation in which it is not in the client's best interest to purchase a financial product. Subsequently, the manager presented an example of what such a dilemma could practically entail, either from a list of pre-constructed cases or from his/her own personal work experience.

Subsequently, employees were encouraged to think of ethical dilemmas to discuss in the upcoming weekly meetings, as well as to discuss these with their

colleagues. In the following weeks, the employees held meetings in which they discussed these dilemmas with each other. These meetings were held in the back office and typically lasted between 10-15 minutes. In addition, the bank developed an intranet page for the employees in the selected bank shops to share dilemmas. This intranet page was only accessible using a specific password that was exclusively provided to employees in the treatment shops. Finally, some treatment shops used a message board in the back office where employees could write down ethical dilemmas they encountered. More details on the implementation compliance are provided in section 3.

As for the revised client greeting, the bank and research team jointly developed a short slogan that employees were instructed to use when engaging in a conversation with a client in the front office. This new greeting read as follows: "Hello, I will do my utmost to serve you to the best of my abilities". However, during the rollout of the program it appeared that many employees did not feel comfortable using this slogan, and reported that it felt unnatural to apply it. Correspondingly, an ex-post survey revealed that the slogan was only applied in approximately ten percent of the encounters. Finally, halfway during the six-week program, employees in the treatment shops were sent cakes with the title of the ethics program printed on them, to serve as an additional reminder about the program and boost worker morale concerning implementation of the program.

## *Randomization*

To test the hypothesis that the ethics program promotes client-centered behavior among bank employees, the treatment was randomly assigned at a total of 92 bank shops, clustered by 26 bank shop managers. The results of the baseline audit study are explained in more detail below. Based on these results, in which every bank shop was visited by six mystery shoppers, the shops were ranked in terms of the level of ethical behavior. Through stratified sampling, a total of 45 bank shops

clustered by 12 managers were assigned to the program, with the remaining 47 shops and 15 managers assigned to the control group. As a result, the treatment and control group were balanced in terms of baseline levels of ethicality. Details on the randomization are provided in the results section. This randomization took place in February 2017, after finalization of the baseline data collection which took place in the preceding two months. Subsequently, the six-week ethics program was implemented in the period March-May 2017, with an approximate two-week variation between shops in terms of the exact start and end dates. Simultaneously, the second round of data collection, consisting of another six mystery shopper visits per shop, was implemented.

After the ethics program was finalized and the data were analyzed, an online survey was sent by the bank to all employees in the control and treatment group in June 2017. It was communicated in the email invitation to the survey that the bank was working together with universities in the research team to implement a study about behavior and attitudes towards work at the bank. Participation in this survey was voluntary and invitees were informed that the survey would last between 5-10 minutes. This survey included questions about attitudes towards ethical matters in the workplace, as well as an additional section of questions about participation in the ethics program, which was only presented to the treatment group.

## *Audit study*

To measure ethical behavior, six different case scripts were developed, with two different versions of each script. These scripts described a query by a client made to a front-office bank employee about various financial products, including credit and insurance. The cases were constructed such that the bank employee, in advising the client, was facing a tradeoff between private benefits and serving the client's best interest. For example, in one case the client signaled interest in buying

a so-called in-house insurance policy to cover damages if their pet damaged their new sofa. However, this particular insurance policy does not cover such damages. Due to confidentiality, the exact formulation of the cases cannot be revealed. However, we do provide a conceptual description of each case.

In case 1 the mystery shopper informs the bank employee they want a loan for consumption purposes, and they indicate that they think they should be eligible because they hold considerable savings. This purportedly signals that they may fail to understand that they can also use part of their savings for this consumption. The outcome variable is whether the bank employee informs them about this possibility or not. In case 2, the mystery shopper wants a loan and signals they think the total loan costs will not increase if they take longer to repay it. The outcome variable is whether the bank employee informs them that loan costs will be higher if the repayment period is longer. In case 3, the mystery shopper wants an insurance product but signals they think this insurance covers a risk which is actually not covered. In case 4, the mystery shopper wants a loan for consumptive purposes but also indicates they already have several other outstanding loans with other banks. The outcome variable is whether or not the bank employee informs them that it might not be prudent to apply for more credit. In case 5, the mystery shopper wants a loan to on-lend funds to a friend who cannot get a loan. The outcome variable is whether the bank employee informs them that they still have to repay the loan to the bank even if their friend does not repay them. In case 6, the client wants a loan with a flexible interest rate and signals they do not understand that total costs will increase if the interest rate rises. The outcome variable is whether or not the bank employee informs them about this possibility.

Table 2.1 Description of cases

| Case | Description | More ethical response |
|---|---|---|
| 1 | client wants a loan but indicates he/she also holds considerable savings | inform client to consider using (part of) their savings instead of a loan |
| 2 | client wants a loan but doesn't understand total cost increase with longer repayment period | inform client that total costs go up with longer repayment period |
| 3 | client doesn't understand that insurance won't cover the relevant risk | inform client that particular risk is not covered by insurance product |
| 4 | client wants consumer loan but has outstanding loans at other banks | inform client that they should consider not taking another loan before repaying outstanding loans |
| 5 | client wants loan to on-lend to a friend who cannot get a bank loan | inform client that they are liable to repay the bank even if friend does not repay them |
| 6 | client wants loan with flexible rate but doesn't understand rate can go up | inform client that costs will go up if interest rate will rise |

All these cases were constructed based on extensive discussions with various bank employees both from the headquarters as well as staff with closer ties to the daily operations of the shops in the field. Furthermore, prior to the implementation of the audit study, these case scripts were pre-tested with N=1198 subjects in Amazon mTurk to see whether people indeed considered the less ethical response in the various cases to be unethical/inappropriate. For each case script they were asked to indicate on a scale ranging from -3 to +3 how inappropriate/appropriate they considered the less ethical response from the bank employee. The average rating was -1.35, significantly lower than zero ($P<0.01$), indicating that people indeed considered the described bank employee responses as inappropriate[10].

Subsequently, a professional "mystery shopping" firm was hired and trained to implement these case scripts. A total of N=233 different mystery shoppers were trained by this firm to implement the various scripts during mystery

[10] More details on this pilot test are provide in appendix 2.3

visits to the bank shops. These mystery shoppers had previous experience with implementing audit studies, typically for corporate clients, for example to test sales staff adherence to company protocols. During both six-week phases of the data collection, every bank shop was visited once per week by a mystery shopper. As such, there were a total of twelve waves of data collection. The same bank shop was never visited by the same mystery shopper more than once, so as to avoid recognition by the bank employees. Furthermore, mystery shoppers were recruited from the specific region in which these banks shops were located, so as to avoid making a suspicious impression, for example by not speaking with the local accent. Before the start of data collection, the different cases were randomly assigned to the bank shop – wave combination, such that the ordering of data collection over cases was counterbalanced.

After visiting the bank shop and implementing the specific case script to which they were assigned, the mystery shoppers recorded - directly after leaving the bank shop - how the bank employee had advised them in this situation. In addition, they rated the perceived ethicality of the bank employee on a -3 to +3 scale, as well as their perception of the extent to which the advisor asked questions to better understand their financial situation. These three questions formed the main outcome variables for the analysis[11]. In addition, the mystery shoppers recorded the age and gender of the bank employee, and the number of other employees and clients present in the office. These variables were used as control measures in the estimation of the treatment effect. Responses were entered using a tablet and directly sent to the mystery shopping firm, who in turn transferred these to the research team.

[11] The exact audit survey instrument in provided in appendix 2.2

## 2.3 Results

### *Descriptive statistics*

A total of N=92 bank shops participated in the experiment. Of these, N=45 were randomly assigned to the treatment group and the remaining N=47 to the control group. More details on the randomization follow later in this section. These shops were managed by a total of N=27 different managers[12]. Each shop was visited by mystery shoppers a total of N=12 times, before and during the program, yielding a total of 1.104 unique observations.

Table 2.2: Data Structure

| | Control | Treatment | Total |
|---|---|---|---|
| Managers (distinct) | 15 | 12 | 27 |
| Bank shops (distinct) | 47 | 45 | 92 |
| Total observations | 564 | 540 | 1.104 |

Bank shops have on average 4.12 employees (SD=1.90) working in the shop simultaneously[13]. Most shops have a greater pool of employees, since most employees do not work all hours. Furthermore, employees may sometimes work in another bank shop that belongs to the same manager group, for example to fill in for a sick colleague. Most bank shops were located in the center of cities and larger towns. Both clients and non-clients could walk into the front office of these bank shops without making an appointment. Front-office staff provide information, sell or request products such as consumer loans and help clients with simple matters such as requesting a new ATM card. They are not licensed to provide advice about more complex products such as mortgages; this requires an appointment with a specialized financial advisor. We restricted our study to

[12] For the analysis, observations for shops belonging to a particular manager were excluded as the headquarters reported that this manager was unwilling to implement the program.
[13] This result was obtained during the post-program employee survey.

queries that could be addressed to front-office staff. In this regard, our study differs from Mullainathan et al. (2012), who focused on investment advice. The mystery shoppers reported in 66.3% cases that they encountered a female bank employee. Approximately 70% of employees were 20-40 years, and 20% were between 40-50 years.

## *Baseline*

We will now report the results of the mystery shopping visits implemented during the six weeks prior to the implementation of the ethics program. The first key result is that bank employees provided the more ethical response in 38.22% of the N=552 mystery visits. In the remaining 61.78% of cases the bank employees did not inform the mystery shopper about the product-related misunderstanding or lack of understanding that they signaled during their query. The bank employee responses differed between the six types of cases. As can be seen in table 3 below, ethical responses were most common for case 3 (insurance) and case 5 (borrowing for friend), and lowest for case 4 (borrowing when already indebted).

Table 2.3 Fraction of ethical responses per case, baseline

| Case nr | Ethical advice | Obs. |
|---|---|---|
| 1 | 39.67% | 184 |
| 2 | 31.52% | 184 |
| 3 | 53.80% | 184 |
| 4 | 15.76% | 184 |
| 5 | 51.09% | 185 |
| 6 | 33.70% | 184 |
| Total | 37.59% | 184 |

Furthermore, mystery shoppers were asked to indicate, on a scale running from -3 to +3, the extent to which they perceived that the bank employee put their interest first. The average rating for this question during the baseline was 0.39, with scores

of zero or below reported in 47.46% of visits. Secondly, mystery shoppers reported the extent to which the bank employee asked them questions to better understand their financial situation, using a scale from -3 (not at all) to +3 (very much). The average rating for this question was -0.43, with scores of zero or below reported in 58% of visits. These three outcomes – ethical behavior (EB), perceived ethicality (PE) and asking question (AQ) – correlate positively and significantly with each other (EB vs. PE=0.563$^{***,}$ EB vs. AQ=0.2960$^{***}$, PE vs. AQ=0.4468$^{***}$).

Based on the baseline results, the ethics program was then randomly assigned to N=26 managers and the respective N=92 bank shops under their management. Randomization was stratified by the bank shops' baseline levels of ethicality, calculated as the fraction of baseline visits in which employees provided the more ethical response.

Table 2.4 Randomization balance

| | Control | | Treatment | | P-val |
|---|---|---|---|---|---|
| | mean | s.d. | mean | s.d. | |
| Employee ethical response (1=y) | 0.369 | 0.166 | 0.396 | 0.199 | 0.473 |
| Perceived ethicality (0-6) | 0.44 | 1.878 | 0.337 | 1.868 | 0.547 |
| Employee asked questions (0-6) | -0.344 | 2.01 | -0.481 | 2.012 | 0.435 |
| Employee age (years) | 33.298 | 7.82 | 34 | 7.957 | 0.475 |
| Employee male (1=yes) | 0.383 | 0.487 | 0.267 | 0.443 | 0.063* |
| Nr of other clients in shop | 0.387 | 0.766 | 0.441 | 0.796 | 0.525 |
| Nr of other employees in shop | 1.28 | 0.883 | 1.289 | 0.844 | 0.905 |

As such, the baseline fractions of ethical responses in the treatment and control group are 39.6% and 36.9% respectively, the difference being not statistically significant (Mann-Whitney test, z=-0.757. Pr>|z|=0.4491). In addition, the randomization was also balanced in terms of baseline levels of the following

variables: (i) mystery shopper's perception of the ethicality of the employee, (ii) mystery shopper's perceptions of the extent to which the employee asked questions to understand their financial situation, (iii) employee demographics, and (iv) the number of other clients and employees in the shop during the mystery shopping visit.

## *Compliance*

At the start of the program, selected bank managers were informed via email by the headquarters that they had been selected to participate in a pilot ethics program. This email indicated that the program built on the bank's ambition to maintain the highest ethical standards and to restore the trust of society in the financial sector. Furthermore, the email informed the selected managers about the various activities they and their team were requested to carry out during the program: (i) implementation of a kickoff meeting, (ii) usage of the new client greeting and (iii) implementation of weekly "dilemma meetings". In the subsequent two weeks, the headquarters checked with all selected bank managers to see whether they had read the email and made the necessary preparations for the program kickoff. During this phase, the research team was informed that one manager was unwilling to implement the program. In the estimation of the treatment effect, this manager and corresponding shops were omitted from the analyses.

We will now report several indicators regarding the compliance with the program. Data about compliance were obtained in two ways. First, mystery shoppers recorded whether the bank employee greeted them with a slogan "similar" to the actual slogan employees were requested to use when starting a conversation with clients. Because the concept of similarity is inherently vague, we cannot refute the possibility that this question captures not only whether employees actually used the slogan, but also more generally whether they used a

client-friendly greeting. Second, an online survey was sent to all shop employees after the program had finished. This survey contained questions about participation in the kickoff meeting and weekly meetings.

Regarding the client greeting, we find that client-focused greetings were observed in 7.2% of visits during the baseline as opposed to 15.3% during the actual program phase. However, it appeared that client-focused greetings were recorded both in the treatment and control shops. For the control shops, client-focused greetings were recorded in 7.6% of baseline visits and 12.7% of program phase visits. For the treatment shops, such client-focused greetings were recorded in 6.7% of baseline visits and 18.3% of program phase visits.

These figures indicate that the slogan was not adopted in most of the treatment shops. This corresponds with anecdotal reports received from the bank during the implementation phase indicating that employees found the slogan to be "artificial", "bothersome" or simply "annoying" and that many simply refused to use it.

Furthermore, these results indicate that some employees – both in the treatment and control group – were already using client-focused greetings. Next, the increase for the control group could imply several things: (i) there may have been some spillovers from the treatment group, for example through employees rotating between shops, however such rotations were purportedly very infrequent according to the bank headquarters; (ii) mystery shoppers may have been more attuned to the greetings during the program phase, and thus more likely to interpret any type of client-focused greeting as sounding "similar" to the actual program greeting. Finally, the fact that usage of such greetings was recorded in both treatment and control shops does speak to the fact that shoppers were not aware of the treatment status of the shops.

We now turn to the results from the ex-post online survey, which was sent to the employees in all bank shops in June 2017. Participation was voluntary, but employees were strongly urged by the bank management to fill it out. In total

N=238 employees opened the survey. These responses corresponded to N=35 control shops (N=110 responses) and N=38 treatment shops (N=128 responses). After filling out several questions related to ethical behavior and social norms – which we report on in the next chapter – employees in treatment shops were asked to fill out several questions about the implementation of the program.

In total, N=63 employees from N=31 bank shops responded to these questions. These responding shops have slightly higher levels of baseline ethicality than shops where no employees responded (44% vs. 35%), so the survey results may be biased although it is unclear which direction such a bias would have. It is conceivable that more ethical shops are also more compliant, while at the same time employees in such shops may be less motivated to engage in the program – for example because they think they already behave ethically. In sum, it remains unclear if and to what extent these survey results reflect the actual compliance with the program. We now proceed to discussing the survey results.

As shown in table 2.5 below, approximately 76% of employees reported they attended the kickoff meeting (q1). Furthermore, employees reported they participated in an average of 3.5 weekly meetings (q2). Most employees participated in 1-5 meetings (60%), while some participated in 5-10 meetings (22%). Approximately 5% participated in more than 10 meetings, while around 12% did not attend any meetings. Next, 37% of weekly meetings were attended by colleagues from other shops in the manager's clusters ("sales region", q3). Employees shared, on average, 1-2 personal dilemmas during these meetings (q4) and reported they were moderately active in terms of speaking vs. listening (q5). Typically, these meetings were held with around 4 employees (q6).

In sum, these figures suggest that compliance with the meeting component was quite successful and that – given the limited compliance with the client-greeting – these meetings are the primary component of the overall ethics program.

Table 2.5 Compliance with program

| Question | mean | s.d. | N |
|---|---|---|---|
| 1. Did you attend the kickoff during which a video message from the director was shown? (1=yes,0=no) | 0.76 | 0.43 | 63 |
| 2. How many times did you attend a weekly meeting, during which you discussed ethical dilemmas? (0-20) | 3.57 | 3.44 | 63 |
| 3. Where there colleagues present from other shops in your sales region during the meetings? (1=yes,0=no) | 0.37 | 0.49 | 63 |
| 4. How many times did you share a personal dilemma during the meetings? (0-20) | 1.57 | 1.88 | 63 |
| 5. Did you mostly listen or mostly speak during the weekly meetings? (0=mostly listen, 6=mostly speak) | 3.87 | 1.24 | 63 |
| 6. How many colleagues (that is, excluding you) were typically present during the weekly meetings? (0-10) | 3.33 | 2.19 | 63 |

## *Average treatment effect*

We estimate the average treatment effect[14] using the following difference-in-difference model, with $Y_{ij}$ denoting the respective outcome variable for bank shop *i* at time *t;* $T_{it}$ denoting treatment status of bank shop *i* at time t*;* $X_{it}$ denoting a vector of control variables[15]; $\lambda_t$ denoting a time fixed effect; $B_i$ denoting a bank shop fixed effect and $M_i$ denoting a mystery shopper fixed effect:

$$Y_{ij} = \beta_0 + \beta_1 T_{it} + \beta_2 X_{it} + \lambda_t + B_i + M_i + \varepsilon_{it} \qquad (3.1)$$

The results in table 5 indicate that the ethics program did not have a statistically significant effect on employee advice, as indicated by the insignificant point estimate on the "ethics program" variable in column 1. The treatment did have a marginally significant positive effect on the perception of bank employee ethicality. However, the treatment did not influence the perception of the extent to

[14] Observations for six shops in the treatment group - where the manager did not implement the program – were excluded from the analysis

[15] The controls are: (i) bank employee age; (ii) bank employee sex, (iii) nr. of clients in bank shop during visit (iv) nr. of employees in bank shop during visit

which bank employees asked the client questions to better understand their financial situation. In sum, these analyses suggest that – on average – the ethics program had limited effects on employee behavior towards the client.

We find a similar null result when considering the treatment effects on the two other key outcome variables: (i) "perceived ethicality"; whether the mystery shopper perceived the bank employee to be placing the client's interest first (table 6 column 2) and (ii) "asking questions"; whether the mystery shopper perceived the bank employee to be asking to questions to better understand the client's financial situation (table 6, column 3).

Table 2.6 Treatment effects on ethical response

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Ethical response | | | |
| Ethics program | 0.010 | 0.024 | -0.023 | -0.066 |
| | (0.045) | (0.045) | (0.059) | (0.077) |
| Post | -0.011 | -0.015 | 0.008 | 0.094* |
| | (0.037) | (0.036) | (0.038) | (0.052) |
| Constant | 0.384*** | 0.371*** | 0.383*** | 0.315 |
| | (0.020) | (0.054) | (0.052) | (0.312) |
| Observations | 1,032 | 1,032 | 1,032 | 1,032 |
| R-squared | 0.000 | 0.008 | 0.006 | 0.310 |
| Controls | No | Yes | Yes | Yes |
| Shop FE | No | No | Yes | Yes |
| Shopper FE | No | No | No | Yes |
| Case FE | No | No | No | No |
| Nr of bank shops | 86 | 86 | 86 | 86 |

OLS regression with clustered standard errors in parentheses. Dependent variable is whether the bank employee gave the more ethical response (1=yes, 0=no). Visit controls include: employee age, employee sex, other clients in shop (yes/no), other employees in shop (yes/no). Post is a dummy variable indicating the period (0=before program, 1=during)*** p<0.01, ** p<0.05, * p<0.1

Table 2.7 Treatment effects on all outcome variables

| | (1) Ethical response (0=no 1=yes) | (2) Perceived ethicality (-3 to +3 scale) | (3) Asking questions (-3 to +3 scale) |
|---|---|---|---|
| Ethics program | -0.100 | -0.136 | 0.013 |
| | (0.077) | (0.271) | (0.309) |
| Post | 0.105** | 0.223 | 0.115 |
| | (0.049) | (0.202) | (0.240) |
| Constant | 0.351 | -2.382*** | 0.907 |
| | (0.276) | (0.850) | (1.309) |
| Observations | 1,032 | 1,032 | 1,032 |
| R-squared | 0.377 | 0.454 | 0.426 |
| Controls | Yes | Yes | Yes |
| Shop FE | Yes | Yes | Yes |
| Shopper FE | Yes | Yes | Yes |
| Case FE | Yes | Yes | Yes |
| Nr of bank shops | 86 | 86 | 86 |

OLS regression with clustered standard errors in parentheses. Visit controls include: employee age, employee sex, other clients in shop (yes/no), other employees in shop (yes/no). Post is a dummy variable indicating the period (0=before program, 1=during)*** p<0.01, ** p<0.05, * p<0.1

## *Heterogeneous treatment effects*

We then estimate the treatment effect separately for the bank shops with below- vs. above-median "baseline ethicality", again using model 3.1. To this end, we apply the same score that was used for the stratified randomization of the program: the fraction of baseline mystery shop visits in which the employee provided the more ethical advice. We split the sample into a group of shops with baseline ethicality scores equal to and above the median of 33% (BE>median) and a group with below-median baseline ethicality scores (BE<median). In this manner, there are N=64 shops with equal-to and above-median baseline ethicality and N=22 shops with below-median baseline ethicality. As a robustness check, we also split the sample into above vs. below and equal to median sub-groups (columns 3-4).

Table 2.8 Treatment effect, by baseline ethicality

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Ethics response(0=no, 1=yes) | | | |
| | EB<median | EB≥median | EB≤median | EB>median |
| Ethics program | 0.107 | -0.188** | -0.000 | -0.061 |
| | (0.157) | (0.086) | (0.099) | (0.154) |
| post | 0.232 | 0.082 | 0.156** | -0.104 |
| | (0.159) | (0.062) | (0.063) | (0.129) |
| Constant | 0.164 | 0.853*** | -0.156 | 1.429*** |
| | (0.354) | (0.174) | (0.215) | (0.293) |
| Observations | 264 | 768 | 600 | 432 |
| R-squared | 0.683 | 0.435 | 0.474 | 0.551 |
| Nr of bank shops | 22 | 64 | 50 | 36 |
| Controls | Yes | Yes | Yes | Yes |
| Shop, Shopper and Case FE | Yes | Yes | Yes | Yes |

OLS regression with clustered standard errors in parentheses. Dependent variable is whether the bank employee gave the more ethical response (1=yes, 0=no). EB≤median (EB>median ) indicates if the fraction of ethical responses during the baseline was equal/below (above) median. Visit controls include: employee age, employee sex, other clients in shop (yes/no), other employees in shop (yes/no). Post is a dummy indicating the period (0=baseline, 1=during) $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

When comparing the treatment effect on these sub-groups, we find that the program had a significantly negative effect on ethical behavior for the equal-to and above-median baseline ethicality shops, as indicated by the coefficient on the treatment variable in column 2. In contrast, the program has a positive but insignificant effect for the sub-group of N=22 shops with below-median baseline ethicality (column 1). However, when splitting the sample into (i) below- or equal-to median vs. (i) above-median baseline ethicality sub-groups (columns 3-4), the negative effect for the latter group is smaller and no longer statistically significant. This result suggests that the negative effect observed in column 2 cannot be ascribed simply to shop baseline ethicality. If, indeed, shops with higher levels of baseline ethicality would react negatively to the program, one would expect to see an even stronger negative effect for the sample of above-median baseline ethicality. However, as shown in column 4, the treatment effect is not significant

for this sub-group. Another possibility is that it's not the baseline average but rather the variance in ethicality that causes heterogeneous effects. The conjecture here is that the dilemma meetings unfolded differently in shops with few very unethical employees compared to shops with a more homogenous group. Although both groups may have the same baseline ethicality average, the former will have a greater variance. In the high variance shop, the more unethical employee may "drag down" the social norms and consequent ethical behavior of colleagues. Furthermore, this effect may be particularly strong in shops with equal- or below-median baseline ethicality, since these shops are more likely to have at least one or two "highly unethical" bank employees.

Since the main outcome measure "ethics score" is a dummy variable, we use the "perceived ethicality", which is rated on a -3 to +3 scale, to construct a more fine-grained measure of baseline shop-level variance in ethicality. Using model 3.1, we then estimate the treatment effect for the group of shops with baseline perceived ethicality below median (column 1) vs. baseline perceived ethicality above median (column 2). We also estimate this model separately for shops with high vs. low baseline ethicality (columns 3-6). In line with our conjecture, we find that indeed, the treatment effect is negative for high-variance shops and positive for low-variance shops. For the full sample, the treatment effect is negative and highly insignificant at the 1% level for shops with above-median baseline variance in perceived ethicality (column 2), while the effect is positive but statistically insignificant for the shops with below-median variance. This differential effect is particularly strong and significant at the 5% level for shops with equal- or below-median baseline ethicality (column 3-4). For high baseline ethicality shops (column 5-6) the direction of the effect is similar, but the size of the effect is smaller and statistically not significant. The result for the main sample (i.e. columns 1-2) is also shown in graph 1.

Table 2.9 Treatment effect, by baseline ethicality

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | | Ethics response | | | |
| | Full sample | | EB≤median | | EB>median | |
| | Var<m | Var>m | Var<m | Var>m | Var<m | Var>m |
| Ethics program | 0.092 | -0.328*** | 0.284** | -0.426** | 0.138 | -0.159 |
| | (0.109) | (0.110) | (0.124) | (0.192) | (0.305) | (0.172) |
| post | 0.014 | 0.141* | 0.023 | 0.272** | -0.329 | -0.077 |
| | (0.071) | (0.083) | (0.107) | (0.120) | (0.241) | (0.178) |
| Constant | 0.321 | 0.735*** | 0.084 | -0.325 | 0.201 | 1.161* |
| | (0.273) | (0.238) | (0.287) | (0.470) | (0.299) | (0.608) |
| Observations | 528 | 504 | 312 | 288 | 216 | 216 |
| R-squared | 0.493 | 0.510 | 0.653 | 0.583 | 0.751 | 0.749 |
| Nr of bank shops | 44 | 42 | 26 | 24 | 18 | 18 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Shop FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Shopper FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Case FE | Yes | Yes | Yes | Yes | Yes | Yes |

OLS regression with clustered standard errors in parentheses. Dependent variable is whether the bank employee gave the more ethical response (1=yes, 0=no). EB≤median (EB>median ) indicates whether the fraction of ethical responses during the baseline was equal/below (above) median. Var<m (var>m) if the variance in the perceived ethicality during the baseline was below (above) median. Visit controls include: employee age, employee sex, other clients in shop (yes/no), other employees in shop (yes/no). Post is a dummy variable indicating the period (0=baseline, 1=during) $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$

## *Survey*

To better understand the insignificant net effect of the ethics program, we turn to the results from the online survey. This survey was sent to all employees approximately one month after the ethics program. A total of N=198 employees from N=68 shops responded, divided equally between treatment (N=103 subjects, N=33 shops) and control (N=95 subjects, N=35 shops).

In this anonymous survey, employees were shown a list of statements and asked to indicate on a 1-7 scale the extent to which they agreed, with a value of 1 corresponding to "I do not agreeing at all" and 7 to "I fully agree". Statements 1-7 concerned issues relating to employees' awareness and knowledge of ethical

dilemmas with clients ("what *is*"), while statements 8-10 pertained to what employees think appropriate behaviors are ("what *should be*").

For statements 1-7, employees were asked to indicate the extent to which *they themselves* agreed, while for statements 8-9 they were asked the extent to which *the people in their team* agreed. The reason for this difference is that the bank management felt it was too suggestive to ask employees personal questions about ethical norms (i.e. statements 8-10). As a result, it remains somewhat unclear to what extent the responses to these latter statements reflect employees' personal ethical views vs. their perception of their colleagues' views.

Table 2.10 Survey responses, treatment effects

| Level of agreement with statement (1-7 scale) | | Control | | Treatment | | Dif. |
|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | |
| Q1 | "Sometimes there is tension between what is best for the client vs. the bank" | 3.87 | 1.74 | 4.17 | 1.71 | 0.29 |
| Q2 | "Not every financial product is good for all clients" | 5.42 | 1.57 | 5.22 | 1.80 | -0.20 |
| Q3 | "I know which questions to ask clients to understand what's best for them" | 6.22 | 0.80 | 6.29 | 0.82 | 0.07 |
| Q4 | "I know the details about all products that clients ask me about" | 5.60 | 1.21 | 5.50 | 1.36 | -0.10 |
| Q5 | "When I have doubts about how to best serve a client, I discuss with colleagues" | 6.47 | 0.85 | 6.21 | 1.07 | -0.25* |
| Q6 | "The bank gives me the freedom to serve clients in the way I see fit" | 5.73 | 1.14 | 5.26 | 1.30 | -0.46*** |
| Q7 | "The bank expects me to place the clients' interests above meeting my targets" | 5.22 | 1.60 | 4.80 | 1.71 | -0.43* |
| Q8 | "The clients' interests should be placed before the bank's interests" | 6.19 | 1.14 | 6.09 | 1.19 | -0.10 |
| Q9 | "Clients should know themselves if a financial product is suitable for them" | 3.37 | 1.57 | 3.13 | 1.45 | -0.24 |

The results for questions 1-2 indicate that the ethics program did not have a significant effect on employees' awareness of ethical dilemmas vis-à-vis clients. Second, the results for questions 3-5 indicate that in addition, the ethics program did not have a significant effect on employees' product knowledge and procedures for obtaining such knowledge. Surprisingly, there was even a marginally negative treatment effect on employees' propensity to discuss doubts with their colleagues.

Third, the results from question 6 indicate that the program undermined employees' sense of autonomy: this corresponds to employees' reluctance to implement the client greeting. In addition, this feeling of undermined autonomy might explain, at least partially, the ineffectiveness of the program. Furthermore, the response to question 7 indicates that the program had a moderately negative effect on employees' perceptions of the importance the bank expects them to place on meeting targets vs. behaving ethically vis-à-vis the client. This result may indicate that the group meetings made employees more aware of these potentially opposing objectives. Another interpretation is that peer influences in the group meetings changed employees' beliefs about the bank's expectations, for example if a less ethical colleague tells them that it's acceptable not to inform clients about certain things.

Fourth, the results from statements 8-9 indicate that the program did not have a significant effect on employees' views regarding how they *should* behave vis-à-vis clients; if anything, there appears to be a slight negative effect, although not statistically significant. These results correspond with the findings from the audit study: although the program had a positive effect on the sub-group of shops with below-median average and variance in baseline ethicality, this effect is cancelled out by negative effects for other sub-groups.

## 2.4 Discussion

### *Summary of main results*

We investigate how an ethics program influences the behavior of front-office employees in a large commercial bank. In particular, we considered employee behavior vis-à-vis clients in situations where it is in the employees' interest to pursue a product sale, whereas it is presumably better for the client not to do so. To this end, we implemented an audit study, whereby professionally trained actors visited bank employees who were unaware of their identity as a so-called "mystery shopper". The two-month program consisted of two parts. Firstly, employees were instructed to use a new client greeting, which served as a moral reminder. Compliance with this part of the program was very limited. Secondly, employees held weekly meetings where they discussed ethical dilemmas they encountered during their work. According to feedback from the bank and results from an ex-post survey, compliance with this part of the program was high, with 85% of employees in treatment shops reporting they participated in a least one of the weekly meetings. Three main results emerge from our study:

First, the overall level of employee ethicality - as measured using audit methodology - is low. Bank employees informed the client about potential problems with the product in less than 40% of the cases implemented during the baseline. This suggests that despite the various measures taken by governments and industry to promote integrity in the financial sector, unethical behavior among employees is still commonplace. Second, we find that the ethics program has no significant effects on employee behavior in terms of both objective and subjective measures of ethical behavior as collected in the audit study. Third, we find heterogeneity in the program effects. In particular, the program has a negative and significant effect for shops with higher variance in baseline ethicality, whereas the program effect is positive – albeit not statistically significant for shops with low

baseline variance. Finally, for shops with below-median average and variance of baseline ethicality, the program actually has a positive effect.

## *Interpretation*

Why do we find a negative treatment effect for shops with high variance in baseline ethicality? One possible explanation is related to the notion of conformism, where people adjust their behavior in the direction of the behavior they observe in others. Conformism is observed in various social behaviors, such as charitable giving (Alpizar et al., 2008), energy consumption (Schultz et al., 2007) and tax compliance (Bobek et al., 2007). Pierce and Snyder (2012) show that inspectors adjust the rate of passing vehicles – an indicator of ethical behavior – to the behavior of their co-workers (Pierce and Snyder, 2008). Similarly, Gino et al. (2009) show in a lab experiment that when people observe another person behaving unethically, they become more inclined to behave unethically themselves (Gino et al., 2009).

In the context of our study, high variance in baseline ethicality implies that within a shop, some employees behaved ethically while others behaved unethically. In these shops, the group discussions implemented during the ethics program may have actually allowed the unethical employees to exert a negative effect on the ethical norms of their colleagues. This, in turn, may have caused these employees to actually behave less ethically towards clients.

However, if this interpretation is correct, the question arises as to why the unethical employees would have "dragged down" the integrity of their colleagues during these meetings, rather than the ethical employees "pushing up" the integrity of their unethical peers? One possibility is that employees may seek motivations to justify self-serving tendencies and that the exposure to unethical peers provided a way for them to engage in "motivated reasoning" (Kunda, 1990). Given such

"excuse-driven" tendencies, employees may have selectively incorporated signals from unethical peers.

This interpretation may also account for the positive treatment effect in shops with below-median variance, particularly in the sub-group of shops with below-median levels of baseline ethicality. According to this view, the positive effects of the group meetings in these shops – such as making the client's interest more salient – were stronger than the negative conformity and contagion effects observed in the high variance shops.

## *Limitations*

Several limitations and caveats apply to the findings of this study. First, the mystery shopping study method is predicated on the research subjects not being aware that they are being studied. We did, however, receive anecdotal reports from the bank headquarters that some bank employees had voiced suspicions. In particular, it was reported that one bank employee that had rotated to a neighboring shop had encountered the same mystery shopper twice, once in her main shop and once in the other shop. We cannot refute the possibility that this suspicion was shared with other colleagues. However, there are two reasons why we consider that these suspicions do not threaten the validity of our main results. Firstly, bank employees have encountered mystery shopping in the past, which is typically used by banks and other companies to test whether shops adhere to sales and marketing protocols. As such, we deem it unlikely that suspicions were directly linked to the topic of ethical behavior. Secondly, if indeed there had been widespread suspicions about such an audit, we would have expected to observe a strong increase in ethical behavior during the last weeks of the program. However, no such trend was observed.

Furthermore, it should be noted that compliance with the program was imperfect. In particular, the new client greeting was only used by a very small

fraction of bank employees. Anecdotal feedback from the bank indicated this was because bank employees did not like being instructed in such a strict way on how to communicate with their clients, and that it undermined their sense of autonomy.

## *Policy relevance and future research*

This study has a number of policy implications. Firstly, it indicates that there is a considerable degree of "misbehavior" in the financial sector, despite the implementation of various measures by government and industry. Although this study considered behavior of employees in one particular bank, and results cannot simply be taken as reflective of the industry at large, the finding that there is still considerable misbehavior does align with findings from several other recent studies (Egan et al., 2016; Mullainathan et al., 2012)

Second, the insignificant average treatment effect suggests that ethics programs and other "soft measures" may not be a panacea for all misbehavior in the financial sector and other industries suffering from unethical behavior. In the absence of revisions of incentive schemes that can promote behavior that is harmful to clients, as illustrated by the recent Wells Fargo scandal, it remains questionable as to whether ethics programs will fully solve the problem of misbehavior in the financial sector.

Third, the study suggests that the details matter when it comes to the manner in which ethics programs are implemented. For example, if employees have the feeling that the ethics program undermines their autonomy, or that it provides a signal about low levels of ethical behavior in the organization, these programs may actually have an adverse effect. This result aligns with the finding that ethical codes of conduct that are not enforced by the organization can be perceived by employees as window dressing, and cause a cynical response that leads to more unethical behavior (Kish-Gephart et al., 2010). Interventions that require less active involvement from employees, such as reminders sent per email

or text message (Konrath et al., 2015), might circumvent the negative sentiments and low compliance associated with programs that require active involvement of employees.

Finally, our finding that the program did cause an increase in ethical behavior among employees in shops with low levels and variance in baseline ethicality suggests that such program do have potential to promote ethical behavior. And that this applies even in the presence of countervailing incentives and with imperfect program implementation. Given the possible "ethical contagion" in interventions that are implemented within groups, an interesting avenue for future research may be to test interventions that individual employees, such as daily "ethics reminders" send to employees via email or text message.

# Appendix 2.1 Mystery Shopping Instrument

| Question | Answer options |
|---|---|
| Did the employee inform you about the potential downside of the product? | yes/no |
| To what extent do you think the employee acted in your interest? | 1-7 scale |
| To what extent did the employee ask questions to better understand your financial situation? | 1-7 scale |
| How capable do you think the employee was? | 1-7 scale |
| How many other clients were present when you left the shop? | Open |
| How many bank employees, including the one you spoke with, were present in the bank shop during your visit? | Open |
| How busy do you think the employee was? | 1-7 scale |
| Could other employees or clients hear the conversation? | yes/no |
| What is the sex of the employee you spoke with? | male/female |
| How old do you think the employee was? | <20 years, 21-30, 31-40, 41-50, >50 |
| Did the employee say at the beginning of the conversation that they "placed your interest first" (or something similar)? | yes/no |

# Appendix 2.2 Online Survey Instrument

1. In which region/location do you usually work? (dropdown menu)
2. How many others work in the shop where you typically work? (0-20)
3. To what extent do you agree with the following statements (0=not at all, 6 =fully)
   a) Sometimes there is tension between what is best for the client vs. the bank
   b) Not every financial product is good for all clients
   c) Pressure at work sometimes makes it difficult to focus on the clients' best interest
   d) I know which questions to ask clients to understand what's best for them
   e) I know the details of all products that clients ask me about
   f) When I have doubts about how to best serve a client, I discuss with others in the office
   g) The bank gives me the freedom to serve clients in the way I see fit
   h) The bank expects me to place the client's interest above meeting my targets
4. To what extent do the people in your team agree with the following statements (0=not at all, 6 =fully)
   a) To advance one's career in the bank one should be flexible on ethical standards
   b) The client's interest should be placed before the bank's interest
   c) Clients should know themselves if a financial product is suitable for them
5. How ethically appropriate/inappropriate do the people in your team consider a bank employee acting as described in the scenario? (0=very unacceptable, 6=very acceptable)
   a) A client asks a bank employee several questions about travel insurance. While doing so, this client mentions that he is a member of the Automobile Association. Next, the bank employee recommends SNS travel insurance to the client without first asking if the client already has travel insurance via the Automobile Association.
6. Finally, we would like to ask you a few questions about financial products:
   a) A client deposits € 10.000 in a SNS savings account which offers a 0.4% interest rate. How much interest will they have earned after 5 years? (A=<€ 200, B=€ 200, C=>€ 200)
   b) A client has a travel insurance policy with SNS. During their holiday in France they accidentally break a vase in a museum. Are these damages covered by the insurance? (yes/no)
   c) A client's income in 2020 will be double that of today, and the price of all products and services will have doubled as well. How much can they buy with their income in 2020? (A=more than today, B=the same, C=less than today)

Recently, several shops participated in the pilot "From Oath to Client"

7. Did your shop participate? (yes/no)

*If yes:*

8. Did you attend the kickoff during which a video message from the director was shown? (yes/no)

9. How frequently did you use the new client greeting/slogan? (0=never, 6=very often)
10. How frequently did your colleagues use the new client greeting/slogan? (0=never, 6=very often)
11. How many times did you attend a weekly meeting, during which you discussed ethical dilemmas? (0-20)
12. Were there colleagues present from other shops in your sales region during the weekly meetings? (yes/no)
13. How many times did you share a personal dilemma during the weekly meetings? (0-20)
14. Did you mostly listen or mostly speak during the weekly meetings? (0=mostly listen, 6=mostly speak)
15. How many colleagues (that is, excluding you) were typically present during the weekly meetings? (0-10)
16. How time-consuming was participation in the program? (0 = not at all 6 = very much)
17. Would you recommend the program to other shops or banks? (0 = not at all 6 = very much)
18. What is the most important thing you learned from this pilot? (Open)
19. If you could change one thing about this pilot, what would it be? (Open)

# Appendix 2.3 Case Pilot Test mTurk

A total of N=1198 subjects were recruited via Amazon mTurk to indicate for each of the six case scripts how ethically appropriate or inappropriate they found the described behavior of the bank employee. A scale ranging from -3 (very inappropriate) to +3 (very appropriate) was used for this purpose. As can be seen in table A1, the average ethicality rating was below zero in all six cases, and a majority of subjects rated the described response as ethically inappropriate.

Table 2.11 Ethicality ratings, per case

| Case | Average rating | Fraction rated as unethical | Obs. |
|---|---|---|---|
| 1 | -0.588 | 0.552 | 1198 |
| 2 | -1.989 | 0.908 | 1198 |
| 3 | -2.428 | 0.945 | 1198 |
| 4 | -0.810 | 0.593 | 1198 |
| 5 | -0.552 | 0.532 | 1198 |
| 6 | -1.763 | 0.848 | 1198 |

Chapter 3

# Professional Expertise and the Reference Group Bias*

"Experto credito" Virgil, 19 B.C.

(*Trust me, for I have experienced*)

## Abstract

The amount of charitable funds allocated to disasters, diseases and other problems is disproportional to the number of people that are affected. One proposed explanation for this is the reference group bias, which is the tendency to place more weight on the fraction of people than on the absolute number of people that can be helped. Evidence for bias is derived from experiments with students. However, it is unclear whether professionals in charitable organizations are similarly prone to this bias. In two studies with students and professionals we find evidence for the reference group bias among both students and professionals. Furthermore, our second study shows that susceptibility to the bias, particularly for non-experts, is reduced when projects are evaluated jointly rather than separately.

* Joint work with Dinand Webbink.

## 3.1 Introduction

A large amount of charitable funds is allocated to problems that afflict relatively few people. For example, over $3 billion in charitable funds were collected in the U.S in the wake of the 2005 Katrina hurricane, which directly affected approximately 2 million people. In comparison, during the same year less than $1 billion was allocated to fight malaria, a disease which annually affects over 300 million people (Epstein, 2006). One proposed explanation for this disproportional allocation of funds is that people prefer helping a larger fraction of victims than a larger absolute number of victims (a "larger drop in a larger ocean"). It is further proposed that people are 'psychologically numbed' when faced with large numbers of victims and that this numbing has resulted in a failure to motivate action in the face of mass murders and genocide (Slovic, 2007).

This tendency to discount the value of lives of people that are part of large groups of victims is known as the 'reference group effect' or 'proportion dominance effect' and has been demonstrated in a number of lab experiments with student populations (Baron, 1997; Fetherstonhaugh et al., 1997; Friedrich et al., 1999; Jenni and Loewenstein, 1997; Slovic et al., 2007).

In practice however, a large fraction of charitable funds are allocated to specific projects by professionals in foundations, governments and development agencies. As various studies show that results from behavioral experiments with student populations do not correspond to results obtained with other non-student populations (Falk et al., 2013; Gaudecker et al., 2012; Peterson, 2001; Stoop et al., 2012) this raises the question whether these professionals are similarly prone to the reference group bias.

In this study we test whether professionals from the charitable sector are similarly prone to the reference group bias as student populations in previous studies. To this aim, we conduct two studies with both students and professionals in the charitable sector. Subjects were presented with a set of hypothetical cases

about funding proposals for charitable projects that addressed various life-threatening problems. This is the same approach as in previous studies on proportion dominance and the reference group bias (Bartels, 2006; Fetherstonhaugh et al., 1997; Jenni and Loewenstein, 1997; Small et al., 2007). The cases varied in terms of both the absolute number of lives that could be saved and the size of the reference group. We find that both studies that experts and non-experts display the reference group bias, although experts are somewhat less biased in the second study.

Furthermore, we study whether enabling subjects to evaluate projects in a comparative rather than sequential manner can alleviate the reference group bias. This part of our study builds on previous research on preference reversals between separate and joint evaluation (Hsee and Zhang, 2010; Ritov and Baron, 2011), which indicates that attributes which are hard to evaluate in isolation, such as absolute numbers, have greater decision weight in joint evaluation because this makes it easier to evaluate these features. In line with this research we expected that in joint evaluation modus subjects would be less susceptible to the reference group bias. Furthermore, we were curious whether the effect of joint evaluation differed between experts and non-experts. If experts would feel more compelled to maximize life savings, their susceptibility might be more strongly reduced by the joint evaluation modus than among non-experts. Our results indicate that indeed the reference group bias was reduced in joint evaluation, although this effect was not significantly greater for experts. A possible explanation for this is that already in the separate evaluation modus the experts were less susceptible to the reference group bias.

As such, this study contributes to the literature on expert decision making and is, to the best of our knowledge, the first study to test the reference group bias among professionals. Furthermore, this study indicates that the evaluation framework can alleviate in part the susceptibility to the bias. Our findings suggest that the disproportionate allocation of charitable funds to identifiable victims is

partially caused by the proportion dominance bias in expert fund allocation. The paper proceeds as follows. In chapter 2 we provide an overview of the literature on the reference group bias and expert decision biases. In chapters 3 and 4 we describe the methodology, estimation strategy and results for our studies. Finally, in chapter 5 we conclude by discussing the results.

## 3.2 Background

### *Reference group bias*

People typically prefer to help specified victims, such as "two young children from Haiti named Maria and Jose whose parents were recently killed due to a flooding", rather than statistical victims, such as an "as-of-yet unknown children somewhere in the tropics whose parents may soon die from a curable disease such as HIV-AIDS". This preference is attributed to multiple factors (Jenni and Loewenstein, 1997). Firstly, identifiable victims are more vivid: people can observe them and thus their empathy can be invoked more, whereas they cannot observe a statistical victim. However, identifiability of victims does not always lead to more willingness to allocate vaccines to hypothetical victims (Wiss et al., 2015). Another reasons why people tend to prefer to help specific rather than statistical victims is that the latter are by definition uncertain – although it is likely that many children will soon be orphaned due to HIV-AIDS it is not absolutely certain. Thirdly and relatedly, identifiable victims are considered after they have become victims, whereas statistical victims are considered as such before they are actually victims. Finally, identifiable victims are often framed as stand-alone cases such as in the "Maria from Haiti" example, whereas statistical victims are by definition part of a larger problem. As such, people thus typically prefer saving 1-out-of-1 identifiable victims over saving 2-out-many statistical victims. Of these various dimensions along which identifiable and statistical victims differ, the

reference group dimension has been shown to have the greatest effect on people's willingness to save lives (Jenni and Loewenstein, 1997).

This 'reference group' bias was replicated in various subsequent experiments. For example, Baron (1997) found that people's willingness to pay for hypothetical cases of treatment of sick people was lower for cases where the number of people that could not be cured was higher, even when the number that could be cured was kept constant. In a second experiment, willingness to pay for reducing deaths for various causes was insensitive to whether the live savings were as framed as percentages or absolute quantities (Baron, 1997).

Comparable results were found in several other experiments. For example, Fetherstonhaugh (1997) measured preference over pairs of hypothetical government programs that aimed at providing drinking water to save the lives of 4500 Rwandan refugees suffering from cholera. The programs only differed in the total size of the refugee camp (11,000 vs. 250,000 lives) and subjects preferred programs for the smaller camps (Fetherstonhaugh et al., 1997). Friedrich et al. (1999) also replicated the reference group bias by an experiment with hypothetical cases about life-saving antilock brake requirement in which support for this requirement decreasing in the size of the reference group (Friedrich et al., 1999). Next, Slovic et al. (2007) asked people to state their support for purchasing of equipment to save lives in the case of airplane crashes. One group was informed that this could save 150 lives in such an event. Other groups were informed that it could respectively save 98%, 95%, 90% & 85% of the 150 lives. The willingness to contribute was higher for the latter groups, even though in those groups the absolute number of lives saved was lower (Slovic, 2007).

The reference group bias is a manifestation of a more general psychological tendency for people to place more decision weight on attributes that can be more easily evaluated (Hsee and Zhang, 2010). People typically find it easier to evaluate proportions than absolute quantities, especially in a context where they lack prior knowledge,. For example, continental Europeans will find it

difficult to assign a prize to 0.043 gallons of beer (an absolute quantity). However, when asked about a half-full regular-sized glass of beer (a proportion) it will be much easier to make this evaluation. As such, proportions are typically given more weight in decisions than absolute numbers.

However, when multiple options are evaluated simultaneously, absolute quantities can be compared and thus became easier to evaluate. As a result, preference reversals can occur when people shift from separate to joint evaluation. To demonstrate this, Fetherstonhaugh (1997) asked subjects to rank treatments for diseases which saved different absolute and relative numbers of people. While subjects gave higher rankings to projects that saved higher proportions of lives under separate evaluation, in the joint evaluation mode this preference switched to treatments that saved more lives in absolute terms (Fetherstonhaugh et al., 1997).

However, all lab-experiments about proportion dominance and the reference group bias that have hitherto been conducted made use of student populations. Bartels (2006) notes that undergraduates are usually not used to making policy choices about life-and-death cases and suggests that lack of expertise might contribute to their susceptibility to the reference group bias (Bartels, 2006). This thus raises the question if professionals in charitable organizations are similarly susceptible to the reference group bias when allocating funds to projects.

## *Expert decision making*

In certain domains experts are less susceptible to cognitive bias. For example, Marteau (1989) shows that medical students, when deciding between treatments, are less susceptible to a framing bias when they were asked to take the perspective of a patient than in the doctor's perspective Furthermore, susceptibility to cognitive biases also differs between groups of professionals and depend on

various factors, including whether the decision task can be decomposed in smaller tasks (Shanteau, 1992)..

In tasks that involve social comparisons, expertise does not always lead to better judgment and decision-making. In a seminal study (Kruger and Dunning, 1999) show that unskilled individuals tend to overestimate their own ability whereas skilled individuals are more likely to underestimate their ability. In a similar vein, expertise has been shown to affect how well people are able to estimate the level of knowledge of others. (Hinds, 1999) shows experts (salespeople) score worse than intermediate users (customers) in terms of predicting the performance of novices in a complex task involving advanced cellphone technology. Bromme et al., (2001) however show that IT professionals are better in predicting the level of knowledge about computer related topics among students than are laypersons. In the domain of finance, (Kaustia and Knüpfer, 2008a) find a negative relation between expertise and financial returns, as investors overweigh personal experience compared to rational Bayesian learning. . In the case of charitable donations, this could mean that professionals are sensitive to cognitive biases when allocating funds on behalf of donors than they would be when allocating their own money.

There are also a number of studies that show that cognitive biases are insensitive to professional experience. Kahneman and Klein (2009) review the literate on expert judgment and highlight various studies in which people who have extensive expertise in their respective fields incorrectly apply various heuristics, resulting in errors and biases in their decision making. They show that faulty statistical intuitions manifest even in experienced statisticians, in an experiment on choosing the number of cases for a psychological experiment. Similarly, Camerer and Johnson (1991) review the literature on expert decision making in clinical and medical domains and conclude the following: "the depressing conclusion from these studies is that expert judgments in these domains are no more accurate than those of lightly trained novices". Northcraft and Neale

(1987) show that real-estate agents' estimates of the value of various properties was biased towards the starting-list prices, an effect known as the anchoring bias. Another bias to which both students and professionals are susceptible is the confirmation bias; the tendency to selectively search for- and attach more weight to- pieces of data supporting pre-existing beliefs. This bias was displayed in lab experiments by students (Jonas et al., 2001; Jones and Sugden, 2001; Nickerson, 1998) as well as by employees of the World Bank (2015). Furthermore, World Bank employees were also susceptible to the sunk cost bias – the tendency to base decisions about future actions based on costs incurred in the past that cannot be changed and undone anymore (ibid).

Given this mixed evidence it is thus unclear ex-ante whether professionals in charitable organizations are less likely to display the reference group bias in allocating funds than novices.

## 3.3 Study 1

### *Participants*

The first session was conducted during an annual conference of the Dutch charitable sector in Rotterdam, the Netherlands in September 2011. Most participants at this conference were professionals from charitable institutions in the Netherlands. During the conference they could sign up to take part in the experiment; 89 individuals participated in the experiment. The second session was conducted in 2012 with undergraduate students of the Erasmus School of Economics. In this experiment 46 students participated. The third experiment was conducted in 2013, through an online survey with participants of the 2013 annual conference of the Dutch charitable sector in Rotterdam. In this session 60 individuals participated. The first two experiments were conducted in a lab setting; the first experiment was conducted in an auditorium during a conference about

philanthropy at Erasmus University, the second experiment was conducted in a classroom with economics students. In both these sessions a hardcopy version of the survey cases was used.

Background characteristics on gender, age and occupational sector were collected in a short questionnaire. Moreover, individuals were asked about the expertise in allocating funds to social projects, see appendix 2. Table 2 shows summary statistics for the three experiments and for the treatment groups within each experiment. The second experiment was done among students at the Erasmus School of Economics which explains the age difference with the other two experiments. More than half of the participants in the first and third experiment reported to have professional experience in allocating funds to social projects: these subjects were classified as 'experts'. The second experiment focused on a sample of students and therefore, does not contain experts.

We tested whether there were differences in the covariates between the treatment groups within each session. The p-values in the Table 2 show that there were no statistically significant differences in gender, age and expertise between the sessions, implying that the random assigned to the two treatment groups was successful. Information on expertise was missing for 14 observations, which were discarded in the regression analyses. For observations for which age and gender information was missing (10 and 9 observations respectively) we imputed the sample averages, and we included a dummy for these imputed values in the analyses.

Table 3.1 Summary statistics per session

| | | *Session 1* | | *Session 2* | | *Session 3* | | *Full sample* | |
|---|---|---|---|---|---|---|---|---|---|
| | Group | Mean | P-val. | Mean | P-val. | Mean | P-val. | Mean | P-val. |
| Male (%) | 1 | 0.47 | 0.45 | 0.67 | 0.92 | 0.58 | 0.11 | 0.55 | 0.18 |
| | 2 | 0.39 | | 0.68 | | 0.36 | | 0.45 | |
| Age (years) | 1 | 45.04 | 0.54 | 21.79 | 0.26 | 48.08 | 0.21 | 40.91 | 0.16 |
| | 2 | 43.27 | | 21.00 | | 43.76 | | 37.83 | |
| Expertise (%) | 1 | 0.67 | 0.25 | 0.00 | 0.00 | 0.58 | 0.36 | 0.49 | 0.16 |
| | 2 | 0.55 | | 0.00 | | 0.45 | | 0.39 | |
| Nr. of obs. | Total | 89 | | 46 | | 60 | | 195 | |

## *Materials and procedures*

To investigate the relationship between professional expertise and the susceptibility to the proportion dominance bias we conducted three experiments where we measured subjects' willingness to contribute (WTC) to various hypothetical health projects targeting victims and potential victims of various diseases in several countries in Sub-Sahara Africa. A detailed description of these cases is provided in the next section.

As is common in studies on the reference group bias we made use of hypothetical cases. Subjects were asked to imagine that they were in charge of evaluating proposals that required a subsidy of €50,000 from a large philanthropic institution. They were informed that the philanthropic institution had enough funds to grant each proposal but that a critical evaluation of each individual proposal was requested. Next, they were informed that they would be shown a number of proposals in sequential manner and asked to state for each of these their willingness to grant the subsidy on a 10 point scale. Subsequently, participants were given eight consecutive proposals for funding for public health programs in Sub-Saharan Africa. The proposals varied on the following dimensions: type of

disease, region, number of lives that could be saved, and size of reference group. Subjects were randomly assigned to one of two groups. Subjects in both groups were presented eight cases with between group variation in the number of lives to be saved and the size of the reference group. This randomization allows for differentiating between the effects of the case specifics (country and disease) and the effects of the absolute and relative number of lives on subjects' willingness to support the projects. An overview of the cases is shown in table 1, for wording of the vignettes please see appendix 3.3.

Table 3.2 Characteristics of cases for subsidy applications

| Case | | | Group 1 | | Group 2 | |
|---|---|---|---|---|---|---|
| Nr. | Disease | Country | Lives | Reference group | Lives | Reference group |
| 1 | Shigellosis | Uganda | 500 | 1.000 | 500 | 500.000 |
| 2 | Typhoid | Tanzania | 750 | 50.000 | 500 | 1.000 |
| 3 | Noma | Zambia | 500 | 1.000 | 750 | 20.000 |
| 4 | Cholera | Cameroun | 500 | 150.000 | 500 | 1.000 |
| 5 | Kala-azar | Nigeria | 750 | 500.000 | 500 | 1.000 |
| 6 | Trypano | Ethiopia | 500 | 1.000 | 750 | 150.000 |
| 7 | Dengue | Sudan | 500 | 1.000 | 500 | 50.000 |
| 8 | Lassa | Mali | 500 | 20.000 | 500 | 1.000 |

Subsequently, the proposals were handed out in hard-copy for the first two experiments, and online for the third experiment. In the first two experiments, participants were asked to evaluate the proposals in the order in which these were presented. It was stressed that each case was to be assessed independently. Participants were asked to imagine that the hypothetical budget, although sufficient to fund each proposal, was finite and that they hence should assess each case critically and specify their WTC for each case accordingly.

## *Empirical strategy*

To identify the effect of the reference group we first estimate the following model:

$$Y_{ij} = \beta_0 + \beta_1 L_{ij} + \beta_2 R_{ij} + \varepsilon \qquad (3.1)$$

in which $Y_{ij,}$ represents the willingness of individual i to contribute to case j (WTC), L is the absolute number of lives that can be saved, and R is the (natural logarithm) of the reference group size, $\varepsilon$ are unobserved factors and $\beta_{0,}\ \beta_{1,}\ \beta_2$ are parameters to be estimated[16]. The parameter $\beta_2$ represents the reference group bias. Estimates of the parameter $\beta_2$ can be interpreted as the causal effect of the size of the reference group because the cases, which have different combinations of L and R, are randomly assigned to individuals. Hence, it is unlikely that unobserved factors will be correlated with R.

Next, we study whether there are difference in the allocation decisions of experts and non-experts. This is done by adding a dummy variable E for being an expert and by interacting expertise (E) with the absolute number of lives saved (L) and with the size of the reference group (R):

$$Y_{ij} = \beta_0 + \beta_1 L_{ij} + \beta_2 R_{ij} + \beta_3 E_i + \beta_4 E_i * L_{ij} + \beta_5 E_i * R_{ij} + \varepsilon_{ij} \qquad (3.2)$$

As a robustness test, we estimate the model in equation (3.2) by adding additional controls for age, gender, case specifics (country and type of disease), and the experimental session. Furthermore, we test the robustness of the results for excluding the last case as we observed in our data that individuals had a higher willingness to contribute to the last case. As a final robustness test we include individual fixed effects, denoted by $\alpha_i$, as specified in equation:

[16] We use the natural logarithm because the reference group size increases non-linearly (1, 20, 50, 150, 500 thousand).

$$Y_{ij} = \beta_0 + \beta_1 L_{ij} + \beta_2 R_{ij} + \beta_3 E_i * L_{ij} + \beta_4 E_i * R_{ij} + \alpha_i + \varepsilon_i \qquad (3.3)$$

We start by estimating these models for the pooled data of all three experiments, and subsequently we estimate equation (3.3) separately for each of the three experiments.

## *Results*

Figure 3.1 shows the average willingness to contribute to the projects (WTC) according to the number of lives that can be saved (500 or 750), for both experts and non-experts. We observe that non-experts are more likely to contribute to projects than experts. Remarkably, the willingness to contribute seems to be independent of the number of lives that can be saved. This holds both for experts as for non-experts; the differences are statistically insignificant.

Figure 3-1 WTC and absolute number of lives

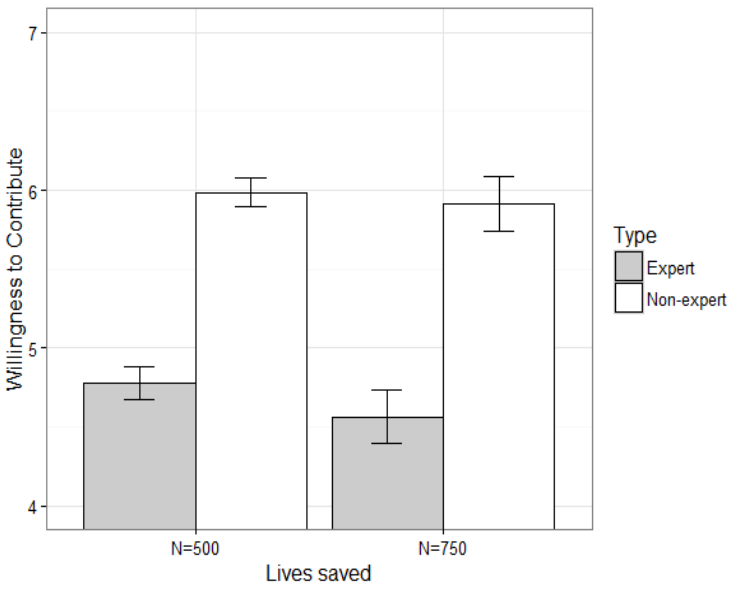
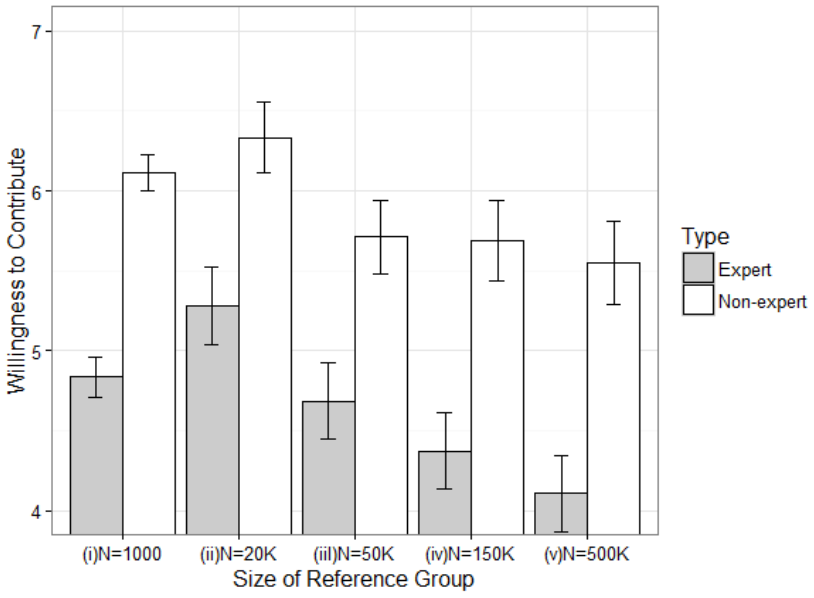


Figure 3-2 shows the relationship between the willingness to contribute (WTC) and the size of the reference group, both for experts as for non-experts. For both groups we observe that the WTC decreases with the size of the reference-group. This suggests that both experts and non-experts suffer from reference-group bias.

Figure 3-2 WTC and reference group size




In all specifications of the regression model in equation (3.1) we observe that willingness to contribute decreases in the size of the logarithm of the reference group ($P<0.01$)[17]. The results from the model specification in equation (3.2) show that, although willingness to contribute is lower for experts, there are no differences between experts and non-experts terms of in susceptibility to the reference group effect, as indicated by the non-significant coefficients on the "reference group*expertise" interaction variable ($p>0.1$). These findings are robust

[17] Detailed results of the regression analysis described in this section are provided in appendix 3.1.

to adding controls and to excluding observations from the eighth case as outlined in equation (3.3).

As a next step in the analysis we investigate the effects within each session. In all three experiments we find a negative effect of the size of the reference group on the willingness to contribute. Statistically significant effects are found for the second experiment ($p<0.1$) and third experiment ($p<0.01$). Again, we find no difference between the decisions of experts and non-experts. Hence, the results for the separate experiments are consistent with the results for the pooled sample.

At the end of the survey, after individuals had decided on the allocation of funds over the cases, they were asked to rank the importance of the following factors in their evaluation of the cases: (i) absolute number of lives saved, (ii) relative number of lives saved, (iii) other contextual factors (disease, region and country). Table 3 shows the answers of experts and non-experts.

Table 3.3 Self-reported decision criterion

| Main criterion[18] | Experts | Non-experts |
|---|---|---|
| Lives, absolute | 27.9% | 32.2% |
| Lives, relative | 24.6% | 31.2% |
| Other[19] | 47.5% | 36.6% |
| *Total* | *100%* | *100%* |

Almost half of the experts reported contextual factors (country, region, disease) as the main criteria for their allocation decisions, whereas non-experts more frequently listed the absolute number of lives. This difference is statistically significant ($t(194)=-1.823$, $P=0.068$). Furthermore, non-experts were more likely

[18] This table includes the observations for which the survey question on decision criteria (see appendix 1) was fully filled out (159 out of 195 observations) and for which the three respective classes of criteria were ranked: observations for which some or all criteria were equally ranked (154 out of the 159).

[19] Other factors are: (i) country, (ii) region, (iii) disease

than experts to list (i) the absolute number of lives and (ii) the relative number of lives as the main decision criteria, but these differences are not significant at the 10 %-level.

Next, we explored whether this self-reported ranking of decision criteria corresponds to actual choices, and whether expertise affects this correspondence. To this purpose we estimated Equation (3.3) according to the reported decision criteria. We find that the ranking does correspond with actual choices: for individuals that listed "absolute number of lives" as the main criterion WTC increases in absolute number of lives ($P<0.01$) but does not increase significantly in the size of the reference group, For subjects that ranked "relative number of lives" as main decision criterion the results are the opposite: WTC decreases in the size of the reference group ($P<0.01$) but is not affected by the absolute lives. Finally, for subjects that ranked "context" as the main criterion the coefficients on absolute or relative number of lives are statistically not significant. Again we find that the estimated effects on the interaction terms "expertise*lives saved" and "expertise*reference group" are statistically insignificant in all regressions.

## 3.4 Study 2

In order to calculate the cost-effectiveness of projects, for example in terms of lives saved, it is essential for donors to have an indication of the total costs. However, in reality the costs of projects are not always clear ex-ante. This was reflected in the vignettes used in the first study, in which estimated project costs were only mentioned at the beginning of the instructions. A concern with this approach is that professionals – based on their working experience – made different inferences and estimates of the project costs, which in turn could have influenced their reported "willingness to contribute" to these project, independent of the absolute and relative number of lives that could be saved. To address this

issue we conducted a second study in which we explicitly stated the costs for each project that subjects were asked to evaluate.

Furthermore, our first study asked subjects to evaluate individual projects, whereas in reality donors as well as professionals often evaluate several projects simultaneously, and make their evaluation in a comparative manner. Research shows that shifting from a separate- to joint evaluation modus can cause preference reversals, whereby the attention of the decision maker is shifted from choice attributes that are easier to assess in the separate evaluation to attributes that become easier to assess in the joint evaluation modus (Hsee and Zhang, 2010). This so-called evaluability bias has been shown to occur in various contexts. For example, (Ritov and Baron, 2011) find that a shift from separate to joint evaluation of public policies moves the focus from emotions that are aroused by the projects towards the more cognitive rating of the priority of the policy. In a similar vein, (Caviola et al., 2014) find that in the context of charitable donations, moving from separate to joint evaluation shifted focus from overhead expenses towards cost-effectiveness.

These results suggest that the reference group bias might be mitigated by joint evaluation of projects. The rationale here is that in the separate evaluation modus people focus more on the size of the reference group, whereas in the joint evaluation condition the number of lives that can be saved becomes an easier-to-evaluate attribute, which in turn would shift preferences towards projects with greater cost-effectiveness. This further raises the questions whether the magnitude of such a preference reversal differs between novices and experts. For example, it could be that experts are more trained to compare projects in terms of cost-effectiveness, in which case the comparison modus would have a greater effect on their evaluations. To explore this question, we include in our second study first a series of projects that are evaluated separately, followed by second part where these projects are presented in pairs. Through this within-subject design, we study whether joint evaluation reduces susceptibility to the reference group bias and

whether this possible preference reversal is more pronounced among professionals.

## *Participants*

The second study was conducted among a group of students at Erasmus University and among a group of professionals with experience in the allocation of funds towards social projects. Both groups were invited per email, students via the University mailing list for economics experiments and professionals via an emailing list of the Erasmus Centre of Strategic Philanthropy. The experiment was conducted through an online survey and each subject could participate in this survey once. The responses were collected in the period Sep. 2016 – April 2017. A total of 121 students and 39 professionals participated in this online survey experiment. Background characteristics on gender, age and occupational sector were collected in a short questionnaire. Moreover, individuals were asked about their experience in allocating funds to social projects, see annex V. Table 3.4 shows summary statistics for the experiments grouped by subject type. As expected, the average age of professionals was significantly higher. We control for this in our regression analyses.

Table 3.4 Subject characteristics, study 2

| | | *Session 1* | |
|---|---|---|---|
| | Group | Mean | P-val. |
| Male (%) | Non-experts | 50.1 | 0.28 |
| | Experts | 48.7 | |
| Age (years) | Non-experts | 25.5 | 0.00 |
| | Experts | 42.2 | |

## *Materials and procedures*

Subjects received a short email inviting them to participate in a short study on decision-making. This email contained the link to the online survey. Furthermore, to promote the response rate, subjects in both groups were informed that among the first 100 participants a total 10 gift vouchers with value of €25 would be randomly assigned. After clicking on the survey link, subjects were guided to the online survey which started with instructions in which subjects were asked to imagine that they were working for an organization that provides support to refugees in developing countries. Our vignette survey was based on the instrument used by (Fetherstonhaugh et al., 1997). Subjects were instructed to evaluate proposals from organizations that try to provide clean drinking water to sick people in refugee camps in South-Sudan. Subjects were instructed that the delivery of clean drinking water would save these patients' from near-certain death otherwise. It was explicitly noted to subjects that all projects costs approximately €50.000 (for a detailed overview see appendix XI). The vignettes indicated for each project the respective location of the refugee camp, the number of lives threatened by the disease, the percentage of lives that could be saved without the project, and the percentage of lives that could be saved with the project. Table 5 provides an overview of the projects. Subjects were first presented these projects separately, and were asked for each project to indicate how the perceived benefits of this project on a 0-8 scale, and whether or not they would be willing to provide a €50.000 subsidy to the project. Subsequently, they were presented with pairs of projects (p1-p3, p2-p4, p5-p7, p6-p8) and asked to indicate which project they would prefer on a seven-point scale ranging from "strong preference for project A" to "strong preference for project B".

The projects were paired such that the absolute number of lives that could be saved was equal within each pair. According to evaluability bias theory, this joint evaluation modus is expected to make subjects (more) indifferent between

the projects within these pairings, as it becomes easier to see that both projects have the same cost-effectiveneness. At the end of these tasks, subjects were asked to fill out a short survey with questions about demographics, as well as several questions to check their understanding of the projects, and a 10-item instrument to measure whether their thinking style was more "rational" or more "experiential" (the REI-10 inventory, see appendix V). This instrument was used since previous research (Bartels, 2006) has shown that susceptibility to the reference group bias strongly correlated with scores on this instrument, with a more "rational" thinking style being associated with lower susceptibility to the reference group bias.

Table 3.5 Overview of cases, study 2

| Project | City | Number of lives threatened by disease | % of lives saved without project | % of lives saved with project |
|---|---|---|---|---|
| 1 | Gok Machar | 10,000 | 5% | 50% |
| 2 | Melut | 10,000 | 50% | 95% |
| 3 | Abiemnom | 225,000 | 5% | 7% |
| 4 | Bentiu | 225,000 | 93% | 95% |
| 5 | Ghaat | 5,000 | 5% | 50% |
| 6 | Paloich | 5,000 | 50% | 95% |
| 7 | Geigar | 112,500 | 5% | 7% |
| 8 | Marial Bai | 112,500 | 93% | 95% |

## *Empirical strategy*

To estimate the effect of expertise on susceptibility to the reference group bias we use the same regression models as specified in the first study (models 3.1-3.3). To estimate the effect of switching from separate to joint evaluation of projects we first compute the differences between the ratings for the two separate projects that were paired in the joint evaluation modus, as to be able to determine whether subjects preferred project A or project B. In each pairing, the absolute number of lives that could be saved was identical, whereas the reference group was smaller in

project A. This ranking variable is denoted as variable $Y_{ij}$. Each pair of cases is denoted by the subscript $_j$. For example, project 1 (reference group of N=10.000) was paired with project 3 (reference group of N=225.000). If a subject gave a higher rating to project 1 the dependent variable was assigned value=1, whereas if the rating of was lower or equal the dependent variable was assigned value=1.

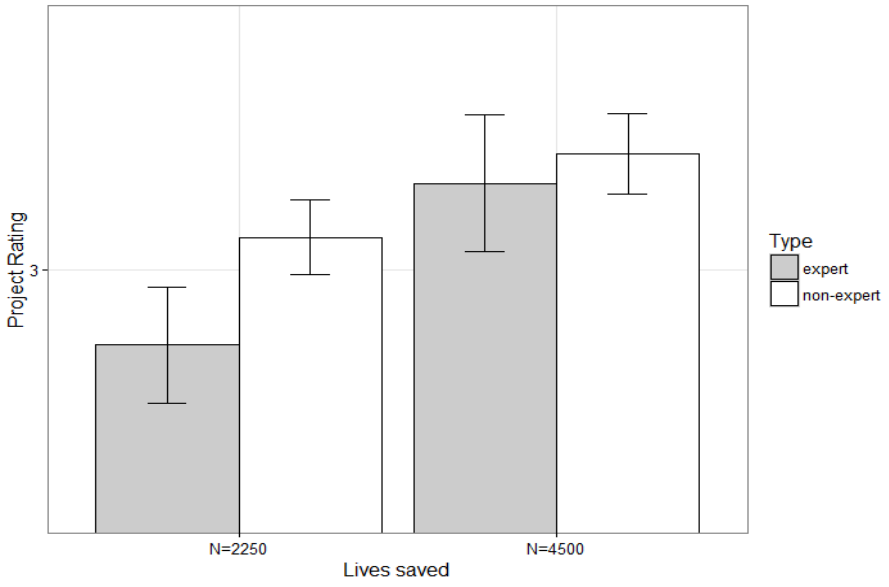
To estimate the effect of the joint evaluation modus we use an OLS model in which we regress variable $Y_{ij}$ on the variable "joint evaluation" ($J_{ij}$) with value equal to one if the projects were presented jointly and zero if separately. We interact this joint evaluation variable with expertise ($J*E_{ij}$) to test if experts benefit more from the joint evaluation modus. As a robustness check we control for age, gender and thinking style as measured with the REI-10 inventory:

$$Y_{ij} = \beta_0 + \beta_1 J_{ij} + \beta_2 J*E_{ij} + \alpha_i + \varepsilon_{ij} \qquad (4.1)$$

## *Results*

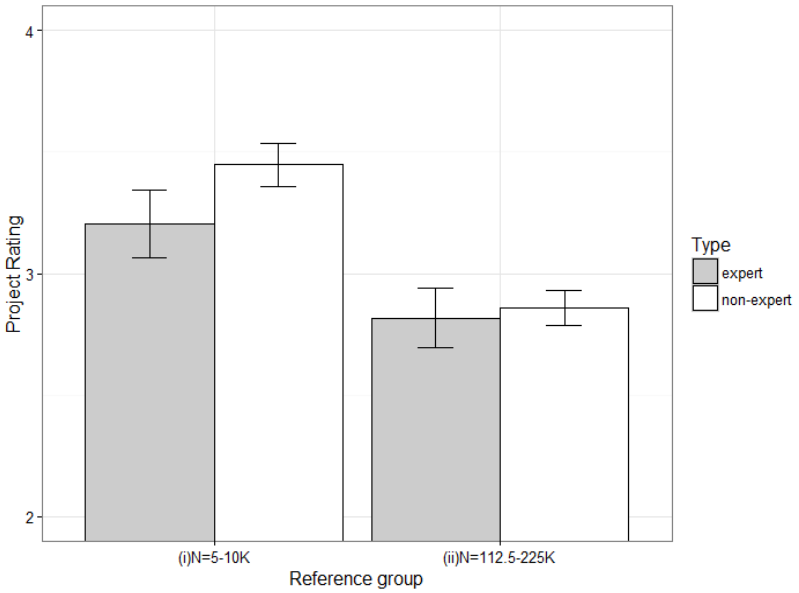
Figure 3.3 indicates the preference rating for professionals and non-professionals, according to the absolute number of lives to be saved. For both groups the project ratings is higher in projects where 4.500 lives can be saved than in cases where 2.250 lives can be saved. A two-sample t-test shows that these differences are statistically insignificant for non-experts ($Pr(|T|>|t|) = 0.1289$) and only significant at the 10 percent level for experts ($Pr(|T|>|t|) = 0.0730$). The results from our regression analyses similarly indicate that project rating increases in number of lives, as indicated by the positive and significant coefficients on the variable "lives saved" in all specifications, both for the OLS and fixed effects models ($P>|t|<0.01$). Furthermore, we find no significant difference between experts and non-experts in terms of this effect, as indicated by the insignificant coefficient on the variable "expertise*lives saved".

Figure 3-3 Project rating and number of lives saved, study 2




Next, we investigate sensitivity to the reference group bias. As reflected in figure 3-4 both experts and non-experts give lower ratings to projects with larger reference groups. These differences are significant both for non-experts ($Pr(|T|>|t|) = 0.0000$) and experts ($Pr(|T|>|t|) = 0.0387$). Similar results are found when comparing responses to the question in which subjects are asked whether or not they wish to support the project. In line with these results we find that both groups are susceptible to the reference group bias, as indicated by the positive coefficient on the variable "reference group size" ($P>|t|<0.01$) in both the OLS and fixed effect model specifications. Furthermore, we find that experts are slightly less susceptible to the reference group bias ($P>|t|<0.1$), as indicated by positive coefficient on the interaction term "expertise*reference group size" in OLS specification 4 and FE specification 3.

Figure 3-4 Project rating and reference group, study 2




We then turn to the issue how the evaluation mode influences project assessment. In the separate evaluation mode subjects are shown eight individual projects sequentially. In the joint evaluation mode subjects are shown these same eight projects grouped in four pairings. Each pairing consists of two projects that differ in the size of the reference group, but have the same absolute number of lives that can be saved. In order to compare the assessment of cases across these two evaluation modes, we construct a variable with value=1 if subjects preferred the project with a smaller reference group, and value=0 if otherwise. For example, if in the separate evaluation modes the subject gave a rating of 6 to project 1 (small reference group) and a rating of 1 to project 3 (large reference group) the dependent variable "preference for project with smaller reference group" is assigned value=1. Similarly, if in joint evaluation mode the subject rated project 3 as being more favorable than project 1 the dependent variable is assigned value=0.

Our main result is that joint evaluation of projects significantly reduces the propensity to prefer projects with a smaller reference group in situations where the

absolute number of lives that can be saved is the same. This result is found in Mann-Whitney test (z-score=4.627, Pr.>|z|=0.00) as well as in the first OLS regression with subject fixed effects as presented in appendix 3 (coefficient= -0.137, robust standard error=0.038). In other words, we find that the joint evaluation mode makes people less susceptible to the reference group bias.

Subsequently, we investigate whether the effect of the joint evaluation mode on susceptibility to the reference group bias is different for experts and non-experts. Our results indicate that the effect is stronger for non-experts. The results from our second and third model specifications show a positive, albeit statistically insignificant, coefficient on the variable "Joint evaluation*expertise" (coefficient=0.145, robust standard error=0.089). This indicates that non-experts benefit more from the joint evaluation mode. This result could be related to the fact that non-experts subject in this experiment were slightly more susceptible to the reference group bias to begin with, so they had more to gain from having the cases presented to them in pairs.

Figure 3-5 Evaluation mode, expertise and bias

| | Preference for smaller reference group | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Joint evaluation | -0.137*** | -0.177*** | -0.177*** |
| | (0.038) | (0.042) | (0.042) |
| Joint evalation*expertise | | 0.145 | 0.145 |
| | | (0.089) | (0.089) |
| Pair controls | No | No | Yes |
| Constant | 0.565*** | 0.565*** | 0.569*** |
| | (0.019) | (0.019) | (0.025) |
| Observations | 1,136 | 1,136 | 1,136 |
| R-squared | 0.033 | 0.040 | 0.044 |
| Number of subjects | 142 | 142 | 142 |

OLS regression with robust standard errors in parentheses. Dependent variable is whether or not for a pair of projects, which were presented first sequentially and later jointly, subjects had a preference for the project with the smaller reference group (1=yes, 0=no). *** p<0.01, ** p<0.05, * p<0.1

## 3.5 Conclusion

In this study we investigate the role of the reference group bias in the disproportionate allocation of charitable funds to projects that target victims in relatively small populations, rather than victims in wider populations. In line with previous studies we use hypothetical cases with variation in the absolute and relative number of lives that can be saved. We find that project evaluations are indeed more strongly affected by their potential for relative life savings rather than their potential for absolute live savings. Through a series of experiments with students as well as professionals from the charitable sector in the Netherlands we find that this bias occurs in both experts and non-experts. It is worth noting that the reference group bias is just one potential driver for the disproportionate allocation of funds to certain causes, next to factors such as social distance between donors and recipients (Engel, 2011a) and purported urgency of the respective problems (Sargeant, 1999). Furthermore, we find that among this professionals with experience in the allocating of funds to social projects are only slightly less susceptible to the bias than professionals without this expertise. This finding is in line with studies that show that cognitive biases are displayed by professionals in a range of domains, including medicine, finance and international development.

Subsequently, we find that susceptibility to the reference group bias can be reduced through joint- rather than separate evaluation of projects. This finding is in line with previous studies that show that subjects attach more decision weight to choice attributes as these attributes become easier to evaluate through between-option comparison (Caviola et al., 2014; Ritov and Baron, 2011). In sum, our study provides evidence that the tendency to favor relative life savings over absolute life savings is pervasive both among students as among professionals within organizations that allocate funds to charitable projects.

# Appendix 3.1 Results and Instructions (study 1)

*Regression tables*

Table 3.6 OLS-estimates of reference group effect, pooled data

| | Willingness to Contribute | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Reference group (log) | -0.103*** | -0.103*** | -0.108*** | -0.083** | -0.100** |
| | (0.026) | (0.026) | (0.037) | (0.039) | (0.041) |
| Number of lives saved | 0.001** | 0.001** | 0.001** | 0.001* | 0.001** |
| | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) |
| Expertise | | -1.246*** | -0.998* | -0.367 | -0.506 |
| | | (0.265) | (0.593) | (0.640) | (0.658) |
| Expertise*ref.group | | | 0.011 | 0.000 | 0.008 |
| | | | (0.052) | (0.053) | (0.053) |
| Expertise*lives | | | -0.001 | -0.001 | -0.000 |
| | | | (0.001) | (0.001) | (0.001) |
| Subject controls | No | No | No | Yes | Yes |
| Case controls | No | No | No | Yes | Yes |
| Session controls | No | No | No | Yes | Yes |
| Last case excl. | No | No | No | No | Yes |
| Observations | 1,556 | 1,556 | 1,556 | 1,548 | 1,354 |
| R-squared | 0.008 | 0.072 | 0.072 | 0.144 | 0.134 |

OLS regression with robust standard errors in parentheses, The dependent variable is the reported likelihood of support the project on a 1-10 scale. *** p<0.01, ** p<0.05, * p<0.1

Table 3.7 Fixed effects estimates of reference group effect, pooled data

| | Willingness to Contribute | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Reference group (log) | -0.103*** | -0.108*** | -0.081** | -0.081** |
| | (0.026) | (0.037) | (0.039) | (0.039) |
| Number of lives saved | 0.001** | 0.001** | 0.001* | 0.001 |
| | (0.000) | (0.000) | (0.001) | (0.001) |
| Expertise*ref.group | | 0.012 | -0.000 | 0.000 |
| | | (0.052) | (0.052) | (0.052) |
| Expertise*lives | | -0.001 | -0.001 | -0.000 |
| | | (0.001) | (0.001) | (0.001) |
| Case controls | No | No | Yes | Yes |
| Excl. last case | No | No | No | Yes |
| Observations | 1,556 | 1,556 | 1,556 | 1,361 |
| R-squared | 0.021 | 0.021 | 0.073 | 0.038 |
| Number of subjects | 195 | 195 | 195 | 195 |

OLS regression with robust standard errors in parentheses, The dependent variable is the reported likelihood of support the project on a 1-10 scale. *** p<0.01, ** p<0.05, * p<0.1

Table 3.8 Fixed effect estimates of reference group effect, per session

| | Willingness to Contribute (WTC) | | | | | |
|---|---|---|---|---|---|---|
| | Session 1 *(profesionals & students,lab)* | | Session 2 *(students,lab)* | | Session 3 *(professionals, online survey)* | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Reference group (log.) | -0.051 | -0.067 | -0.090 | -0.085 | -0.127** | -0.123* |
| | (0.067) | (0.071) | (0.068) | (0.066) | (0.063) | (0.063) |
| Nr of lives saved | 0.000 | 0.000 | 0.002** | 0.002** | 0.001 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Expertise*ref. group | 0.001 | 0.012 | | | 0.052 | 0.056 |
| | (0.078) | (0.081) | | | (0.080) | (0.077) |
| Expertise*lives | 0.000 | 0.000 | | | -0.001 | -0.001 |
| | (0.001) | (0.001) | | | (0.001) | (0.001) |
| Case effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Last case excl. | No | Yes | No | Yes | No | Yes |
| Observations | 711 | 622 | 365 | 319 | 480 | 420 |
| R-squared | 0.119 | 0.060 | 0.069 | 0.043 | 0.085 | 0.089 |
| Number of subjects | 89 | 89 | 46 | 46 | 60 | 60 |

OLS regressions with individual fixed effects, robust std. errors in parentheses. The dependent variable is the reported likelihood of support the project on a 1-10 scale. *** p<0.01, ** p<0.05, * p<0.1

*Survey Instrument*

Can you please provide the following information about yourself?

1) What is your age? (years)
2) What is your gender? (F/M)
3) What type of organization do you work for? (private sector, public sector, non-profit, other)
4) In your work, are your involved with grants/investments for social projects? (yes/no)

Can you please indicate how you rank the following criteria of social projects in your assessment of the cases in the experiment (1=least important, 3=most important)?

- Absolute number of lives saved (e.g. 500 people)
- Relative number of lives saved (e.g. 500 out of de 150.000)
- Other factors (disease, country, region)

## *Wording of cases*

The section below is the English translation of the first two cases that was presented to the subjects in group one. Other cases, for both groups, followed the same structure and wording. The description about the disease was taken from Wikipedia. The information about the number of people in the respective region and number of people suffering from the disease is hypothetical.

CASE 1 (Group 1)
Each year over 1000 people in the Ugandan region Luweero die from Shigellosis. This is an infection with bacteria of the Shigella genus. This species typically multiplies in the intestines. Infection with the rod-shaped gram-negative Shigella bacteria is a common cause of dysentery. The disease is usually contracted in overpopulated areas with limited hygiene. The main treatment consists of administering oral rehydration salts to combat dehydration. A Dutch NGO that works in this region requests a grant to provide oral rehydration salts to the local population. With the proposed project an estimated 500 lives will be saved in the saved in the region in the upcoming year.

| | |
|---|---|
| On a 1-10 scale, how likely is it that you would honour this grant?<br>(1 means very low chance, 10 means very high chance) | |

CASE 2 (Group 1)
Sleeping sickness is caused by a protozoan parasite, Trypanosoma Brucei, which is spread by the tsetse fly. This disease can be dormant until the parasite has entered the brain. Every year around 50.000 people in Africa die as a result of sleeping sickness. Patients in the advanced stages of the disease can be cured with the recently developed medicine 'Eflornithine'. A Dutch NGO active in the Tanzanian region Ruvuma request a grant to distribute the medicine there. With the proposed project an estimated 750 lives will be saved in this region in the upcoming year.

| | |
|---|---|
| On a 1-10 scale, how likely is it that you would honour this grant?<br>(1 means very low chance, 10 means very high chance) | |

# Appendix 3.2 Results and Instructions (study 2)

*Regression tables*

Table 3.9 OLS-estimates of reference group effect, pooled data

| | Project rating | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Number of lives saved | 0.860*** | 0.860*** | 0.790*** | 0.607*** |
| | (0.077) | (0.077) | (0.084) | (0.146) |
| Reference group (log) | -0.423*** | -0.423*** | -0.467*** | -0.496*** |
| | (0.048) | (0.048) | (0.057) | (0.076) |
| Expertise | | -0.296 | -1.031** | -0.776 |
| | | (0.304) | (0.452) | (0.533) |
| Expertise*lives | | | 0.275 | 0.074 |
| | | | (0.192) | (0.181) |
| Expertise*reference group | | | 0.170* | 0.177 |
| | | | (0.101) | (0.110) |
| Subject controls | No | No | No | Yes |
| Case controls (incl. order) | No | No | No | Yes |
| Observations | 1,212 | 1,212 | 1,212 | 1,072 |
| R-squared | 0.103 | 0.106 | 0.110 | 0.140 |

OLS-regression, robust std. errors in parentheses. Dependent variable is the rating of the project benefit on a 0-8 scale (8=high benefit). Subject controls include: age, gender, REI (rational vs. experiential inventory) score. Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.10 Fixed effects estimates of reference group effect, pooled data

| | Project rating | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Number of lives saved | 0.860*** | 0.860*** | 0.790*** | 0.584*** |
| | (0.077) | (0.077) | (0.084) | (0.149) |
| Reference group (log) | -0.423*** | -0.423*** | -0.467*** | -0.498*** |
| | (0.048) | (0.048) | (0.057) | (0.075) |
| Expertise | | -0.296 | -1.031** | -0.783 |
| | | (0.304) | (0.452) | (0.540) |
| Expertise*lives | | | 0.275 | 0.085 |
| | | | (0.192) | (0.181) |
| Expertise*reference group | | | 0.170* | 0.177 |
| | | | (0.101) | (0.110) |
| Subject controls | No | No | No | Yes |
| Case controls (incl. order) | No | No | No | Yes |
| Observations | 1,212 | 1,212 | 1,212 | 1,072 |
| R-squared | 0.103 | 0.106 | 0.11 | 0.14 |

OLS-regression with fixed effects, robust std. errors in parentheses. Dependent variable is the rating of the project benefit on a 0-8 scale (8=high benefit). Subject controls include: age, gender, REI (rational vs. experiential inventory) score. Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Survey instrument*

1) Was it clear to you that projects 1-4 were the same in the sense that 4500 lives could be saved (45% of 10,000 or 2% van 225,000), and similarly that projects 5-8 were the same in the sense that 2250 lives that could be saved (45% van 5,000 or 2% van 112,500)? *(Yes/No)*
2) According to you, how does the value of a human life in smaller refugee camp compare to the value of a human life in a larger refugee camp? (*Same value regardless of camp size, Larger value in larger camp, Larger value in smaller camp)*
3) In case you think the value of a human life depends on the camp size, can you explain why you think this is the case? *(Open question)*
4) Did you think that in large camps there would be more chance of conflicts? *(Yes/No)*
5) Was it clear to you that patients in the camp couldn't infect others (as the disease was non contagious)? *(Yes/No)*
6) What is your age? *(Years)*
7) What is your sex? *(F/M)*
8) Do you donate to charities? *(Yes/No)*
9) What do/did you study? *(Economics, Natural Sciences/Mathematics, Social Sciences, Law, Medicine, Other)*
10) Are you currently a student? *(Yes/No)*

11) Do you have professional experience in providing subsidies to projects/programs? (*Yes, in deciding which program/projects receive subsidies; Yes, in supporting these decisions (e.g. preparing documents); Yes, in a different way (fund raising, admin etc); No)*
12) Do you have professional experience specifically with projects aimed at refugee camps? *(Yes/No)*
13) Finally we will show you a few statements Can you please indicate to what extent you agree with these statements (1=completely disagree, 5=completely agree)
    - I don't like to have to do a lot of thinking
    - I try to avoid situations that require thinking in depth about something
    - I prefer to do something that challenges my thinking abilities rather than something that requires little thought
    - I prefer complex to simple problems
    - Thinking deeply gives me little satisfaction
    - I trust the first impressions I have in people
    - I believe I can trust my feelings
    - My first impression of people is usually correct
    - In terms of trusting others I can usually follow my instincts
    - I usually feel if people are right or not, even though I can't explain

## *Wording of cases*

Instructions

Imagine, you are working for a charitable foundation that provides support to refugees in developing countries. According to the United Nations Refugee Agency (UNHCR) and the World Health Organization (WHO) there is a threat of a humanitarian disaster in refugee camps in South-Sudan. Many refugees in the refugee camps have contracted a disease that results from drinking contaminated water. The disease is not infectious. Those refugees will die quickly if they will not get access to clean drinking water. You are asked to consider a number of projects that provide clean drinking water to the refugees in the refugee camps in South-Sudan. All projects provide clean drinking water with a Dash-8 airplane. Once the plain arrives at the refugee camps, local aid-workers will distribute the water to the designated patients, as a result of which the patients' lives are usually saved. All projects cost about € 50.000. However, the charitable organization does not have enough funds to finance all projects. Now we would like to ask you a few questions about eight specific projects.

Example case, separate evaluation mode

In the city of Paloich in South-Sudan there are about 5.000 refugees suffering from the lethal disease. Currently, with the available amount of clean drinking water in the Paloich refugee camp 50% of the patients will be saved. If the Dash-8 waterplane is sent to Paloich, 95% of the patients will be saved.

- What would be the benefit of sending the Dash-8 waterplane to this refugee camp? (slider bar: 0=extremely low benefit, 8=extremely high benefit)

- Given the benefit you indicated, would it be worth the €50.000 subsidy to send the plane to this camp? (yes, no)

Example case, joint evaluation mode

- *Project A)* In the city of Ghok Machar in South-Sudan there are approximately 10.000 refugees that are suffering from the lethal disease. Currently, with the available amount of clean drinking water in the Ghok Machar refugee camp 5% of the patients will be saved. If the Dash-8 water-plane is sent to Ghok Machar, 50% of the patients will be saved

- *Project B)* In the city of Abiemnom in South-Sudan are about 225.000 refugees suffering from the lethal disease. At the moment, there is clean drinking water in the Abiemnom refugee camp to save 5% of the patients. Once we allow the Dash-8 airplane to go to Abiemnom, 7% of the patients will have access to clean drinking water, so they will be safe.

    - Which project do you prefer? (slider bar)
        - strong preference for project A
        - preference for project A,
        - slight preference for project A,
        - no preference,
        - slight preference for project B,
        - preference for project B,
        - strong preference for project B

Chapter 4

# Free to help?

## An experiment on free will belief and altruism*

"We must believe in free will, we have no choice."

- Isaac Bashevis Singer –


## Abstract

How does belief in free will affect altruistic behavior? In an online experiment we undermine subjects' belief in free will through a priming task. Subjects subsequently conduct a series of binary dictator games in which they can distribute money between themselves and a charity that supports low-income people in developing countries. In each decision task, subjects choose between two different distributions, one of which is more generous towards the charity. In contrast to previous experiments that report a negative effect of undermining free will on honest behavior and self-reported willingness to help, we find an insignificant average treatment effect. However, we do find that our treatment reduces charitable giving among non-religious subjects, but not among religious subjects. This could be explained by our finding that religious subjects associate more strongly with social norms that prescribe helping the poor, and might therefore be less sensitive to the effect of reduced belief in free will. These findings indicate that the effects of free will belief on prosocial behavior are more nuanced than previously suggested.


* Joint work with Kellie Liket, John Protzko and Vera Schölmerich.

## 4.1 Introduction

### *Free will belief and social behaviors*

Do humans have free will? This is the topic of an ancient debate that remains unresolved up to this day (Aquinas, 1981; Aristotle and Book, 1933; Dennett, 2015; Schopenhauer, 1839). The implications of this debate extend beyond the intellectual realm. Despite the lack of consensus in the academic community, people young and old across the world believe that they have free will (Kushnir, 2012; Sarkissian et al., 2010) and most people even believe they have more free will than others (Pronin and Kugler, 2010). Some scholars argue that widespread belief in free will has evolved as it allows for larger and more complex societies to function and thrive (Baumeister and Monroe, 2014). Instilling in people a sense of control over their actions, this belief has allowed for the justification of rules and institutions that punish anti-social behavior and reward pro-social behaviour. Thus, rather counter-intuitively, the belief in free will is proposed to have enabled humans to become better at adhering to social norms.

Greater belief in free will has been associated in observational research with a range of positive outcomes, including better career prospects and higher job performance (Stillman et al., 2010). Furthermore, experimental evidence also points towards benefits of greater belief in free will, for example by promoting appreciation towards acts of kindness by others, who were "free" to also be unkind (MacKenzie et al., 2014).

However, the level of people's belief in free will and their locus of control [i.e. the extent to which people attribute control to themselves vs. their environment] have been declining in recent decades (Twenge et al., 2004). Belief in free will and locus of control are strongly correlated (Stillman et al., 2010) and conceptually related: without a belief in free will it is more difficult to attribute control to oneself. The decline of these beliefs coincides with the popularization of insights from neuroscience, for example the famous Libet experiments (Libet,

1993) which conclude that free will is an illusion (Harris, 2012; Montague, 2008). Neuroscience experiments by Libet and others demonstrate that information about brain activity can be used to predict decisions before the decision-maker becomes aware of making a decision. Various philosophers contest the claim that free will does not exist (Bourget and Chalmers, 2013; Dennett, 2015). Regardless of whether the inferences from neuroscientific evidence to the supposed impossibility of the existence free will are correct, the changing attitudes in society about free will and self-control have been shown to influence various social behaviours.

Various lab experiments have undermined people's belief in free will by priming tasks, in which subjects read texts about neuroscientific evidence implying the non-existence of free will. These studies tend to show that exposure to such primes undermines honesty and willingness to help others. One study shows that undermining belief in free will causes increased cheating in tasks where subjects could earn more money by lying (Vohs and Schooler, 2008b). A study by Baumeister and colleagues 2009 finds that experimental reduction of belief in free will through a reading task lowers people's likelihood of reporting to be willing to help others in various hypothetical scenarios (Baumeister et al., 2009). It should be noted that this study did not look into the relationship between free will disbelief manipulations and *actual* helping of others. This question has yet to be empirically investigated.

Next to their experimental findings, Baumeister and colleagues 2009) also reported that subjects with a stronger *dis*belief in free will were less likely to sign up for volunteer work absent of experimental manipulations). One interpretation for these findings is that a disbelief in free will gives people an excuse to justify their selfish tendencies (Smilansky, 2000). In other words, people can refrain from engaging in prosocial behavior and then justify this by explaining that they have very little control over their own behavior. Further supporting these findings, studies show that undermining free will belief reduces people's ability to control themselves (Baumeister et al., 1994; Baumeister and Vohs, 2007; Vohs and

Heatherton, 2000). This reduced self-control would, in turn, diminish people's willingness to act prosocially (Baumeister et al., 2009). Taken together, the current evidence to support the view that free will disbelief undermines prosocial behavior is still rather limited.

While studies have shown that undermining free will belief leads to increased cheating and reduced likelihood of reported willingness to help others, other studies have found that undermining free will belief can actually promote sympathy for others. For example, lowered free will beliefs have been shown to reduce the attribution of blame of criminal offenders (Krueger et al., 2014; Shariff et al., 2014). This finding suggests that undermining people's belief in free will increases their perception that other people are shaped by forces outside of their own control, such as their genetic composition or their upbringing. Indeed, another experiment shows that the tendency to blame others depends on the perceived level of control that others have. In this experiment (Monroe et al., 2016), subjects were shown a video in which an 'active' person interacted with another 'passive' person, from whom they could steal money. In both treatments the active person stole the entire endowment [$10] from the passive person. In one treatment, this action was the result of a random process outside of the control of the active person [die roll], whereas in the other treatment it was the result of an active choice made by the passive person. The observing subject was then asked to judge the blameworthiness of the active person, and the authors [ibid] found that more blame was attributed in the condition were an active choice was made. In a similar vein, Fong (Fong, 2007) finds that people, even those that perceive themselves as being unconditionally altruistic, donate more to welfare recipients if they are informed that these recipients are actively looking for work as opposed to waiting for a work opportunity to arise. The results of these suggest that undermining free will belief may also promote prosocial behavior by increasing people's perception of the lack of control other people have on their own lives and consequently increasing their willingness to help.

Another open question is whether the effects of free will beliefs on social behavior are homogenous across different groups. Various studies have shown that a range of behavioral patterns observed in lab experiments with college students do not generalize to other populations that are less educated, rich and westernized (Arnett, 2008; Henrich et al., 2010b). Since all previous experiments with free will manipulations were conducted among college students, it remains unclear how universal the effects of free will beliefs on social behavior are. One group that might respond differently to free will manipulations are religious people, as belief in free will is higher among religious people. Various studies have shown that religious affiliation is associated with higher charitable giving (Bekkers and Wiepking, 2011), higher propensity to volunteer to help the poor and elderly (Phillips et al., 2011) and higher giving to charities in dictator games – an activity where subjects are given a sum of money and can decide how much they want to donate to another subject (Nahmias et al., 2014). It should be noted that these studies apply to western countries where Christianity is the main religion. Furthermore, experimental studies provide causal evidence to support the theory that religious primes can promote honesty and prosocial behavior, both in the lab (Mazar et al., 2008; Shariff and Norenzayan, 2007) and in the field (Duhaime, 2015). Moreover, there is some evidence that religious primes have different effects on social behaviors according to the religious status of subjects. For example, the willingness to engage in costly punishment of free-riders in a public-goods game was increased by a religious prime, but only among subjects that had previously made religious donations (McKay et al., 2011).

To summarize, there is currently mixed evidence on whether undermining free will belief will lead to more prosocial behaviour, it is not clear how this effect will play out amongst a diverse population, and there is no experimental evidence on whether undermining free will belief influences people's actual likelihood of helping others. In this study we tested the effects of free will manipulations on actual behaviour among a diverse population. We did this by conducting an online

experiment using a dictator game with subjects recruited via the Amazons mTurk platform.

One of the most commonly used methods to study actual [as opposed to self-reported] prosocial behavior is the dictator game. In this two-player game, one subject – the dictator - is given a sum of money, and can decide how much of this money they want to donate to the other person playing – the recipient. This recipient is passive and can do nothing but accept whatever fraction of the sum of money they are given. Under conditions of anonymity, the rational strategy for a purely selfish dictator is to give nothing, but a meta-analysis with data from hundreds of dictator game experiments shows that people donate on average between 25-30% of their money (Engel, 2011b). Given its simplicity, the dictator game is a useful tool to study the factors that shape prosocial behavior.

The Amazon mTurk platform allows people to earn money by completing small tasks see Methods - Procedures for more details and is increasingly used in social science experiments (Buhrmester et al., 2011; Casler et al., 2013) and provides access to a more population that is more diverse than college students in terms of demographics, socio-economic and cultural background. For example, an experiment on mTurk used a dictator game with subjects in the U.S. and India, and found that the latter group were more sensitive to the size of the endowment (Raihani et al., 2013). Another dictator game experiment with subjects from different countries recruited through mTurk found substantial heterogeneity in dictator game play across cultures (Raihani and Bshary, 2012). Furthermore, such experiments can be used not only to test whether behavior in standardized experiments differs across groups, but also whether these various groups respond differently to experimental manipulations.

Our paper aims to address the following question: "How does undermining belief in free will affect altruistic behavior?" We measure altruism in terms of behavior in a binary dictator game where subjects can divide money between themselves and a charity. In line with several previous experiments about

free will belief, we hypothesize that undermining belief in free will make people less inclined to engage in charitable giving. Furthermore, we hypothesize that this manipulation could have different effects among a more diverse sample of subjects.

## 4.2 Methods

### *Procedure*

Subjects were recruited via Amazon mechanical Turk (mTurk), a crowdscourcing website that is increasingly used in the social sciences. Although providing less control over experimental conditions than lab experiments, various studies show that results obtained through mTurk are comparable to results from the lab (Crump et al., 2013; Horton et al., 2011; Paolacci et al., 2010). At the onset of the experiment subjects were informed that the study was about the effect of exposure to text on happiness. The purpose of this was to reduce socially desirable responses due to observer bias and to prevent subjects from making a connection between the manipulation we performed and our dependent variable. Following the introduction to the experiment, subjects were asked to rate their happiness on a 1-10 scale. Subsequently, subjects were exposed to the treatment or control text. To ensure that subjects read this text, they could only click to the next page one minute after opening the page with the treatment or control text. Furthermore, subjects were requested to write a short summary of the text and they were informed that their payment could be affected if they did not do so. They were then again asked to rate their happiness. Subsequently, they continued to a set of 24 decision tasks, followed by a short survey.

### *Treatment*

The treatment consisted of subjects being asked to read a 1-page article from the popular science journal "NewScientist" in which neurological scientific evidence

is presented to support the notion that humans do not have free will. The control group was shown another 1-page article from the same magazine about sustainable energy technologies (see S1 Appendix for the full text of both treatment and placebo). Subjects were asked to write a 1-2 sentence summary of the text in order to demonstrate that they had carefully read the text. This manipulation has successfully been previously (Shariff et al., 2014).

## *Decision task*

In a second step of this experiment, subjects were told that they could allocate monetary tokens to themselves or to GiveDirectly, a charity that provides direct cash transfers to low-income households in sub-Saharan Africa. Subjects were informed that these cash transfers would be given to "people like Beatrix", followed by a short description of this woman's situation accompanied by a photograph of the illustrative recipient and her two children. This information was taken from the website of the charity. The exact wording to describe the example recipient was as follows: "An example of a family benefiting from GiveDirectly is Beatrice (31yrs) and her two young children, living in Kenya." Subjects then completed 24 binary Dictator Games (DGs), which are a widely used tool to measure social preferences (Krajbich et al., 2015; Schulz et al., 2014). In our experiment, each binary DG consisted of two different distributions of tokens between the subject (i.e. the dictator) and the charity. For example, subjects could choose between option A) keep 50%, give 50% to charity or option B) keep 0%, give 100% to charity. Henceforth, we refer to each DG as a 'decision task'. The number of experimental tokens that could be earned per task ranged between 0-60, with a conversion rate of 1 dollar cent/token. In other words, subjects had the chance of earning up to 60 cents per task. Previous research on mTurk has shown that dictator games with stakes of max. $1 yield similar outcomes to higher stakes (Raihani et al., 2013) To ensure that subjects had an incentive to reveal their true

preference in each task as they were informed that one of the games would be randomly selected at the end of the experiment and then paid out according to the choices the dictator made (Becker et al., 1964). The order of the decision tasks was randomized to control for order effects. Subjects were informed that they had 10 seconds per task. They were also informed that if they would not choose within this timeframe, then they or the charity would not receive any money in case this task was randomly selected to be played for real money at the end of the experiment.

In each decision task, one of the two options provided a higher payoff to the charity but a lower payoff to the dictator (see Table 1). In our analysis, we classify this as the more altruistic option. The first 12 decision tasks consisted of choices between an equal distribution and an unequal distribution. For example, the fair allocation for task 1 is option A: 50% dictator/50% charity and the unfair allocation is option B: 100% dictator/0% charity. The second set of 12 tasks consisted of choices between two unequal allocations. For example, task 13 option A was 100% dictator/0% charity and option B was 0% dictator/100% charity.

Decision tasks also differed in terms of whether the more altruistic option consisted of the same, less or more money than the less altruistic option. By varying the size of the allocation we wanted to investigate whether the treatment effect was greater for options where the less altruistic option was more efficient, i.e. resulted in larger potential earning. Moreover, we varied whether or not both options provided at least some payoff to both players. In this way we wanted to investigate whether the treatment had a greater effect when dictators could "excuse" their selfish behavior by selecting an option that provided at some payoff to the charity.

Table 4.1 Overview of decision-tasks

| Task | Option A (self, charity) | Option B (self, charity) | Task | Option A (self, charity) | Option B (self, charity) |
|---|---|---|---|---|---|
| 1 | (50,50) | (100,0) | 13 | (100,0) | (0,100) |

| | | | | | |
|---|---|---|---|---|---|
| 2 | (50,50) | (80,20) | 14 | (100,0) | (20,80) |
| 3 | (50,50) | (20,80) | 15 | (80,20) | (0,100) |
| 4 | (50,50) | (0,100) | 16 | (80,20) | (20,80) |
| 5 | (60,60) | (100,0) | 17 | (100,20) | (0,100) |
| 6 | (60,60) | (80,20) | 18 | (100,20) | (20,80) |
| 7 | (60,60) | (20,80) | 19 | (80,40) | (0,100) |
| 8 | (60,60) | (0,100) | 20 | (80,40) | (20,80) |
| 9 | (40,40) | (100,0) | 21 | (80,0) | (0,100) |
| 10 | (40,40) | (80,20) | 22 | (80,0) | (20,80) |
| 11 | (40,40) | (20,80) | 23 | (60,20) | (0,100) |
| 12 | (40,40) | (0,100) | 24 | (60,20) | (20,80) |

## *Questionnaire*

After the decision tasks, subjects were asked to indicate their sex, age, and perceived socio-economic status. In addition subjects were asked whether they identified with a religion: Christianity, Islam, Hinduism, Buddhism, Judaism, other, or no religion. Subjects who responded to not identify with a religion were labelled as non-religious. After presenting these survey questions we also asked subjects to indicate their self-reported level of free will, measured on a 100-point scale. This question served as a manipulation check. In addition, subjects were asked if they considered whether recipients of the charity had control over their personal situation and whether they thought that "one ought to help" people such as these recipients (see 'Mechanisms' in the Results section for exact phrasing). This last item allowed us to investigate whether adherence to social norms might moderate the effect of the free will disbelief treatment on altruistic behavior. For more details about these items please see S1 Appendix.

## *Subjects and randomization*

Subjects were recruited from U.S. members of Amazon mTurk in August 2016. As can be seen in the 'Full sample' column of table 2, a total of 108 subjects participated. This sample consisted of 64% females. The average age was 34.2 years. Approximately half of our subjects considered themselves religious, the

majority of religious subjects subscribing to various Christian denominations. Moreover, most subjects report to perceive their own socio-economic status to be right in the middle between most and least successful. Differences between treatment and control group in terms of age and sex were not statistically significant. A Mann-Whitney U-test shows that our random assignment did result in a higher fraction of non-religious subjects being allocated to the treatment group (Pr.>|z=0.0075) as well as a higher level of self-perceived socio-economic status in the treatment group (Pr.>|z=0.0928). To account for this imbalance we add these variables as controls in our regressions.

Table 4.2 Summary statistics

| | Control | | | Treatment | | | Full sample | | |
|---|---|---|---|---|---|---|---|---|---|
| | Obs. | Mean | SD | Obs. | Mean | SD | Obs. | Mean | SD |
| Female (1=yes) | 51 | 0.59 | 0.5 | 57 | 0.68 | 0.47 | 108 | 0.64 | 0.48 |
| Age (years) | 51 | 33.53 | 11.59 | 57 | 34.81 | 10.3 | 108 | 34.2 | 10.89 |
| SE status (0-10) | 51 | 4.94 | 1.61 | 57 | 5.47 | 1.6 | 108 | 5.22 | 1.62 |
| Religious (1=yes) | 51 | 0.37 | 0.49 | 57 | 0.63 | 0.49 | 108 | 0.51 | 0.5 |

## *Ethics statement*

All participants in the experiments reported in the manuscript were informed: first, about the protocols of the study that ensure anonymity and confidentiality; second, about the content of the experiment (and the potential monetary earnings) prior to participating. Written consent was obtained from all participants included in this paper. Only those who accepted completed the experiment. Those who did not accept did not continue the experiment. Anonymity was preserved as participants signed up through their Amazon mTurk account number. No association was ever made between their real names/addresses and the results. As is standard in socio-economic experiments, no ethic concerns are involved other than preserving the anonymity of participants. This procedure (including the consent process) was

checked and approved by the Office of Research and Human Subjects at the University of Santa Barbara, the institution hosting the experiments.

## 4.3 Results

### *Manipulation check*

For the full sample, the average level of free will belief was 69.9 on a 0-100 scale (SD=23.6), with 100 indicating total agreement with the statement "I fully believe I have free will", and zero indicating full disagreement. The average level of free will belief in the control group was 72.2 (SD=22.2), and in the treatment group it was 67.8 (SD=24.8). To estimate the effect of the treatment on self-perceived free will, we run an OLS regression of the treatment on the free will measure, controlling for religiosity, demographics and self-perceived socio-economic status. In line with previous studies (Shariff et al., 2014), we find that our treatment reduced belief in free will by 8.9 points on a 0-100 scale, as shown in table 3.

Table 4.3 Manipulation check

| | Free Will Scale (0-100) |
|---|---|
| Treatment "free will disbelief" | -8.930** |
| | (4.115) |
| Constant | 37.306*** |
| | (9.827) |
| Observations | 108 |
| R-squared | 0.264 |

OLS regression with robust standard errors clustered on subjects in parentheses. The dependent variable is the level of agreements subjects reported on a 0-100 scale to the statement "I fully belief I have free will". Controls for age, sex, religiosity and socio-economic status. *** p<0.01, ** p<0.05, * p<0.1

### *Treatment effects*

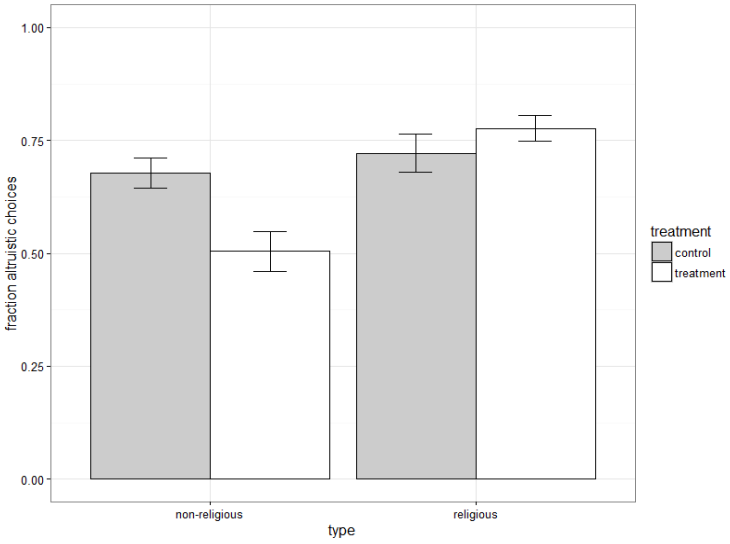
First, we estimate the effect of the treatment on the fraction of decision tasks in which the more altruistic distribution was chosen.

Table 4.4 OLS regression of treatment effects

| | Fraction of altruistic choices | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | full sample | | religious | non-rel. |
| Treatment "free will disbelief" | -0.016 | -0.058 | 0.091 | -0.214** |
| | (0.061) | (0.062) | (0.087) | (0.087) |
| Level of free will belief | 0.000 | -0.002 | -0.001 | -0.001 |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Controls | No | Yes | Yes | Yes |
| Constant | 0.694*** | 0.370*** | 0.609*** | 0.196 |
| | (0.107) | (0.134) | (0.183) | (0.167) |
| Observations | 107 | 107 | 54 | 53 |
| R-squared | 0.001 | 0.186 | 0.293 | 0.299 |

OLS regression with robust standard errors clustered on subjects in parentheses. The dependent variable is what fraction of subjects' choices was for the more altruistic option. Controls include age, sex, religiosity and socio-economic status. *** p<0.01, ** p<0.05, * p<0.1

Figure 4-1 Treatment effect, by religiosity




As can be seen in table 4.4, the treatment did not have a significant effect for the pooled sample. However, as shown in columns 3-4, we find that the treatment did result in a statistically significant decrease in altruistic choices for the non-

religious subjects, of more than 21 percentage points ($P<0.05$). This result is also reflected graphically in figure 4-1.

We then consider the subjects' choice behavior for each individual choice task. To this purpose we use a random effects probit model, with robust standard errors clustered on individuals. The main outcome variable in these regressions is a dummy variable indicating whether or not the subject selected the more altruistic of the two options (1=yes, 0=no).This model allows us to test not only the treatment effect, but also control for specifics of the choices task. In particular, we control for: (i) whether the choice was between two unequal distributions or between an unequal vs. an unequal distribution, (ii) whether one of the choices was more efficient in terms of the total amount to be distributed, (iii) the cost of altruism – i.e. how much extra the subject could earn from selecting the less altruistic option, (iv) the benefit of altruism – i.e. how much extra the charity could earn if the subject selected the more altruistic option. Furthermore, we control for subject characteristics and task order In additional analyses, we also control for moral self-licensing effects as studied in a 2013 paper by Brañas Garza et al. (Brañas-Garza et al., 2013) by including a lag of the variable "altruistic choice" as a control variable in a probit model without random effects. However, we find that no indication of moral self-licensing, as the coefficient on the lag variable is significant and positive.

Table 4.5 Probit regression of treatment effects

| | Altruistic choice | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | full sample | | religious | non-religious |
| Treatment "free will disbelief" | 0.001 | -0.096 | 0.070 | -0.254** |
| | (0.084) | (0.076) | (0.080) | (0.105) |
| Level of free will belief | 0.001 | -0.002 | -0.001 | -0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Unequal vs. unequal | | 0.119*** | 0.103*** | 0.128*** |
| | | (0.030) | (0.039) | (0.046) |
| Selfish option efficient | | -0.067*** | -0.055*** | -0.076*** |
| | | (0.016) | (0.017) | (0.028) |
| Cost of altruism | | -0.000 | 0.000 | -0.001** |
| | | (0.000) | (0.000) | (0.001) |
| Controls | No | Yes | Yes | Yes |
| Subjects | 108 | 108 | 55 | 53 |
| Observations | 2,532 | 2,532 | 1,292 | 1,240 |
| Wald $\chi^2$ | 0.19 | 77.54 | 50.49 | 36.80 |

Marginal effects of probit model with subject random effects. The dependent is whether or not the subject selected the more altruistic distribution (value=1 if yes). Standard errors clustered at the subject-level. Observations where subjects did not make a decision within 10 seconds (n=60) were excluded from this analysis. The variable "unequal vs. unequal" has value=1 if both options consisted of unequal distributions, (e.g. 100/0 vs. 0/100) and value=0 if only 1 option consisted of an unequal distribution (e.g. 50/50 vs. 100/0). The variable "selfish option" efficient has value=1 if the total number of tokens to be distributed was greater in the less altruistic option, and value=0 if the amount of tokens was equal in both options. The variable "cost of altruism" indicates the difference between the two options in terms of the number of tokens that could be earned by the dictator. Controls include: age, sex, religiosity, socio-economic status, and task-order.
*** p<0.01, ** p<0.05, * p<0.1

In line with the results from the OLS regression model we find that only among the sub-group of non-religious subjects did the treatment reduce the propensity of subjects to choose the more altruistic distribution, as can be seen by the negative coefficient on the variable "Treatment free will disbelief" in table 5, column 4. The treatment reduces the propensity of non-religious subjects to select more altruistic option by over 25% (significant at P<0.05), whereas for religious subjects the treatment had no significant effect. Furthermore, we find that subjects were more likely to choose the altruistic distribution in choices between two

unequal distributions – one favoring the dictator and one favouring the recipient - than in choices between one equal and one unequal distribution. In addition, we find that subjects are less likely to choose the more altruistic distribution when this option is inefficient - in terms of the total size of the pie – in comparison to the other distribution. This effect occurs for both religious and non-religious subjects. Finally, we find that among non-religious subjects altruism is also influenced by the size of its costs (to the subject), as indicated by the coefficients in column 4 on the variable "cost of altruism".

## *Thou Shalt Help?*

Subjects were asked to indicate which of the following two statements they most agreed with: (i) *"One ought to help people such as Beatrix"* or (ii) *"People such as Beatrix ought to help themselves".* Using the probit regression analyhsis shown in table 4.6 we find that non-religious subjects are approximately 10% less likely to adhere to the helping norm than their religious counterparts (P>|z|=0.098).

Table 4.6 Probit analysis of determinants of helping norm

| | Adherence to helping norm |
|---|---|
| Treatment "free will disbelief" | -0.083 |
| | (0.065) |
| Free will scale | -0.002 |
| | (0.001) |
| Religious (1=yes) | 0.101* |
| | (0.061) |
| Controls | Yes |
| Observations | 108 |
| Pseudo-$R^2$ | 0.1857 |

Marginal effects of probit estimation with robust standard errors clustered on subjects in parentheses. The dependent variable is a dummy indicating whether the subject identified more with the statement "one ought to help poor people" (value=1) or with the statement "poor people ought to help themselves" (value=0). Controls include: age, sex and socio-economic status
*** p<0.01, ** p<0.05, * p<0.1

## 4.4 Discussion

With this experiment we aimed to explore how undermining belief in free will affects altruistic behavior in terms of charitable giving. On the basis of previous studies (Baumeister et al., 2009; Vohs and Schooler, 2008b) that found free will disbelief to be associated with reduced prosocial behavior we expected that undermining people's belief in free will would reduce the probability that subjects would select the more generous distribution. Our results indicate, however, that this was not the case. While our treatment did reduce belief in free will by 8.9 points on a 100-point scale, this did not significantly influence the likelihood of subjects selecting the more generous distribution. This null result is robust to controlling for sex, age, perceived socio-economic status, task characteristics and the order in which the decision tasks were presented.

Contrary to previous experiments on the effects of free will beliefs, we did not work with a sample of college students, but included a more diverse population. Whereas the average treatment effect was insignificant, our results indicate that the treatment did significantly reduce charitable giving among non-religious subjects. We considered several possible explanations for why religious people seem to be buffered from the treatment effect of our experiment. One possible explanation is that religious people are less open to scientific evidence, and thus less easily influenced by the free will disbelief treatment, which consisted of the presentation of neuroscientific evidence. However, this does not seem plausible as the treatment equally affected the belief in free will among both religious and non-religious subjects. Another explanation we explored is related to religion-based social norms. In line with previous studies that show an association between religious affiliation and the socialization of social norms promoting the helping of others (Batson et al., 1985; Benson et al., 1989; Ritzema, 1979; Youniss et al., 1999), our results indicate that religious subjects had a stronger adherence to the social norm of helping the poor. More strongly identifying with this norm

might buffer religious people against the effect of undermining belief of free will on charitable giving. Since our treatment was equally effective in reducing belief in free will amongst religious and non-religious people, we conclude that among religious subjects the negative effects of free will disbelief on charitable giving were cancelled out by their adherence to religion-based helping norms. The notion that religion-based helping norms affect giving behavior is further supported by our finding that only non-religious subjects' choices are influenced by the opportunity cost of the more altruistic option compared to the other option, whereas religious subjects' decisions are not influenced by this opportunity cost.

## *Limitations and strengths*

Several limitations apply to our study. Firstly, we have limited insight into the mechanisms by which religious affiliation might moderate the influence of free will beliefs. Although our data point towards the role of helping norms, we cannot exclude the possibility that other aspects of religious affiliation play a role. Future research in which both free will beliefs and the salience of religion-based helping norms are manipulated could shed more light on this. Compared to the main other study that used an experiment to investigate the effect of free will belief on altruistic behavior, by Baumeister and colleagues (Baumeister et al., 2009), our study has the advantage that it measures actual behavior rather than self-reported behavior in hypothetical scenarios. As previous research has indicated a significant bias in self-reported donation behavior (Bekkers and Wiepking, 2010), we think our study design offers a more reliable estimate of the effect of free will beliefs on charitable giving.

## *Practical implications and future research*

Our results warrant further caution for drawing the simplistic conclusion that a reduction of free will beliefs will automatically undermine pro-social

behaviors. For one, our null result suggests that the previously reported finding that undermining free will belief reduces pro-social behaviors reported might be more nuanced. Second, our finding that this effect only occurs among non-religious subjects suggests that beliefs regarding free will do not operate in isolation, but rather interact with pre-existing social norms and religious beliefs. As our sample of religious subjects consisted largely of Christians, it would be interesting to test the effect of free will belief manipulation among other subjects adhering to other religions. Furthermore, as free will disbelief can promote appreciation for the lack of control others have over their situation, another interesting avenue for future research would be to investigate how free will disbelief affects social preferences towards others that vary in perceived "helplessness", along the lines of an experiment where subjects could donate towards different kinds of welfare recipients that varied in their degree of perceived deservingness of welfare aid (Fong, 2007).

In sum, our study shows that undermining free will beliefs reduces charitable giving in binary dictator games, but that this effect only applies to non-religious subjects. Our results suggest that religion-based adherence to helping norms might interact with belief in free will and jointly shape altruistic behavior. Future research should shed further light on how belief in free will interacts with pre-existing social norms, such as the norm to help the poor.

# Appendix 4.1

Subjects were shown the following screens during the experiment which was conducted online through the Amazon mTurk platform:

**Page 1 (Welcome and informed consent)**

Principal Investigator: Dr. Jonathan Schooler, Department of Psychology, University of California, Santa Barbara Phone: (805) 893-5969

You have been asked to participate in a 2-part study, supervised by Dr. Jonathan Schooler, a psychology professor at the University of California at Santa Barbara. The first part of the study is about differences in mood following exposure to passages of text. You will then engage in a second study in which you can earn money. We will provide you with specific instructions prior to the starting the study.

The study will take approximately 10 minutes and you will receive $1 in exchange for your participation. Additionally, there is also the opportunity for you to receive up to an additional $1 based on the outcome of the second part of the study (Rewards will be granted through M-Turk's "Bonus" system at the same time the Hit is approved). You will only get an opportunity to receive this money if you have filled out all questions.

Your participation in this research will be kept strictly confidential. To preserve your anonymity, you will be provided with a Participant ID number. Any information that you provide will be available only to members of the research team for approved research purposes. If you feel uncomfortable at any time during the study please inform us immediately. Participation in this study is voluntary and you are free to discontinue your participation at any time.

If you decide not to participate, your refusal will involve no penalty and you will still receive full compensation. If you have any questions about this research or concerns about your participation, please contact the Human Subjects Committee at (805) 893-3807 or hsc@research.ucsb.edu. Or write to the University of California, Human Subjects Committee, Office of Research, Santa Barbara, CA 93106-2050

Participation in research is voluntary. Checking the box below will indicate that you have decided to participate as a research subject in the study described above. You may request an original and dated copy of this form to keep. Checking the box below indicates that you consent to participate in this study.

❍ By checking this box I agree that I have read and understood the terms and conditions above and certify that I am at least 18 years of age or older

**Page 2 (Happiness question)**

How happy do you feel at this very moment? Please indicate with the slider below your current level of happiness (0=very unhappy, 10=very happy)

_____ Current happiness

**Page 3 (Treatment group)**

We now ask you to read the following passage from the magazine "New Scientist" Please read it carefully as we will afterwards ask you to write a short summary of the text. If your summary doesn't relate to the text this might affect your payment.

---

New Scientist | Space | Technology | Environment | New Scientist Jobs | Subscribe to New Scientist

NewScientist SEARCH

NEWS | EXPLORE BY SUBJECT | SPECIAL REPORTS | LAST WORD | SUBSCRIBE | BLOGS | VIDEO | ARCHIVE | RSS

Print Send Feeds

# Neuroscientists Discover that Free Will is an Illusion

By Julia Reed
September 2, 2011

Your decision to read this article was entirely the product of your brain, which indicates to scientists that you do not have free will. In a study published this month in the *Journal of Cognitive Neuroscience*,[1] researchers using new brain scanning technology could see exactly which brain processes occurred as people made decisions. "We discovered that our decisions are caused entirely by complex brain processes," says lead researcher Dr. Peter Bernstein of Princeton University's Center for Neuroscience. "This shows that there is no free will."

Such discoveries seem to vindicate Dr. Francis Crick—a Nobel Prize winning scientist and co-discoverer of DNA—who once wrote: "Most people take free will for granted, since they feel that they are usually free to act as they please. Although we appear to have free will, in fact, our choices have already been predetermined for us by our brains and we cannot change that. Of course, myths such as free will seem only too plausible. Eventually, we will find the belief in free will every bit as outdated as the belief that the earth is flat."

Professor Bernstein's study used neuroimaging techniques pioneered by Dr. Martha White at Yale University's Center for Advanced Neuroscience. Participants in this experiment were students considering classes to take the next semester. They read descriptions of three psychology courses, considered reasons for and against each, and then pressed one of four buttons indicating their decision to take one of the courses (or none of them). They did all this while lying in a new type of functional Magnetic Resonance Imaging (fMRI) scanner, which measures where and when brain activity occurs, as well as the connections between specific brain processes. The researchers were able to measure how earlier brain processes, such as activity in the temporal striatum, predicted which course participants ultimately decided to take.

Dr. Bernstein explained, "Since their behavior was completely predicted by their brain processes, they clearly did not have free will about which course they selected. Their decisions were nothing more than the inevitable outcome of the processes we observed in their brain." This skeptical view about free will is now the dominant position among the world's leading scientists and philosophers. Dr. White concludes: "These studies confirm that our brains cause our decisions and then we consciously experience the outcome, much like a spectator observing a play. As scientists continue to demystify the mind by uncovering the neural mechanisms that drive our thoughts and behaviors, it has become increasingly clear that there is no role left for free will to play."

[1] Bernstein, P., Yin, R., Smith, J., White, M., and Snyder, H. 2011. "Mapping neural activity during a complex decision-making task." *Journal of Cognitive Neuroscience* 23(5): 1042-1051. | Article |

---

Please write a very short summary of the text (1-2 sentences):

**Page 3 (Control group)**

We now ask you to read the following passage from the magazine "New Scientist" Please read it carefully as we will afterwards ask you to write a short summary of the text. If your summary doesn't relate to the text this might affect your payment.

---


## Do we need to go nuclear to stay green?

Print Send Feeds

by Julia Grant

It's the billion-dollar question. Will nuclear energy save the world from global warming? Nuclear power plants produce virtually zero carbon emissions throughout their lifecycle, but they are costly to build and environmentalists claim the money would be better spent on building renewable resources.

However, increased energy demands and rising fossil fuel prices may persuade governments to adopt nuclear energy anyway, according to a new report published by the Nuclear Energy Agency in France, which considered three scenarios.

Firstly, if renewable energy sources and carbon-capture technology proves to be highly successful, and public acceptance of nuclear energy is low, ageing reactors will be replaced with newer reactors of a slightly higher capacity, but no additional power stations would be built. Overall, the NEA estimates that nuclear installed capacity would increase to 580 gigawatts of electrical energy worldwide by 2050, compared to just 370 GW in 2007.

In the second scenario, renewable supplies don't meet their expected capacity, so our reliance on nuclear energy will need to be even higher. In this case, nuclear energy would provide 1400 GW of electrical in 2050 – roughly 12.5% of the predicted energy demand.

Finally, the authors considered whether governments could phase out nuclear energy, but they believe it "unlikely" that other sources could completely fulfill the increased demands.

Please write a very short summary of the text (1-2 sentences):

**Page 4 (Happiness question)**
Please consider again how happy you feel at this very moment (0=very unhappy, 10=very happy)
_____ Current happiness

**Page 5 (Instructions to decision task)**
We will now proceed to a part of the study where you can earn additional money. You can earn up to 100 tokens, each token being worth 1 dollar cent. There are 24 decision-tasks. In each task you are asked to select 1 out of 2 possible allocations. By selecting an allocation, you decide how tokens are divided between you and another person. This other person is a poor person in a developing country, who will be selected by the charity "GiveDirectly". This charity directly sends cash to impoverished people in developing countries. An example of a person that is supported by GiveDirectly is Beatrice (31years) from Kenya, who has two young children and no income.

*[Here, subjects are shown a picture of Beatrice and her children]*

At the end of the study, 1 decision-task from the list of 24 decision-tasks will be randomly selected to be paid out in real money. Based on which allocation you selected in this choice, both you and/or the poor person will receive some amount of money. It is thus important that you pay close attention in all 24 choice-tasks! You have 10 seconds for each decision-task. If you fail to make a choice within this time neither you nor the poor person can receive any money in case this choice is selected.

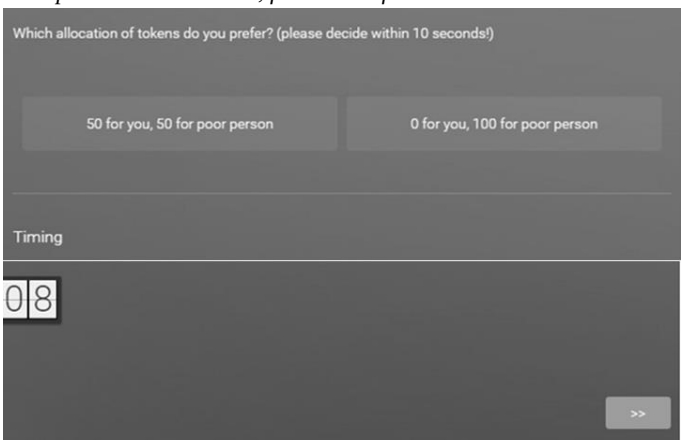
Before we begin with the decision-task please answer a few questions to confirm that you have understood the instructions. Your answers to these questions do not influence your payment.

Can you earn real money with the decision-tasks?
- ❍ *Yes*
- ❍ *No*

You will allocate money between yourself and another person. Who is this other person?
- ❍ *A poor person in a developing country*
- ❍ *Another m-Turk participant in this experiment*

**Page 6 (Decision tasks)**

*[Example screenshot below, full series of decision-tasks shown in table1*]




[*After these checks, subjects proceed to decision tasks 1-20*]

**Page 7 (Survey)**

What is your age?
- *18-29*
- *30-39*
- *40-49*
- *50-59*
- *60 and above*

Are you a male or female?
- *Male*
- *Female*

Do you consider yourself as belonging to any particular religion or denomination?
- *No*
- *Yes, Roman Catholic*
- *Yes, Protestant*
- *Yes, other Christian denomination*
- *Yes, Jewish*
- *Yes, Muslim*
- *Yes, Hindu*
- *Yes, Buddhist*
- *Yes, other, please specify:*

Please consider the following ladder. The bottom represents the least successful people in society, the top represents the most successful people. Where do you see yourself on this ladder (0=least successful, 10=most successful)?
*Position on ladder (0-10)*

To what extent you agree with the following statement: "I fully believe I have free will"
_____ *0=totally disagree, 100=totally agree*

When deciding how much to donate to GiveDirectly, did you consider whether Beatrice has any control over her lack of income?
- *Yes*
- *No*

Thinking about it now, how much control do you think Beatrice has over her lack of income?
_____ *0=no control, 100=complete control*

Which of the following two statements do you most agree with?
- *One ought to help people such as Beatrice*
- *People such as Beatrice ought to help themselves*

Chapter 5

# Choosing for Colleagues

## Experimental from the Field*


## Abstract

Decisions in small businesses, which commonly involve uncertainty, are often made on behalf of others. A growing literature shows that people exhibit different uncertainty preferences depending on whether they decide for themselves or for others. In this study, we investigate self-other differences in risk and ambiguity preferences among entrepreneurs in Bangladesh. We apply a between-subject design and use a vignette to elicit risk and ambiguity preferences in both self- and other-decisions. The results indicate lower ambiguity aversion in other-choices, but only for a sub-group of entrepreneurs that believe the ambiguous probability to be lower than fifty percent. Furthermore, we find that entrepreneurs are more risk averse when choosing for others. Finally, our results indicate that social distance does not drive self-other differences in ambiguity attitudes.


*Joint work with Haki Pamuk and Karen Maas.

## 5.1 Background

### *Ambiguous returns to business training*

Firms in developing countries are much less productive than their northern counterparts. This productivity difference is attributed in part to lower human capital (Bloom and Van Reenen, 2010). This human capital can be increased by investments in external advice and business trainings, which have been shown to promote firm productivity and employment creation both in developed countries (Robson and Bennett, 2000) and in developing countries (Bloom et al., 2013; Karlan and Valdivia, 2010; McKenzie and Woodruff, 2014). Consequently, many governments and donors have implemented subsidized business training and entrepreneurship programs for small firms in developing countries.

This begs the question why not more firms invest in business training? One factor that is likely to play a role here is that most small firms in developing countries have limited financial resources as well as limited access to credit (Beck and Demirgüç-Kunt, 2006). However, access to finance is likely not the only constraint. After all, if firms do not know the potential benefits of business training they might be unwilling to invest in it even if they have the financial resources.

In order to investigate this possibility, (Higuchi et al., 2015) conducted a field experiment with entrepreneurs in Vietnam. Their results show low demand for the business training prior to the random assignment of a business training program, but a significant increase in demand afterwards, through learning experiences from program participation. The authors suggest that this finding indicates that lacking knowledge about the potential benefits training benefits is one reason for low investments in business training. In other words, whereas the benefits of business training are ambiguous ex-ante, experience resolves – at least in part – this ambiguity, and consequently increases demand.

It is important to note here the difference between risk and ambiguity, where the latter involves probabilities and outcomes that are vague and cannot be

quantified (Knight, 2012). As first demonstrated through the famous "urn experiments" conducted by (Ellsberg, 1961) people typically prefer risky prospects the probabilities associated with payoffs can be quantified, to ambiguous prospects where these probabilities are unknown. For example, most people are ambiguity averse in the sense that they prefer a lottery with a fifty percent chance to win some prize to a lottery where the chance of winning this same prize is unknown. This conceptual distinction is also reflected in the fact that different parts of the brain are associated with processing risky and ambiguous prospects (Huettel et al., 2006).

There is an extensive literature on the role of risk preferences in business-related decisions such making investments, developing products and hiring employees (Burmeister-Lamp et al., 2012; Caliendo et al., 2009; Cramer et al., 2002; Ekelund et al., 2005; Elston and Audretsch, 2011; Forlani and Mullins, 2000; Gloss et al., 2017; Kan and Tsai, 2006). However, the role of ambiguity attitudes in such decisions has only become a topic investigation in empirical research more recently. Given that most business-related decisions involve probabilities that cannot be easily quantified a deeper understanding in the role of ambiguity attitudes can provide valuable insights in a range of questions about business decision-making.

For example, (Barham et al., 2014) show that among farmers in the U.S., those with higher levels of ambiguity aversion are more inclined to adopt genetically modified crops, which reduce the uncertainty associated with pest damages. Similarly, (Engle et al., 2011) show that among Peruvian farmers those with higher ambiguity aversion are less likely to diversify to new crops for which it was more difficult to assess the risk of harvest failure. However, it's not only entrepreneurs that dislike ambiguity, this aversion seems to be pretty similar between entrepreneurs and non-entrepreneurs. In a study among Chinese subjects, (Holm et al., 2013)find that whereas entrepreneurs are more tolerant towards uncertainty in strategic interactions, such as negotiating with another person, they

are equally ambiguity averse as non-entrepreneurs in non-strategic decision-tasks, such as choosing between risky and uncertain lottery. Similarly, (Koudstaal et al., 2015) study a population of entrepreneurs and non-entrepreneurs in the Netherlands and find no differences between these groups in terms of ambiguity aversion in non-strategic tasks.

These studies all consider ambiguity attitudes in choices that people make by themselves. However, there are also many ambiguous outcomes about which people in small firms make decisions on behalf of others. For example, consider a manager that has to decide on behalf of his team whether the company should invest in acquiring new knowledge, e.g. through a business training, of which the potential gains are uncertain.

In such other-decisions, the outcome is not merely a result of personal ambiguity attitudes, but also depends on the attitudes projected onto the other person on whose behalf the decision is made. Furthermore, since other-decisions involve psychological distance between the decision-maker and the subject of the decision, it has been argued that this mode of decision-maker promotes more a more abstract and generalized mode of thought (Polman, 2010; Trope and Liberman, 2010). Indeed, these differences between self- and other-decisions are reflected in a number of studies.

## *Deciding for others*

For example, (Zikmund-Fisher et al., 2006) ask subjects to consider various active medical treatments and passive strategies of non-intervention, either from the perspective of the patient (i.e. the self-perspective) or from the perspective of a doctor (i.e. the other-perspective). They find that when choosing from the self-perspective people display are more willing to accept risks that result from non-intervention than risks resulting from treatment, whereas in the other-perspective this bias is less prevalent. In a similar fashion (Polman, 2012) and (Andersson et

al., 2013) find that subjects display lower levels of aversion - the tendency to overweight the negative impact of losses relative to the positive impact of equivalent gains - when making decisions on behalf of others.

Differences in self-other decision have also been observed in the domain of decision-making under uncertainty. (Pollmann et al., 2014) find that people are less risk averse when they make investments for others than for themselves. They also find that this self-other difference in risk aversion disappears when the decision-maker can be held accountable by a reward to be made by the principal, the owner of the money, after the agent, the investor of the money, made the investment. A similar pattern of reduced risk aversion in other decisions is observed in other studies (Chakravarty et al., 2011; Polman, 2012). Two other studies that also make use of a similar investment task however find increased risk aversion when choosing for others (Füllbrunn and Luhan, 2015; Kvaløy et al., 2014). Although it remains unclear how these seemingly contradictory results can be reconciled, there is a growing literature pointing to important self-other differences in various domains of decision making.

Less is known about potential self-other differences in decision making under ambiguity. To our best knowledge, the only paper that considers this question was published by (König-Kersting and Trautmann, 2016a). In contrast to the literature on self-other differences in the domain of risk, the authors do not find significant differences in ambiguity attitudes between self- and other-decisions. This results stands in contrast with (Charness and Gneezy, 2010) who find lower ambiguity aversion in group decisions as compared to individual decisions, suggesting that there may be a normative appeal to ambiguity neutrality. However, the results of (König-Kersting and Trautmann, 2016a) does correspond to earlier work by (Trautmann et al., 2008) which shows that being observed by others results in increased ambiguity aversion. In sum, it remains unclear under which conditions ambiguity attitudes differ between self- and other-decisions, and how this may influence business-related decisions.

### *Research question*

This study builds on the following two results described in the previous section. First, small firms' demand for business training and other technologies with uncertain benefits is limited by ambiguity aversion. Second, it remains unclear under which conditions people exhibit self-other differences in choices that involve risk and uncertainty. We thus seek to answer the following question:

- Is demand for business training with ambiguous returns higher when people choose on behalf of others than when they choose for themselves?

## 5.2 Methods

### *Setting*

The experiment was conducted among a population of owners, executive and high level managers of the registered businesses from information technology (IT) and information technology enabled services (ITES) sector in Dhaka, Bangladesh. The experiment was embedded in a larger study about an international development program which provided business support to small and medium sized enterprises in various countries. Our subjects were selected from the businesses that did not participate or applied to this business support program.[20] From a total of 861 entrepreneurs that were called to arrange an interview for the survey, 441 participated in the survey, and out of 441 survey participants 371 subjects participated in the experiment. The majority of subjects are males, with on average 6.5 years of work experience after they were graduated from school. The largest group consisted of owner-managers (38.8%) with the second and third largest

[20] We excluded from our experiment those firms that participated or applied to- and/or participated in the business support program since these firms might have confused our experiment with the actual business trainings offered to them in the consulting program. These businesses constitute about 10 percent of total population.

groups being owners (30.7%), and executive managers (24.8%). Finally, another 5.7% of respondents had different types of managerial functions in their company.

Table 5.5.1 Respondent characteristics

| | | Obs | Mean | S.D. | Min | Max |
|---|---|---|---|---|---|---|
| Respondent gender (1=male) | | 371 | 0.96 | 0.2 | 0 | 1 |
| Tenure (years) | | 371 | 7.24 | 5.38 | 0 | 41 |
| Position | Owner / exec. manager | 371 | 0.39 | 0.49 | 0 | 1 |
| | Owner | 371 | 0.31 | 0.46 | 0 | 1 |
| | Exec. manager | 371 | 0.25 | 0.43 | 0 | 1 |
| | Manager, other | 371 | 0.06 | 0.23 | 0 | 1 |

## *Procedure*

The experiment was implemented by Bangladeshi enumerators in the local language Bengali. These enumerators were trained by the research team. Each enumerator visited subjects in April-May 2016 at their offices to conduct a face to face interview that took about an hour. Each enumerator first directed questions to the respondents to collect information about subject and business characteristics, and then passed to the experiment section. In our experiment we were especially concerned about religious tensions for questions involving uncertain outcomes in a predominantly Muslim country. In the beginning of the experiment section, each enumerator therefore first read an informed consent script prepared by research team and asked subjects if they would like to answer questions about uncertain outcomes. After getting the consent from survey participants to participate in the experiment, they started to implement the experiment by explaining the decision task in the experiment.[21]

[21] Out of 440 respondents to the surveys 371 respondents accepted to participate in the experiment.

## *Decision task*

To measure attitudes towards risk and uncertainty we used a survey instrument based on a double multiple price list (DMPL) instruments. This instrument consists of two lists. The first list consists of choices between a certain and an ambiguous option. The second list consists of choices between a certain and a risky option (see figure 2). Given that attitudes towards risk and ambiguity have been found to be positively correlated (Abdellaoui et al., 2011; Bossaerts et al., 2010; Charness and Gneezy, 2010)[22] this instrument allows for the estimation of ambiguity attitudes net of risk attitudes.

In contrast to commonly used DMPL instruments in which choices are financially incentivized, our experiment did not use financial incentives. The reason of this was twofold. First we wanted to construct an ecologically valid scenario pertaining to real-life business decisions. Second, we were worried that subjects may think that risk and ambiguity game involve gambling and this might have created religious tensions among the subjects, the majority of whom are Muslim,[23] for these reasons, the choices in the prize list were embedded in a script about hypothetical business trainings.

Subjects were asked to imagine they are asked by their director that they are given a choice between participation in two different trainings, both of which have the same characteristics (i.e. automatic passing of training, one-day duration, in their city of residence). It is explained to them that the payoffs associated with both training programs differ.

[22] For a comprehensive review of the literature on ambiguity and risk attitudes see (Trautmann and van de Kuilen, 2015)

[23] Our concerns are similar to those experienced in other field studies, for example (Callen et al., 2014) who elicit risk preferences in Afghanistan - another predominantly Muslim society -, using a DMPL similar to ours, without financial incentives.

Table 5.2 Decision tasks

| *Price List 1: Ambiguity Attitude* | | |
|---|---|---|
| Task | Training A | Training B |
| 1 | ?? chance to get 50K salary increase | 100% chance to get 5K salary increase |
| 2 | ?? chance to get 50K salary increase | 100% chance to get 10K salary increase |
| 3 | ?? chance to get 50K salary increase | 100% chance to get 15K salary increase |
| 4 | ?? chance to get 50K salary increase | 100% chance to get 20K salary increase |
| 5 | ?? chance to get 50K salary increase | 100% chance to get 25K salary increase |
| 6 | ?? chance to get 50K salary increase | 100% chance to get 30K salary increase |
| 7 | ?? chance to get 50K salary increase | 100% chance to get 35K salary increase |
| 8 | ?? chance to get 50K salary increase | 100% chance to get 40K salary increase |
| 9 | ?? chance to get 50K salary increase | 100% chance to get 45K salary increase |
| 10 | ?? chance to get 50K salary increase | 100% chance to get 50K salary increase |
| *Price List 2: Risk Attitude* | | |
| Task | Training C | Training B |
| 1 | 50% chance to get 50K salary increase | 100% chance to get 5K salary increase |
| 2 | 50% chance to get 50K salary increase | 100% chance to get 10K salary increase |
| 3 | 50% chance to get 50K salary increase | 100% chance to get 15K salary increase |
| 4 | 50% chance to get 50K salary increase | 100% chance to get 20K salary increase |
| 5 | 50% chance to get 50K salary increase | 100% chance to get 25K salary increase |
| 6 | 50% chance to get 50K salary increase | 100% chance to get 30K salary increase |
| 7 | 50% chance to get 50K salary increase | 100% chance to get 35K salary increase |
| 8 | 50% chance to get 50K salary increase | 100% chance to get 40K salary increase |
| 9 | 50% chance to get 50K salary increase | 100% chance to get 45K salary increase |
| 10 | 50% chance to get 50K salary increase | 100% chance to get 50K salary increase |

In the first choice list (certainty vs. ambiguity) it is explained that choosing for training A will result in an increase of monthly salary of 50.000 Bangladeshi Taka (approximately 600€). It is further explained that the probability of this salary increase actually occurring is unknown, and it depends on an uncertain prospect that the firm wins a contract for a project. In contrast, participation in training B - which prepares for work on a project that has already been secured - will result in a certain increase in monthly salary, varying from 10-50.000 BDT. The second choice list has the same structure, but the uncertain probability of a 50K payoff to training A is replaced with a 50% probability of this

payoff. In both these choice lists, the point at which subjects switch from the option with ambiguous/risky payoff, i.e. training A, to the option with certain payoff is used as an indicator of their ambiguity and risk attitudes. For example, consider a subject who prefers in the first price list training A for tasks 1-2 and switches to a preference for training B in tasks 2-9. This subject, as well as and any other subjects that switch to Training B before task 4 can be classified as ambiguity averse, since he/she only prefers the uncertain option if the expected payoff is higher than the certain payoff in option B. Similarly, risk attitudes can be derived from the switchpoint in the second choice list

## *Treatments*

We randomly allocated a business code from 0 to 860 for each firm in our population of businesses before enumerators arranged the interviews with the firms and according to the firm business code each firm was assigned to one of the following treatment groups:

1. Self if the firms' business code ends on a "1", "4", "7" or "0";
2. Colleague/friend if the firms business code ends on a "2", "5", or "8";
3. Colleague the firms' business code ends on a "3 ", "6", or "9".

In the first group, subjects were asked to imagine they made the choice for themselves. The second group was asked to imagine making the choices for a colleague with whom they were also closely befriended (low social distance). The third group was asked to imagine making the choices for a new colleague whom they did not know very well yet (high social distance). The exact wording of these instructions can be found in annex 1. The main focus in the paper is on comparing decisions in the self-condition (i.e. group 1) with the other-condition (i.e. group 2 and 3).

We will compare groups 2 and 3 to explore whether ambiguity attitudes are moderated by the social distance towards the subject of the decision. Subject population of businesses were randomly assigned to one of the three experimental conditions, A total of 345 subjects were assigned to the self-condition, 258 subjects to the colleague low social distance condition and 258 subjects to the colleague high social distance condition. Randomization was implemented with the random generation function of the Microsoft Excel.

As can be seen in Table 3, of 371 subjects 159 subjects participated in the self-condition, 112 subjects participated in the colleague low social distance condition and 100 subjects participated in the colleague high social distance condition.

Table 5.3 Randomization balance

| | Self | | | Low social distance | | | High social distance | | |
|---|---|---|---|---|---|---|---|---|---|
| | Obs. | Mean | S.D. | Obs. | Mean | S.D. | Obs. | Mean | S.D. |
| Male (1=yes) | 159 | 0.96 | 0.21 | 112 | 0.97 | 0.16 | 100 | 0.95 | 0.22 |
| Tenure (years) | 100 | 6.60 | 4.82 | 83 | 6.10 | 4.76 | 73 | 6.37 | 4.49 |
| Position | | | | | | | | | |
| Owner, exec. man. | 159 | 0.39 | 0.49 | 112 | 0.38 | 0.49 | 100 | 0.39 | 0.49 |
| Owner | 159 | 0.37 | 0.48 | 112 | 0.26 | 0.44 | 100 | 0.26 | 0.44 |
| Exec. Manager | 159 | 0.18 | 0.38 | 112 | 0.30 | 0.46 | 100 | 0.30 | 0.46 |
| Manager , other | 159 | 0.06 | 0.24 | 112 | 0.05 | 0.23 | 100 | 0.05 | 0.22 |

The randomization resulted in partial balance. In terms of gender and tenure the randomization was balanced. In terms of position in the firm the randomization was not fully balanced. Compared to the other-treatment (i.e. groups 2 and 3) the fraction of subjects that were firm owner but not executive manager was higher in the self-treatment, whereas the fraction of subjects that were executive manager, but not owner, was higher in the other-treatment. We control for imbalances by controlling for firm position (see next section).

## *Identification*

We estimate all models introduced in the previous paragraphs through OLS regression and report heteroscedasticity robust standard errors. To investigate whether entrepreneurs risk attitudes are different for self and other decisions we estimate following model and test $\alpha_1 \neq 0$:

$$Risk_i = \alpha_{0+} \alpha_1 Other_i + \alpha'_2 X_i + u_i \quad (3.1)$$

where $Risk_i$ is the inverse of the switchpoint in the second choice list for subject i that takes values from 1 and 10. As such this variable indicates the degree of risk aversion. $Other_i$ equals to 1 (0 otherwise) if the participant is either in the low social distance (befriended colleague) or high social distance (colleague) group. If the entrepreneur deciding for others is more (less) risk averse than entrepreneurs deciding for themselves, then $\alpha_1 > 0$ ($\alpha_1 < 0$). $X_i$ is a vector of subject characteristics including gender, tenure (years since graduation of highest completed education), position in the firm, and reported level of concern when choosing for others.

Second, to probe whether entrepreneurs' ambiguity attitudes are different for self and other decisions we estimate following model and test $\beta_1 \neq 0$:

$$Ambiguity_i = \beta_{0+} \beta_1 Other_i + \beta_2 Risk_i + \beta'_3 X_i + \varepsilon_i \quad (3.2)$$

where $Ambiguity_i$ is the inverse of the switch-point in the second choice list for subject i that takes values from 1 and 10. As such this variable indicates the degree of ambiguity aversion. If the entrepreneur deciding for others is more ambiguity averse than subjects when he/she decides for others than when he/she decides for him/herself, then $\beta_1 > 0$. To isolate ambiguity attitude from risk attitude we control for risk attitude by adding $Risk_i$. Again we control of subject characteristics through $X_i$. Subjects might have formed different beliefs about the ambiguous

prospect for training A in the first choice list: the uncertain probability that the firm wins a contract for a project through which training A leads to a salary increase. Do subjects have different risk and ambiguity attitudes for self and other decision when they have different believes about the payoff probability in this ambiguous prospect? To answer this, after the experiment, we asked subjects whether they believed the uncertain probability of winning a contract for the project was either smaller, equal or greater than 50%.

Through this we construct three sub-groups for subjects with smaller, equal or greater than 50% believes about uncertain outcome. We then estimate (3.2) for those three sub-groups and compare $\beta_1$ estimates among groups. The rationale for this question is to allow us to control how beliefs about ambiguous probabilities might relate to self-other differences in ambiguity attitudes. The intuition here is that if subjects believe that the ambiguous prospect occurs with a probability of exactly 0.5, their attitude towards this option should be identical to that in the risk task. In contrast, if subjects believe the probability to not be 0.50 we suspect that self-other differences might be distinct from those observed in the risk task. Furthermore, this question links to the causes and implications of overconfidence among entrepreneurs (Forbes, 2005).

Does increasing social distance mediate the difference between self and other decisions for trainings? Are entrepreneurs' risk and ambiguity attitudes different when they decide for socially close friends and when they decide for socially distant other colleagues? To answer these questions we reformulate equations (3.1) and (3.2) as follows, estimate this model for two of the groups, namely (i) subjects deciding for socially distance colleagues and (ii) subjects deciding for socially closer colleagues, and test $\lambda_1 \neq 0$ for risk attitudes and $\pi_1 \neq 0$ for ambiguity attitudes:

$$Risk_i = \lambda_0 + \lambda_1 CloseC_i + \lambda'_2 X_i + \varepsilon_i \quad (3.4),$$

$$Ambiguity_i = \pi_0 + \pi_1 CloseC_i + \pi_2 Risk_i + \pi'_3 X_i + \varepsilon_i \quad (3.5)$$

where $CloseC_i$ equals 1 if the participant is deciding on behalf of a socially closer colleague and equals to 0 if deciding on behalf of a socially distant colleague. If entrepreneurs are more risk averse (lover) when they decide for socially close colleague than when they decide for socially distant colleagues, then $\lambda_1<0$ ($\lambda_1>0$). If entrepreneurs are more ambiguity-averse (ambiguity-loving) when they decide for colleagues in the lower social distance condition, then $\pi_1<0$ ($\pi_1>0$).

## 5.3 Results

### *Self-other decisions*

We then turn to regression results. Our main results regarding self-other decisions for risk and ambiguity attitudes are summarized in Table 4. Column 1 of the Table includes the coefficient estimates from the estimation of equation (3.1) where Risk aversion is the dependent variable. Columns 2-5 of the Table include coefficient estimates from the estimation of equation (3.2) where *Ambiguity* is the dependent variable. Columns 2-5 respectively report coefficient estimates for all entrepreneurs and three sub-groups including subjects with smaller, equal or greater than 50% believes about uncertain outcome of the project. We control for gender, tenure, position and concern for others.

Several results are reflected in table 5-4. Firstly, as indicated in column 1, we find that entrepreneurs are more risk averse when they decide for others than when they decide for themselves. Second, as indicated in column 2, we find no strong evidence that – on average – ambiguity attitudes differ significantly between self and other decisions for entrepreneurs. In this column, the estimate for *Other* decisions is negative but not statistically significant. Furthermore, and in line with the literature (Abdellaoui et al., 2011; Bossaerts et al., 2010; Charness and Gneezy, 2010) we find that risk aversion – as indicated by the coefficient on the variable "switchpoint choice list 2" is a strong predictor of ambiguity aversion.

Table 5.4 Self-other differences in risk and ambiguity attitudes

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dependent variable | Risk aversion | Ambiguity aversion | Ambiguity aversion | Ambiguity aversion | Ambiguity aversion |
| Other | 0.570* | -0.285 | -0.779*** | 0.414 | -0.224 |
| | (0.308) | (0.200) | (0.258) | (0.289) | (0.395) |
| Risk aversion | | 0.682*** | 0.507*** | 0.695*** | 0.723*** |
| | | (0.041) | (0.059) | (0.100) | (0.068) |
| Constant | 6.885*** | 2.847*** | 5.002*** | 2.318** | 1.871** |
| | (0.584) | (0.511) | (0.598) | (0.940) | (0.944) |
| Observations | 359 | 359 | 137 | 83 | 131 |
| R-squared | 0.024 | 0.501 | 0.384 | 0.719 | 0.474 |
| Sample | Full sample | Full sample | Belief payoff<0.5 | Belief payoff=0.5 | Belief payoff>0.5 |
| Controls | Yes | Yes | Yes | Yes | Yes |

OLS regression of other-treatments (other friend/colleague and other colleague), comparing with self-treatment. Control variables: gender, tenure, position and concern for others. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

However, a more nuanced picture emerges when we test self-other differences in ambiguity attitudes according to subjects' beliefs about the payoff probability in the ambiguous prospect. Columns 3 to 5 of the Table show estimates for group of subjects that believed the payoff probability in the ambiguous prospect to be less than 50%, equals to 50%, and more than 50%. We find that subjects that believed the payoff probability in the ambiguous prospect to be less than 50% are less ambiguity averse when they choose for others than when they decide for themselves. This finding corresponds with our hypothesis of lower ambiguity aversion for other-decisions. We find no strong evidence of difference between ambiguity attitudes in self and other decisions regarding decisions for subjects that believed the payoff probability in the ambiguous prospect to be equal or more than 50%. These results imply that depending on how entrepreneurs interpret the ambiguous prospect and its pay-off, they make different choices.

## *Social Distance*

We then turn to explore if perceived social distance towards the other differently influences ambiguity attitudes in the colleague/friend vs. colleague condition. We test whether entrepreneurs are less ambiguity averse when deciding for socially closer colleagues compared to socially distant colleagues. Our main results regarding social distance and risk and ambiguity attitudes are summarized below.

Table 5.5 Social distance and risk and ambiguity attitudes

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dependent variable | Risk aversion | Ambiguity aversion | Ambiguity aversion | Ambiguity aversion | Ambiguity aversion |
| Close colleague | 0.232 | -0.280 | -0.690 | 0.332 | -0.108 |
| | (0.397) | (0.284) | (0.429) | (0.488) | (0.554) |
| Risk aversion | | 0.668*** | 0.610*** | 0.589*** | 0.703*** |
| | | (0.055) | (0.059) | (0.198) | (0.095) |
| Constant | 7.538*** | 2.759*** | 3.482*** | 2.233 | 2.212* |
| | (0.750) | (0.636) | (0.597) | (1.718) | (1.239) |
| Observations | 207 | 207 | 84 | 43 | 72 |
| R-squared | 0.019 | 0.455 | 0.407 | 0.586 | 0.436 |
| Sample | Full sample | Full sample | Belief payoff<0.5 | Belief payoff=0.5 | Belief payoff>0.5 |
| Controls | Yes | Yes | Yes | Yes | Yes |

OLS regression of other-colleague/friend treatment, comparing with other-colleague treatment. Analysis excludes the subjects in the self-treatment group. Control variables: gender, tenure, position and concern for others. Robust standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1

Column 1 of table 5-5 includes the coefficient estimates from the estimation of equation (3.3) where Risk is the dependent variable. Columns 2-5 of the Table show the coefficient estimates from the estimation of equation (3.4) where *Ambiguity* is the dependent variable. Columns 2-5 respectively reports coefficient estimates for all entrepreneurs and three sub-groups including subjects with smaller, equal or greater than 50% believes about uncertain outcome of the

project. We again control for gender, tenure, position and perceived social distance to others in all estimates.

In sum, we find no evidence that self-other differences in risk and ambiguity attitudes are moderated by social distance, neither for the full sample (columns 1 and 2) nor for sub-groups specified according to beliefs about payoff probabilities (columns 3-5).

## 5.4 Discussion

In this study with owners and managers of small firms in the IT-sector in Bangladesh we investigate if people are more willing to invest in business training with uncertain returns if they make such decisions on behalf of others in their company. In doing so, we aimed to contribute to the understanding of the role of ambiguity attitudes and self-other decisions in demand for new technologies with uncertain returns among small firms in developing countries (Engle-Warnick et al., 2007) Furthermore, given that small and medium sized firms constitute the largest part of the private sector in most developing countries (Ayyagari et al., 2007) our study aims to contribute to the understanding of the mechanisms of economic growth in developing countries.

Given the mixed findings on self-other differences in previous lab experiments on decision making under uncertainty (Chakravarty et al., 2011; Füllbrunn and Luhan, 2015; König-Kersting and Trautmann, 2016a; Kvaløy et al., 2014; Pollmann et al., 2014; Polman, 2012) we asked the question whether owners and managers in our sample would be more willing to invest in business trainings if the uncertain benefits would accrue to a colleague.

Our results indicate that this is only partly the case. On the one hand we find that – for the full sample – ambiguity attitudes do not differ between self- and other-decisions. This finding aligns with a recent result obtained by (König-Kersting and Trautmann, 2016a) who study ambiguity attitudes in an incentivized

lab experiment with students and find no differences between self- and other-decisions. However, when accounting for beliefs about ambiguous probabilities, we find a self-other difference in ambiguity attitudes to occur among those who believe that payoff probability in the ambiguous prospect is lower than 50%. For this sub-group, constituting approximately 40% of the sample, ambiguity aversion is lower when choosing for others.

In contrast, we find a higher level of risk aversion in other-choices. This finding is in line with previous studies that use incentivized tasks to measure risk attitudes in self- and other-decisions
(Füllbrunn and Luhan, 2015; Kvaløy et al., 2014; Reynolds et al., 2011), but differs from various others studies that find opposite results (Chakravarty et al., 2011; Pollmann et al., 2014; Polman, 2012). Taken together, these results correspond to the notion that risk and ambiguity are distinct phenomena, as reflected in the fact that different brain areas are involved in processing known vs. unknown probabilities (Huettel et al., 2006).

Furthermore, we find that social distance between the agent and the subject of the decision, in our case either a befriended colleague or a new colleague, does not predict the level of ambiguity aversion. This result is in contrast to our expectation that self-other differences in ambiguity aversion are mediated by social distance. A potential reason for this finding is that because of the hypothetical nature of our survey instruments, subjects were simply not able to "imagine" different degrees of social distance depending on the specifications of the hypothetical subjects for whom they were deciding.

## *Limitations and strengths*

To the best of our knowledge, our study is the first experiment to investigate self-other differences in ambiguity outside of lab a setting. Contrary to most previous studies on self-other differences that have been conducted with populations of

students in a laboratory context, our experiment specifically focuses on a group of entrepreneurs and managers in the IT-sector in Bangladesh. As such, our study contributes to a growing literature that takes standard behavioral economics experiments outside of the convenience populations in university labs to subjects from different countries, cultures and socio-economic backgrounds (Henrich et al., 2001, 2010c; Herrmann et al., 2008). Furthermore, our study contributes to the literature on the role of ambiguity attitudes in the context of entrepreneurship. Various lines of research have shown that higher ambiguity tolerance can be beneficial for firms, including better coping with stress faced by entrepreneurs (Hai Yap Teoh and See Liang Foo, 1997) and higher willingness to adopt technologies of which the potential benefit is ambiguous, such as business training and improved crop varieties (Barham et al., 2014; Engle-Warnick et al., 2007; Higuchi et al., 2015). Whilst experiments on ambiguity attitudes commonly make use of financial incentives, we chose in our study to use vignettes instead. This approach has two main limitations. Firstly, there are no actual consequences – financial or otherwise – resulting from the subjects' decision. Secondly, we rely on subjects' capacity to imagine different levels of social distance for the colleague/friend vs. the colleague treatment. However, hypothetical scenarios have been used in numerous studies on ambiguity attitudes (Fox and Tversky, 1995; Ritov and Baron, 1990; Trautmann et al., 2008). A key benefit of using a vignette is that it allows for the construction of a decision-task with higher ecological validity and relevance to the entrepreneurial context in which our experiment was implemented.

Another limitation of our study is that in the other-condition it cannot be fully ruled out that the decision also can have indirect economic consequences on decision-maker, other than direct psychological effects such as warm glow (Andreoni, 1990) that the decision-maker might derive from doing what they think will help the other person. For example, if the colleague benefits from having followed a training which was selected for them, it could be that they would feel

inclined to reciprocate and return the favour. The best way to exclude such effects would to use a fully anonymized lab experiment. However, this approach would again limit the ecological validity which we deem relevant for the study of ambiguity attitudes in an entrepreneurial context. Furthermore, it was made very explicit in our instructions for subjects in the other-treatment that the primary benefits – in the form of a salary increase – would accrue only to the colleague and not to the decision-maker.

## *Policy implications and future research*

The main policy implications from this study is that in situations where ambiguity aversion can constrain demand for new services and technologies, such as for example in the case of business trainings (Higuchi et al., 2015) technology uptake may be promoted if decisions regarding its uptake are made not by the main beneficiary, such as the director, but rather by someone deciding on behalf of their behalf, such as the manager.

In order to further explore how delegated decision-making in small firms can counter the tendency to underinvest in inputs with ambiguous returns, further research is warranted. Firstly, it would be valuable to further investigate whether our result holds under conditions in which there are real financial incentives at stake. Recently published work by (König-Kersting and Trautmann, 2016a) making use of a controlled lab experiment with incentivized choice tasks, finds no significant self-other differences in ambiguity attitudes. However, this study does not control for beliefs with respect to the ambiguous prospect. It would be interesting to study in future incentivized experiments whether the self-other differences in ambiguity attitudes do manifest for sub-sets of subjects that hold particular beliefs regarding the uncertain probability.

Another direction for future research is to further explore the role of social distance in explaining self-other differences in ambiguity aversion. Whereas our

study failed to find such a relation, a more controlled manipulation of social distance between the decision-maker and the subject of the decision might serve to provide a more precise answer to this question.

In sum, our experiments shows that entrepreneurs in Bangladesh display a lower level of ambiguity aversion when making decisions on behalf of colleagues compared to making such decisions for themselves. Our results do not support our expectation that social distance moderates this effect. Future research could investigate in more detail the mechanisms by which self-other differences in ambiguity aversion operate and influence entrepreneurial decisions.

# Appendix 5.1 Instructions

*Enumerator read the text:* Now we would like to ask you some hypothetical questions about choices under uncertainty. The research team will use this questions will be used to assess the demand for managerial training services from different support organizations

**H.0**. Do you want to answer those questions?

- ❍ Yes (1)
- ❍ No (2) Skip to Section I

**For enumerator**

- ***if the firm ID code of the firm ends on*** *a "1", "4", or "7" or "0" we ask you to ask GROUP 1,*
- *If the firm ID-code ends on a "2", "5", or "8" we ask you to go to GROUP 2*
- *If the firm ID-code ends on a "3", "6", or "9" we ask you to go to GROUP 3*

# Group 1: Self

Choice 1 (CERTAINTY vs. AMBIGUITY)

Please imagine that you are asked by your company to participate in one of the following two management trainings offered by a training company. Both trainings take place in Dhaka, last one day. All participants will pass the training automatically. Depending on which training you follow, you will be qualified to work on different projects, each of which will be awarded by your company with a different level of salary increase.

*Training A*

- Employees that follow training A will be invited to work on project GREEN.
- Your company is currently trying to obtain the contract for project GREEN
    - The chance that your firm will obtain this contract is unknown. For example, it could be that the chance is only 10%, or it could be 50%, or it could be 90% etc. To summarize, the chance that your firm will obtain the contact is unknown.
- Your company promised that all employees that follow training A can work on project GREEN & will receive a 50.000BDT salary increase, *if your company gets the contract for this project.* If your company does not get the contract your salary will stay the same.

*Training B*

- Employees that follow training B will be invited to work on project BLUE.
- Your company recently obtained the contract for project BLUE.
    - The chance that your firm will obtain this contract is 100%
- Your company promised that all employees that follow training B can work on project BLUE & will receive a 5.000BDT salary increase

→ Please indicate which training you would choose? (A or B)

| **Q.** | **Training A** | **Training B** | **Choice (A or B )** |
|---|---|---|---|
| 1 | ?? chance to get 50K BDT salary increase | 100% chance to get 5K BDT salary increase | |

→ Now imagine the salary increase for training B is not 5K BDT but 10K. For training A it is constant like before. Now which training would you choose? (A or B)

| **Q.** | **Training A** | **Training B** | **Choice (A or B )** |
|---|---|---|---|
| 2 | ?? chance to get 50K BDT salary increase | 100% chance to get 10K BDT salary increase | |
| 3 | ?? chance to get 50K BDT salary increase | 100% chance to get 15K BDT salary increase | |
| 4 | ?? chance to get 50K BDT salary increase | 100% chance to get 20K BDT salary increase | |
| 5 | ?? chance to get 50K BDT salary increase | 100% chance to get 25K BDT salary increase | |
| 6 | ?? chance to get 50K BDT salary increase | 100% chance to get 30K BDT salary increase | |
| 7 | ?? chance to get 50K BDT salary increase | 100% chance to get 35K BDT salary increase | |
| 8 | ?? chance to get 50K BDT salary increase | 100% chance to get 40K BDT salary increase | |
| 9 | ?? chance to get 50K BDT salary increase | 100% chance to get 45K BDT salary increase | |
| 10 | ?? chance to get 50K BDT salary increase | 100% chance to get 50K BDT salary increase | |

Choice 2 (CERTAINTY vs. RISK)

Please imagine that you are asked by your company to participate in one of the following two management trainings offered by a training company. Both trainings take place in Dhaka, last one day. All participants will pass the training automatically. Depending on which training you follow, you will be qualified to work on different projects, each of which will be awarded by your company with a different level of salary increase.

*Training C*

- Employees that follow training C will be invited to work on project ORANGE.
- Your company is currently trying to obtain the contract for project ORANGE
    - The chance that your firm will obtain this contract is 50%.
- Your company promised that all employees that follow training C can work on project ORANGE & will receive a 50.000 BDT salary increase, *if your company gets the contract for this project.* If your company does not get the contract your salary will stay the same.

*Training B*

- Employees that follow training B will be invited to work on project BLUE.
- Your company recently obtained the contract for project BLUE.
    - The chance that your firm will obtain this contract is 100%.
- Your company promised that all employees that follow training B can work on project BLUE & will receive a 5.000 BDT salary increase

→ Please indicate which training you would choose? (C or B)

| **Q.** | **Training C** | **Training B** | **Choice (C or B )** |
|---|---|---|---|
| 1 | 50% chance to get 50K BDT salary increase | 100% chance to get 5K BDT salary increase | |

→ Now imagine the salary increase for training B is not 5K BDT but 10K. For training A it is constant like before. Now which training would you choose? (A or B)

| **Q.** | **Training C** | **Training B** | **Choice (C or B )** |
|---|---|---|---|
| 2 | 50% chance to get 50K BDT salary increase | 100% chance to get 10K BDT salary increase | |
| 3 | 50% chance to get 50K BDT salary increase | 100% chance to get 15K BDT salary increase | |
| 4 | 50% chance to get 50K BDT salary increase | 100% chance to get 20K BDT salary increase | |
| 5 | 50% chance to get 50K BDT salary increase | 100% chance to get 25K BDT salary increase | |
| 6 | 50% chance to get 50K BDT salary increase | 100% chance to get 30K BDT salary increase | |
| 7 | 50% chance to get 50K BDT salary increase | 100% chance to get 35K BDT salary increase | |
| 8 | 50% chance to get 50K BDT salary increase | 100% chance to get 40K BDT salary increase | |
| 9 | 50% chance to get 50K BDT salary increase | 100% chance to get 45K BDT salary increase | |
| 10 | 50% chance to get 50K BDT salary increase | 100% chance to get 50K BDT salary increase | |

## Group 2: Colleague/Friend

Choice 1 (CERTAINTY vs. AMBIGUITY)

Please imagine that *another employee in your company, who is also a close friend of yours and have the same monthly income with you*, is asked by your company to participate in one of the following two management trainings offered by a training company. Both trainings take place in Dhaka, last one day. All participants will pass the training automatically. Depending on which training your colleague will follow, he/she will be qualified to work on different projects, each of which will be awarded by your company with a different level of salary increase.

*Training A*

- Employees that follow training A will be invited to work on project GREEN.
- Your company is currently trying to obtain the contract for project GREEN
    - The chance that your firm will obtain this contract is unknown. For example, it could be that the chance is only 10%, or it could be 50%, or it could be 90% etc. To summarize, the chance that your firm will obtain the contact is unknown.
- Your company promised that all employees that follow training A can work on project GREEN & will receive a 50.000BDT salary increase, *if your company gets the contract for this project.* If your company does not get the contract your friend's salary will stay the same.

*Training B*

- Employees that follow training A will be invited to work on project BLUE.
- Your company recently obtained the contract for project BLUE.
    - The chance that your firm will obtain this contract is 100%
- Your company promised that all employees that follow training A can work on project BLUE & will receive a 5.000BDT salary increase

→ Please indicate which training you would choose? (A or B)

| **Q.** | **Training A** | **Training B** | **Choice (A or B )** |
|---|---|---|---|
| 1 | ?? chance to get 50K BDT salary increase | 100% chance to get 5K BDT salary increase | |

→ Now imagine the salary increase for training B is not 5K BDT but 10K. For training A it is constant like before. Now which training would you choose? (A or B)

| **Q.** | **Training A** | **Training B** | **Choice (A or B )** |
|---|---|---|---|
| 2 | ?? chance to get 50K BDT salary increase | 100% chance to get 10K BDT salary increase | |
| 3 | ?? chance to get 50K BDT salary increase | 100% chance to get 15K BDT salary increase | |
| 4 | ?? chance to get 50K BDT salary increase | 100% chance to get 20K BDT salary increase | |
| 5 | ?? chance to get 50K BDT salary increase | 100% chance to get 25K BDT salary increase | |
| 6 | ?? chance to get 50K BDT salary increase | 100% chance to get 30K BDT salary increase | |
| 7 | ?? chance to get 50K BDT salary increase | 100% chance to get 35K BDT salary increase | |
| 8 | ?? chance to get 50K BDT salary increase | 100% chance to get 40K BDT salary increase | |
| 9 | ?? chance to get 50K BDT salary increase | 100% chance to get 45K BDT salary increase | |
| 10 | ?? chance to get 50K BDT salary increase | 100% chance to get 50K BDT salary increase | |

Choice 2 (CERTAINTY vs. RISK)

Please imagine that *an employee in your company, who is also a close friend of yours and have the same monthly income with you*, is asked by your company to participate in one of the following two management trainings offered by a training company. Both trainings take place in Dhaka, last one day. All participants will pass the training automatically. Depending on which training your colleague will follow, he/she will be qualified to work on different projects, each of which will be awarded by your company with a different level of salary increase.

*Training C*

- Employees that follow training C will be invited to work on project ORANGE.
- Your company is currently trying to obtain the contract for project ORANGE
    - The chance that your firm will obtain this contract is 50%.
- Your company promised that all employees that follow training C can work on project ORANGE & will receive a 50.000 BDT salary increase, *if your company gets the contract for this project.* If your company does not get the contract your friend's salary will stay the same.

*Training B*

- Employees that follow training C will be invited to work on project BLUE.
- Your company recently obtained the contract for project BLUE.
    - The chance that your firm will obtain this contract is 100%.
- Your company promised that all employees that follow training B can work on project BLUE & will receive a 5.000 BDT salary increase

→ Please indicate which training you would choose? (C or B)

| **Q.** | **Training C** | **Training B** | **Choice (C or B )** |
|---|---|---|---|
| 1 | 50% chance to get 50K BDT salary increase | 100% chance to get 5K BDT salary increase | |

→ Now imagine the salary increase for training B is not 5K BDT but 10K. For training A it is constant like before. Now which training would you choose? (A or B)

| **Q.** | **Training C** | **Training B** | **Choice (C or B )** |
|---|---|---|---|
| 2 | 50% chance to get 50K BDT salary increase | 100% chance to get 10K BDT salary increase | |
| 3 | 50% chance to get 50K BDT salary increase | 100% chance to get 15K BDT salary increase | |
| 4 | 50% chance to get 50K BDT salary increase | 100% chance to get 20K BDT salary increase | |
| 5 | 50% chance to get 50K BDT salary increase | 100% chance to get 25K BDT salary increase | |
| 6 | 50% chance to get 50K BDT salary increase | 100% chance to get 30K BDT salary increase | |
| 7 | 50% chance to get 50K BDT salary increase | 100% chance to get 35K BDT salary increase | |
| 8 | 50% chance to get 50K BDT salary increase | 100% chance to get 40K BDT salary increase | |
| 9 | 50% chance to get 50K BDT salary increase | 100% chance to get 45K BDT salary increase | |
| 10 | 50% chance to get 50K BDT salary increase | 100% chance to get 50K BDT salary increase | |

## Group 3: Colleague

Choice 1 (CERTAINTY vs. AMBIGUITY)

Please imagine that *another employee in your company, whom recently joined the company, whom you don't know so well yet, and who have the same monthly income with you*, is asked by your company to participate in one of the following two management trainings offered by a training company. Both trainings take place in Dhaka, last one day. All participants will pass the training automatically. Depending on which training your colleague will follow, he/she will be qualified to work on different projects, each of which will be awarded by your company with a different level of salary increase.

*Training A*

- Employees that follow training A will be invited to work on project GREEN.
- Your company is currently trying to obtain the contract for project GREEN
    - The chance that your firm will obtain this contract is unknown. For example, it could be that the chance is only 10%, or it could be 50%, or it could be 90% etc. To summarize, the chance that your firm will obtain the contact is unknown.
- Your company promised that all employees that follow training A can work on project GREEN & will receive a 50.000BDT salary increase, *if your company gets the contract for this project.* If your company does not get the contract your colleague's salary will stay the same.

*Training B*

- Employees that follow training B will be invited to work on project BLUE.
- Your company recently obtained the contract for project BLUE.
    - The chance that your firm will obtain this contract is 100%
- Your company promised that all employees that follow training B can work on project BLUE & will receive a 5.000BDT salary increase

→ Please indicate which training you would choose? (A or B)

| **Q.** | **Training A** | **Training B** | **Choice (A or B )** |
|---|---|---|---|
| 1 | ?? chance to get 50K BDT salary increase | 100% chance to get 5K BDT salary increase | |

→ Now imagine the salary increase for training B is not 5K BDT but 10K. For training A it is constant like before. Now which training would you choose? (A or B)

| **Q.** | **Training A** | **Training B** | **Choice (A or B )** |
|---|---|---|---|
| 2 | ?? chance to get 50K BDT salary increase | 100% chance to get 10K BDT salary increase | |
| 3 | ?? chance to get 50K BDT salary increase | 100% chance to get 15K BDT salary increase | |
| 4 | ?? chance to get 50K BDT salary increase | 100% chance to get 20K BDT salary increase | |
| 5 | ?? chance to get 50K BDT salary increase | 100% chance to get 25K BDT salary increase | |
| 6 | ?? chance to get 50K BDT salary increase | 100% chance to get 30K BDT salary increase | |
| 7 | ?? chance to get 50K BDT salary increase | 100% chance to get 35K BDT salary increase | |
| 8 | ?? chance to get 50K BDT salary increase | 100% chance to get 40K BDT salary increase | |
| 9 | ?? chance to get 50K BDT salary increase | 100% chance to get 45K BDT salary increase | |
| 10 | ?? chance to get 50K BDT salary increase | 100% chance to get 50K BDT salary increase | |

Choice 2 (CERTAINTY vs. RISK)

Please imagine that *another employee in your company, whom recently joined the company, whom you don't know so well yet , and who have the same monthly income with you*, is asked by your company to participate in one of the following two management trainings offered by a training company. Both trainings take place in Dhaka, last one day. All participants will pass the training automatically. Depending on which training your colleague will follow, he/she will be qualified to work on different projects, each of which will be awarded by your company with a different level of salary increase.

*Training C*

- Employees that follow training C will be invited to work on project ORANGE.
- Your company is currently trying to obtain the contract for project ORANGE
    - The chance that your firm will obtain this contract is 50%.
- Your company promised that all employees that follow training C can work on project ORANGE & will receive a 50.000 BDT salary increase, *if your company gets the contract for this project.* If your company does not get the contract your colleague's salary will stay the same.

*Training B*

- Employees that follow training B will be invited to work on project BLUE.
- Your company recently obtained the contract for project BLUE.
    - The chance that your firm will obtain this contract is 100%.
- Your company promised that all employees that follow training B can work on project BLUE & will receive a 5.000BDT salary increase

→ Please indicate which training you would choose? (C or B)

| **Q.** | **Training C** | **Training B** | **Choice (C or B )** |
|---|---|---|---|
| 1 | 50% chance to get 50K BDT salary increase | 100% chance to get 5K BDT salary increase | |

→ Now imagine the salary increase for training B is not 5K BDT but 10K. For training A it is constant like before. Now which training would you choose? (A or B)

| **Q.** | **Training C** | **Training B** | **Choice (C or B )** |
|---|---|---|---|
| 2 | 50% chance to get 50K BDT salary increase | 100% chance to get 10K BDT salary increase | |
| 3 | 50% chance to get 50K BDT salary increase | 100% chance to get 15K BDT salary increase | |
| 4 | 50% chance to get 50K BDT salary increase | 100% chance to get 20K BDT salary increase | |
| 5 | 50% chance to get 50K BDT salary increase | 100% chance to get 25K BDT salary increase | |
| 6 | 50% chance to get 50K BDT salary increase | 100% chance to get 30K BDT salary increase | |
| 7 | 50% chance to get 50K BDT salary increase | 100% chance to get 35K BDT salary increase | |
| 8 | 50% chance to get 50K BDT salary increase | 100% chance to get 40K BDT salary increase | |
| 9 | 50% chance to get 50K BDT salary increase | 100% chance to get 45K BDT salary increase | |
| 10 | 50% chance to get 50K BDT salary increase | 100% chance to get 50K BDT salary increase | |

Chapter 6

# Choice Complexity, Benchmarks and Costly Information*


## Abstract

We conducted a lab experiment to investigate the effect of information interventions on decision-making in complex choices. Choice options were framed as financial products and could be objectively ranked, but time constraints made this impractical for subjects, forcing them to use more efficient heuristics. In our benchmark treatments, one option was revealed as having average values for all attributes, with attribute values presented either in relative or absolute terms. In our costly information treatment, two options were revealed as being suboptimal. We find that costly information and relative benchmarks improve decision quality. Finally, benchmarks have limited effects on demand for advice.


*Joint work with Mark Sanders and Stephanie Rosenkranz.

## 6.1 Introduction

Many choices are complex. This complexity reflects not just the number of options, but also the number of attributes per option[24]. Such complex choices occur for example when consumers have to choose between mortgage products, financial retirement plans, health insurance policies, or mobile phone subscriptions.

Rational choice theory assumes that it is better to have more rather than fewer options. Empirical evidence, however, suggests that people have difficulties in selecting a suitable product when choices are complex.[25] Recent studies found that complex choices result in choice inertia and reduced decision quality (Besedeš et al. 2012, 2010; Heiss et al. 2013; Huberman et al. 2007; Friesen and Earl 2015; Greifeneder et al. 2010), as well as lower satisfaction with the decision (Iyengar and Kamenica, 2006; Iyengar and Lepper, 2000).

In this context, decisions concerning complex financial products are of specific interest, as mistakes in such decisions can be very costly for the individual as well as for society. Financial literacy seems to be key to financial well-being: financially literate individuals make fewer mistakes and are in better financial condition than financial illiterates (Lusardi and Mitchell, 2011)[26]. However, many individuals generally score poorly on financial literacy, and it remains an open question what aids would effectively help these individuals make better decisions. Where increasing financial literacy may be first best, it is doubtful if this can ever materialize and be sufficient. Experimental research, however, has shown that in

[1] Choice complexity is defined as the amount of information a choice involves: a choice between objects with one or two important attributes is simple, whereas a choice between objects for which many attributes are important is complex (Dijksterhuis et al., 2006)

[25] See Chernev et al. (2015) for an overview on empirical studies. In an experimental study, Besedeš et al. (2010) find that specifically for older subjects the probability of a person selecting the optimal option declines in the number of options, and that older subjects rely more on suboptimal decision rules.

[26] Many people have limited capacity to interpret numerical information and low numeracy is shown to be associated with suboptimal financial decisions (Kirsch and And Others, 1993; Reyna and Brainerd, 2007).

redesigning the choice architecture, decision quality can be significantly improved without removing options (Besedeš et al., 2010, 2012; Heiss et al., 2013). Building on this literature, the present paper provides novel evidence about a specific type of decision aid in complex choices; benchmarks.

There seems to be an emerging literature testing interventions in choice architectures, but the empirical evidence on the effectiveness of such decision aids in complex choice situations remains scarce. Besedeš et al. (2015) study two forms of choice architecture which reduce a large decision problem into a series of smaller ones without reducing the choice set (Thaler and Sunstein, 2003). The authors show that a tournament-style choice architecture, in which a large choice set is broken down in several smaller choice sets from which respective optimal options are selected into a final choice set, reduces choice overload and thereby improves decision making.

Presenting subjects with different choice architectures Besedeš et al. (2015b) find that subjects' preferences for choice architectures are negatively correlated with performance. This suggests that providing choice over architectures might reduce the quality of decisions. The authors cannot exclude that this problematic result is related to self-sorting, suggesting benevolent regulators should make that choice for them.

Samek et al. (2016) test a somewhat different approach to promote improved decision-making in complex choices. The authors provide subjects with different presentation formats of a matrix representing the attribute values of different products. In one treatment, subjects could sort values per attribute (column), and in the second treatment all attributes were automatically ranked in descending order. Compared to a baseline where values are not and cannot be sorted, the second treatment caused a significant improved in decision quality and reduced decision time. In a related fashion, Peters et al. (2009) show that difficult-to-evaluate attributes are taken into account more by decision makers when graphical decision aids make it easier to map these attributes on a good/bad scale.

And Agnew and Szykman (2005) find that subjects with higher levels of financial knowledge are less likely to suffer from choice overload when information about various products is presented in a table than when it is presented in a booklet format.

In addition to these rather stylized architecture interventions, Soll et al. (2013) investigated the effect of the so-called "Credit Card Accountability Responsibility and Disclosure" (CARD) act, which forces providers to give information on the relationship between repayment amounts and loan duration in a prescribed format that makes the information easier to understand. In an online survey experiment the authors then found that this information indeed improves decision quality (it reduced people's bias towards underestimating the duration of repayment periods).

With the present paper we focus on two other approaches to improve decision making in complex choices; benchmarks and advice. In our individual choice task subjects make a sequence of 10 choices among four different multi-attribute options, one of which being objectively optimal. Choices are incentivized and framed as a choice among financial products, and are made under a time constraint to mimic complexity[27]. One treatment variable is the provision of a benchmark. In the benchmark treatments we label one choice option as the "average product", for which we use two different frames: in one frame the attributes of all other products are presented in absolute positive or negative deviations from the average product. In the other frame attributes for each option

[27] The ranking of options does not depend on subjects' risk preferences and but on the assumption that they are money maximizers. The full choice set is clearly defined, as is the value of each option. While the optimal option is always unique, its identity is concealed from subjects by manipulating the value of 5 attributes that need to be combined to calculate the value of each option. Complex decisions are typically made without a time constraint, but are characterized by a large number of alternatives, parameters, variables, and uncertainties, such that it is usually difficult to make the optimal decision. Setting a time constraint makes it difficult for subjects to make the calculations that are needed to find the optimal option with certainty. Several studies how that decision-making under time pressure indeed leads to suboptimal decision making (Ben Zur and Breznitz, 1981; Payne et al., 1988).

are presented as relative deviations from the average product, for which all attributes are set to 100[28]. In all treatments we generated three products with random attributes and computed the attributes of the fourth as the average of the three randomly generated. This randomization allows us to control for product similarity[29].

Providing a benchmark option with average values for each attribute, changes the presentation of decision-relevant information without altering the actual choice set. With this treatment we change the presentation of the choice attributes, by explicitly providing a benchmark as a reference for comparison[30]. Rational decision making in complex choices over options with multiple attributes requires a comprehensive evaluation strategy. Often a high value for one attribute can compensate for a low value for another, as for example in the case of a phone subscription that consists of various costs components such as the co-payment for the phone, fixed monthly costs, flexible monthly costs, etc. In the face of constraints to time and cognitive capacity, decision-makers might benefit from a well-known benchmark or reference product as it makes it easier to compare the available options. By varying the way the information is presented and by identifying a benchmark, the decision-maker is steered towards using an elimination-by-aspects strategy.

Alternatively, a decision-maker may be triggered to substitute a comprehensive evaluation strategy by simpler alternative strategies (weighted additive, satisficing, lexicographic), or by (e.g. fast and frugal) heuristics relying

[28] Note, strictly speaking we thus reduce the available information in the relative benchmark treatment.

[29] Note that in this choice experiment the decision makers were asked to select the optimal product and pay-off depended on the rank of the product they selected, not how close or far the selected product was from the optimal choice. Very similar products make this task more complex as the optimal product is harder to spot in the short time given.

[30] Note that we do not claim that the benchmark product serves as a reference point in the sense of(Kahneman and Tversky, 1982). Our benchmark product is part of the choice set (which does not need to be the case for a reference point) and does not necessarily determine clear domains of gains and losses.

on only a few comparisons[31]. This strategy may be particularly effective for subjects who are less financially literate, i.e. individuals with higher cognitive costs. In line with the notion that presentation of information influences how easily it can be processed (Bettman and Kakkar, 1977), we conjecture that the relative benchmark treatment will be more effective, as it steers the decision-maker towards using a heuristic of comparing products in terms of the number of attributes with above- and below-average values.

We contribute to the literature on decision aids with our finding that benchmarks promote decision quality, but only when attribute values of the products are expressed in relative terms. Our second treatment consists of subjects being given the option to buy information (which we label "advice") about the relative profitability of various options that can be chosen. When advice is bought, the computer truthfully marks two of the suboptimal options, the worst and randomly one of the other non-optimal options, in the set of four. This information is costly and can be ignored but subjects are informed the advice is truthful. When the decision is taken randomly, the two scenarios with and without costly information differ only in the variance not in the expected value of the consequence of the decision, with the variance when information is acquired being lower.

Providing the possibility to discard sub-optimal options from the choice set reduces the cognitive burden of the decision maker because fewer options need to be evaluated. With this treatment we target the structure of the choice task as we vary the number of relevant alternatives. However, we present the decision maker with the possibility to receive additional information on suboptimal options at the expense of the benefits of selecting the optimal option. The effectiveness of this approach relies on the assumption that people who request the additional information are benefited by it, and those who do not are benefited by the original

[31] Regarding strategies in complex choices see (Thaler et al., 2014), (Gigerenzer et al., 1999) or for an overview (Payne et al., 1988).

choice set.[32] A rational decision-maker following a comprehensive evaluation strategy should only make use of the advice option when the reduction in marginal cognitive costs - due to reduced difficulty of the decision task – more than compensates for the reduction in marginal benefits. The advice option is thus attractive for subjects who are less financially literate or who are susceptible to choice overload and thus have higher marginal cognitive costs, as this increases especially their chance to choose the optimal option.

Our paper therefore also contributes to the literature on the effects of advice on decision quality. In an experimental study Gino and Moore (2007) show that demand for costly advice in complex choice tasks sub-optimally high while costly advice is underutilized in simple choice tasks. In a hypothetical choice experiment Hung and Yoong (2010) compare the effect of unsolicited and solicited advice. They find that unsolicited advice does not affect investment behavior, but when advice is optional, individuals with low financial literacy are more likely to ask for it and use it. The authors also find that notwithstanding this negative selection on ability, individuals who actively solicit advice indeed make better choices. Hackethal et al. (2012) find that self-selection largely explains their finding of better outcomes for advisees in the context of German Internet brokerage accounts. A robust finding in the literature is that individuals who receive unsolicited advice tend to significantly discount it (Bonaccio and Dalal, 2006; Yaniv, 2004a, 2004b; Yaniv and Kleinberger, 2000). Furthermore, Gino (2008) shows that individuals are more likely to use decision-related information when they pay for this information compared to when they get it for free. In sum, while advice that is explicitly solicited and paid for is perceived as helpful, unsolicited and free advice is perceived as intrusive and might even lead to worse decisions (Deelstra et al., 2003; Goldsmith, 2000; Goldsmith and Fitch, 1997). Our

[32] Note that providing costless information regarding the ranking of options, i.e. without affecting consequences, would simplify the decision problem for all decision makers in a trivial way without allowing us to differentiate between decision makers applying different choice strategies.

results indicate that access to advice improves decision making. Not only does it result in more optimal decisions and fewer suboptimal decisions, it also leads to a higher payoff even net of the advice cost. Furthermore, we find that subjects benefit more from the option to buy advice when the products presented are more similar and therefore harder to distinguish. Then simplifying the choice set to two options and reducing the cognitive effort needed for optimal decision making has greater marginal effect.

We conduct the experiment with a 3 × 2 factorial design (relative, absolute, no benchmark × advice, no advice), allowing us to also study possible interaction effects of benchmarks and advice. We are especially interested in understanding whether these two approaches indeed improve decision making in complex choices, and whether the instruments are complements or rather substitutes. Thus, our study also contributes to the literature on the demand for financial advice. Various studies show that demand for financial advice, such as consultation of a bank advisor, is positively correlated to individuals' level of financial literacy, even when controlling for income and education levels (Calcagno and Monticone, 2015; Hackethal et al., 2012). One interpretation of this result is that individuals with higher financial literacy better understand the potential benefits of seeking costly advice, for example to avoid even higher costs resulting from poor financial decisions (Robb et al., 2012). In contrast, Gino and Moore (2007) find that subjects are not more likely to seek even costless advice in a difficult version of the task compared to an easy version of the task.

We find that absolute but not relative benchmarks increase the propensity to buy advice. This result corresponds with the finding that only relative benchmarks improve decision quality and suggest that absolute benchmarks were harder for subjects to use, thus driving increased demand for costly advice. This paper proceeds as follows. Section 2 describes the experimental design, and procedures. Section 3 develops the hypotheses we test in our study and our

empirical strategy. Our results are presented in Section 4 and we discuss our findings and conclude in Section 5.

## 6.2 Experimental design

### *Design*

The experiment consisted of three parts. The first part of the experiment was an individual choice task, where subjects made choices among different options, which were framed as financial products. Subjects were presented a table with four different options, labelled product A, B, C and D, and were instructed and incentivized to select the option with the lowest total 'costs'. To mimic decision making for complex financial products, the total costs of the products were not explicitly given, but the subjects were presented with five different cost elements, framed as costs and tax deduction, of each product. Subjects were informed that the product had a maturity of one year (12 months) and that it was the subjects' task to determine which product has the lowest total costs. The optimal product could be calculated using the formula:

Total cost = start costs + (12 x monthly costs) + maturity costs + management fee (percentage of start costs) - tax deduction (percentage of monthly cost)

This formula was not given explicitly, but in the instructions all cost elements and their influence on the total costs were carefully explained (see the instructions in Appendix II). The values for the cost elements were randomly generated for three of the products while for the fourth product the cost elements were calculated as the average of the three randomly generated others. The intervals in which the cost elements were randomly varied were displayed on a whiteboard in the room. This design allows for an objective evaluation and ranking of options, independent of subjects' tastes and risk preferences as long as

subjects are not satiated in money (see (Besedeš et al., 2012b). The products and an example of cost elements were presented as in Table 6.1.

Table 6.1 Example Payoff Matrix

| | Product | | | |
|---|---|---|---|---|
| | A | B | C | D |
| Starting costs | 87 | 92 | 103 | 94 |
| Monthly costs | 35 | 49 | 64 | 49 |
| Maturity costs | 72 | 91 | 2 | 52 |
| Management fee (%) | 15 | 31 | 16 | 21 |
| Tax deduction (%) | 11 | 10 | 10 | 10 |

The task of choosing the optimal product was repeated ten times, and each time the cost elements for three of the products were randomly determined. The position of the average product within the table was randomly assigned by the computer every round. The treatment to which subjects were assigned determined whether they were informed about the existence of this average product or not. Subjects had to perform the task of choosing the optimal product within 30-seconds, which were presented by a timer counting down in the upper right corner of the screen. This time limit was introduced to simulate the complexity of the financial product choice. If a subject did not choose a product within these 30 seconds, the computer automatically implemented the product with the highest total costs as the subject's choice.

We tested the effects of two benchmark treatments and one advice treatment, in a 3 × 2 factorial design. All treatments were employed in a between-subject design. Regarding the benchmark we varied the information subjects had on the existence of the average product as well as the presentation of this average product. In the control treatment, subjects were not informed that the cost elements of one of the options reflected the respective average of the other three options. This information was also not revealed to them in the instructions. This average product was indicated by a blue font and it was also explicitly stated that the

respective product is an average product. In one set of treatments the other products were represented in absolute deviations from the average product (ABSOLUTE) while in another treatment (RELATIVE) the cost elements of all other products were presented in percentages relative to the average product. In RELATIVE the average product had all cost components set to 100 while cost components of other products were expressed relative to this.

We varied the information regarding the ranking of the options by giving subjects in one set of treatments (ADVICE) the opportunity to have the worst option and another randomly chosen suboptimal option indicated by the computer before making a choice, against the payment of a fixed price. While this potentially left the subject with only two options, the advice could be ignored and the indicated suboptimal options could still be chosen. We framed this decision as 'buying advice' and the price for this information was set such that the decision was only profitable if indeed the optimal option was selected. The choice on whether or not to purchase this information was to be made in each round before the options were presented.

Payoffs are shown in Table 2. The subjects were rewarded based on how well they made their decisions. The options were ranked in order from optimal (cheapest) to worst (most expensive). Payoffs were directly linked to whether the optimal, second best, second worst or worst product was chosen. Subjects started with an initial endowment of €8 and the payment was added or deducted from this amount depending on the choices the subject made. Choosing the second best option after buying advice led to a payoff of zero, while choosing a worse product led to deductions from the endowment. Note that it was made clear that the worst option and one of the remaining suboptimal options would be indicated by the advice, such that it was not clear to subjects if the remaining worst option would be the second or third best option. For purely random choices the expected value

of buying advice is thus 0.5*5+0,5*(0,5*-2.5+0,5*2.5) – 2.5 = 0; identical to the expected value of not buying advice.[33]

Table 6.2 Payoffs by choice

| Choice | No advice | Advice (cost=2.5) |
|---|---|---|
| Optimal | 5 | 2.5 |
| 2nd best | 2.5 | 0 |
| 3rd best | -2.5 | -5 |
| Worst | -5 | -7.5 |

At the end of the experiment one of the ten rounds was randomly drawn and used to determine the subjects' actual monetary earnings. Subjects earned on average €9.16 in this part of the experiment. In the second part of the experiment, to assess participants' risk aversion, a sequence of binary lottery decisions was administered, which was equivalent to the one introduced by (Holt and Laury, 2002). The task asked the subjects to choose between two lotteries in ten different cases. The lottery choice screen is shown in Appendix II, Figure A7. The amounts that could be won did not change in the ten lotteries, only the probability of occurrence of each amount changed. One lottery paid either €3 or €0 and the other paid either €1.50 or €1. The switching point from one lottery to the other is the crucial point that reveals the individual's risk aversion. Even when people switch back and forth between lotteries, the number of safe options gives a good indication of the subject's level of risk aversion (Holt and Laury, 2002). After ten decisions were made, one of the ten lotteries was chosen randomly by the computer and played to determine the payoff for the subject in this round. The payoff received in this part of the experiment was added to the amount the subject earned in the first part. Subjects earned on average €1.74 in the lottery part.

[33] Note that a risk-averse decision-maker employing such a random choice strategy should acquire information as this reduces the variance of the expected outcome.

Note that we deliberately abstracted from heterogeneity in subjects' tastes. Of course this limits the external validity of our findings. However, this allowed us to implement a benchmark with an objective value that is independent of subjects' tastes and therefore increase internal validity by abstracting from subjective beliefs. The same argument holds for the absence of risk in the choices in our treatments. The possibility to objectively rank options allows us to observe decision quality without having to control for risk preferences. In the real world it would perhaps be more suitable to set benchmarks that satisfy the preferences of the average consumer. In our simplified choice situation with homogeneous tastes such a benchmark could obviously not be implemented. Nevertheless, we think it is important to study the psychological effects of the presence of a benchmark product in the simplest setting first. Finally, in the real world, complex decisions often have no time constraint, but rather require the assessment of uncertain absolute and relative merits of multiple attributes of available options. In order to be able to objectively determine decision quality in our lab setting we presented the options with attributes that were relatively simple to assess. However, adding a time constraint made it difficult for subjects to process all relevant information and possibly forced them to concentrate on the most important attributes. This evaluation strategy is similar to strategies employed in real complex choices (Payne et al., 1988).

The experiment ended with a questionnaire on demographics, buying behavior, self-reported personality traits (BFI-10(Rammstedt and John, 2007) and attitudes regarding financial products (AFM, 2014). See Appendix III for these survey questions. The survey was not incentivized but anonymity was guaranteed.

## *Experimental procedures*

The experiments were conducted in the Experimental Laboratory for Sociology and Economics (ELSE) at Utrecht University. They were programmed and

conducted with the experimental software 'z-Tree'(Fischbacher, 2007). In seven sessions, a total of 158 subjects (average of 23 subjects per session) participated in the experiment.

The subjects were mostly undergraduate students from various fields at Utrecht University and Hogeschool Utrecht. Over 1000 potential subjects from the pool of the ELSE lab were approached by email to participate in the experiment, using the ORSEE recruitment system (Greiner, 2015). Upon arrival in the lab, subjects were randomly assigned a seat behind a computer. Subjects were randomly divided into 3 groups, where each of the three groups played a different treatment[34]. Treatments with and without the option to buy advice were administered in different sessions but on the same days in alternating order. This allows for comparison of the differences between the treatments and control for all other factors as these stay the same between treatments. Before the start of every experiment, general written instructions in English were given, which were kept identical across sessions (see Appendix II). Additional instructions were displayed on the screen. The first part of the experiment started when all subjects had fully read and understood the instructions. One full experimental session lasted on average 45 minutes and subjects earned an average of €10.90.

## 6.3 Hypotheses and estimation strategy

### *Hypotheses*

Decisions in complex choices can be sub-optimal because of boundaries to cognitive abilities, decision time and/or tractability of the decision problem. As such, interventions that alleviate these constraints may improve decision quality. Our benchmark interventions, which provide subjects with information about the

[34] For groups 1,3 and 5 some an IT-related problem caused in a reduced number of tasks being implemented in periods 9-10 of the experiment (10 missing obs./group for period 9 and 20 missing obs. for period 10). We have controlled for these missing observations by testing our regression models both with and without these periods and the results do not change significantly, We thus report the estimates for the large dataset including rounds 9-10.

way in which attributes compare across products, is expected to improve decision quality by making the decision problem more tractable.

> *Hypothesis 1:* In the treatments with absolute and relative benchmarks, the number of optimal decisions is higher, the number of worst decisions is lower, and average payoffs are higher than in the treatments without a benchmark.

Next, we expect that benchmarks are particularly effective at improving decision quality when attribute values are presented in *relative* terms compared to the "average product" rather than in *absolute* terms. The main reason for this is that the relative benchmark makes it easier for subjects to see in one glance if product attributes have a value above (>100) or below (<100) market average.

> *Hypothesis 2:* In the treatments with a relative benchmark, the number of optimal decisions is higher, the number of suboptimal decisions is lower, and average payoffs are higher than in the treatments with an absolute benchmark.

Regarding the effect of the costly information treatment, we expect that advice will be bought mostly by subjects that find the decision problem particularly challenging and that this will consequently result in improved decision making:

> *Hypothesis 3:* In the treatments with the possibility to buy additional information (advice), the number of optimal decisions is higher, the number of suboptimal decisions is lower, and payoffs are higher than in the treatments without this option.

Finally, regarding the effect of benchmark products on demand for costly information, the existing literature does not provide us with a clear direction of the expected effect. However, if indeed benchmarks make decision problems more

tractable and given that advice is costly, we would expect subjects to have lower demand for advice in the benchmark treatments.

> *Hypothesis 4:* In the treatments with absolute and relative benchmarks there is lower demand for costly information than in treatments without a benchmark.

## *Estimation strategy*

We test these hypotheses using a series of-OLS regression analyses. Regarding hypotheses 1-3, we first test how payoffs – both gross and net – are influenced by the five treatment combinations (i) relative benchmark, (ii) absolute benchmark, (iii) advice, (iv) relative benchmark * advice, (v) absolute benchmark * advice. To this purpose we estimate the following regression model:

$$Y_{ij} = \beta_0 + \beta_1 A_{ij} + \beta_2 BM_{ij} + \beta_3 BM_{ij}*A_{ij} + \beta_4 C_{ij} + \varepsilon_{ij} \qquad (3.1)$$

in which $Y_{ij,}$ represents the payoff individual *i* in decision-task *j*, $BM_{ij}$ is a dummy for the benchmark treatment, $A_{ij}$ is a dummy for the advice treatment, $BM_{ij}*A_{ij}$ is the interaction of the benchmark and advice treatment, $\varepsilon$ are unobserved factors and $\beta_{0,}\, \beta_{1,}\, \beta_2$ are parameters to be estimated. Estimates of the parameter $\beta_2$ can be interpreted as the causal effect of benchmark treatment, which is randomly assigned to individuals. In addition, we control for the possibility of non-random assignment of treatments to subjects by including a vector of subject-level control variables ($C_{ij}$): age, gender, self-reported preferences with respect to financial products and self-reported psychological traits from the survey and risk aversion as measured with the Holt-Laury lottery task. As we expect that subjects require some practice to get acquainted with the benchmark treatments we test for learning effects. To this purpose, we run separate regressions for the first part (round 1-6) and second part (round 7-10) of the experiment.

Having tested the effect of the treatment on payoffs, we then turn to exploring how the treatments affect the actual decision. To this purpose, we apply a probit model with same specification as 3.1. This model is estimated for the following binary outcome variables: (i) *best choice*: if subject selects the most optimal, i.e. least expensive, product; (ii) w*orst choice:* if subject selects the least optimal, i.e. most expensive, product; (iii) *no choice:* if subjects select no product within the allocated 30 seconds. In order to test hypothesis 2 – that relative benchmarks are more effective than absolute benchmarks – we run the same regression analyses to the sub-sample of subjects in either of these two treatments. By comparing the relative to the absolute benchmark treatment we can identify to what extent possible effects derive from the benchmark as such vs. from the format in which the information was presented. Finally, in order to test hypothesis 4 – that demand for costly information is lower in the benchmark treatments – we implement the same regression as specified in 3.1 while restricting the sample to the subjects in either of the two benchmark treatments. Setting the absolute benchmark as the reference category, we then estimate the effect of the relative benchmark compared to the absolute benchmark.

## 6.4 Results

### *Descriptive statistics*

Table 3 describes the demographics of the sample across all sessions and treatments. There are no statistically significant differences between groups, indicating that randomization was successful.

Table 6.3 Demographics by treatment

| Group | # obs. | Age (yr) | Female % |
|---|---|---|---|
| Control | 27 | 23.03 | 0.60 |
| Advice | 25 | 22.40 | 0.60 |
| Benchmark, absolute | 27 | 23.17 | 0.52 |
| Benchmark, absolute + Advice | 25 | 23.04 | 0.56 |
| Benchmark, relative | 28 | 24.08 | 0.68 |
| Benchmark, relative + Advice | 26 | 22.92 | 0.65 |
| Total | 158 | 23.10 | 0.60 |

Table 6.4 provides a summary of choices and performance per treatment. The total number of observations is larger than in table 3, given that each subject engaged in a series of choice tasks. The column “Decision time” indicates the average decision time in seconds, and the column “In-time” indicates the fraction of subjects who successfully made any decision within the allocated 30 seconds. The Column "Payoff" indicates the average payoff for subjects, and finally the Column "Net payoff" indicates the average payoffs corrected for the cost of buying advice. The Column "Best choice" indicates the percentage of subjects that chose the financial product with the lowest costs. The Column "Worst choice" indicates the percentage of subjects that chose the financial product with the highest costs. As can be seen for the groups in the non-advice treatment (Rows 1, 3 and 5) the payoff is equal to the net payoff, because no advice was available.

We see several results in Table 6-4. First, the advice treatment is associated with improved decisions, as indicated by the high fraction of "best choices" and lower fraction of "worst choices" compared to the non-advice treatment groups. Second, the improved decision-making in the advice treatment also results in higher payoffs, even if corrected for the cost of buying advice, as can be seen by comparing payoffs and net payoffs for the treatment group with the advice-only group (second row). Third, we find that – in comparison to the control group - both the absolute benchmark (ABM) and relative benchmark (RBM) treatments are associated with higher payoffs, with a corresponding reduction in

the fraction of worst choices. Fourth, we observe that both payoffs and the fraction of optimal choices is higher in the advice treatment than in the benchmark treatments. Finally, we observe that the combined advice and benchmark treatments (ABM and RBM) are associated with somewhat lower payoffs, fewer optimal decisions and more worst decisions than the advice treatment. In the next section we test the statistical significance of these differences.

Table 6.4 Decisions and payoffs by treatment

| Group | Obs. | Decision time | Payoff | Payoff net | Best choice | Worst choice |
|---|---|---|---|---|---|---|
| Control | 240 | 22.58 | 1.13 | 1.13 | 0.53 | 0.32 |
| Advice | 250 | 22.93 | 3.52 | 3.07 | 0.73 | 0.06 |
| Absolute benchmark (ABM) | 238 | 22.69 | 1.99 | 1.99 | 0.53 | 0.21 |
| ABM + Advice | 250 | 22.58 | 3.14 | 2.45 | 0.69 | 0.09 |
| Relative benchmark (RBM) | 250 | 21.67 | 1.65 | 1.65 | 0.52 | 0.22 |
| RBM + Advice | 260 | 20.33 | 3.08 | 2.45 | 0.65 | 0.08 |
| Total | 1488 | 22.11 | 2.48 | 2.16 | 0.61 | 0.16 |

## *Main estimation results*

We first test hypothesis 1; that benchmarks promote improved decision-making in complex choices. As can be seen in regression table 6.5, the relative benchmark (RBM) treatment has a significantly positive effect on payoffs, but only in the later rounds of the experiments, suggesting that subjects require some rounds of learning to becoming acquainted with the information presentation. Furthermore, the result in column 8 suggests that this effect is mainly driven by reduction in the fraction of worst choices. In contrast, the absolute benchmark (ABM) treatment does not have a statistically significant effect on payoffs, neither in the first and later rounds of the experiment, and only causes a marginally significant reduction in the fraction of worst choices.

Furthermore, as can be seen in tables 6.5 and 6.6, the advice treatment has a strong positive effect on decision quality and corresponding payoffs[35]. In addition, the advice treatment has a marginally significant negative interaction effect with the relative benchmark treatment in later rounds of the experiments, as indicated by the negative coefficient on the variable "A*RBM" in columns 2, 4, and 6. This indicates that in the presence of the relative benchmark, the advice treatment contributes less to improved decision-making. Moreover, the size of the coefficients is such that the positive effect of the advice treatment is absent in the RBM treatment.

Table 6.5 Treatment effects on payoffs

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Payoff | | Payoff (net) | |
| | Pt. 1 | Pt. 2 | Pt. 1 | Pt. 2 |
| Advice (A) | 2.025** | 2.715*** | 1.499* | 2.065** |
| | (0.842) | (0.925) | (0.844) | (0.939) |
| RBM | 0.133 | 1.946** | 0.105 | 1.884** |
| | (0.881) | (0.943) | (0.867) | (0.936) |
| ABM | 0.450 | 0.959 | 0.447 | 0.893 |
| | (0.863) | (0.921) | (0.878) | (0.932) |
| A*RBM | -0.649 | -2.194* | -0.812 | -2.171* |
| | (1.106) | (1.185) | (1.096) | (1.178) |
| A*ABM | -0.371 | -1.567 | -0.659 | -1.726 |
| | (1.143) | (1.202) | (1.144) | (1.223) |
| Controls | Yes | Yes | Yes | Yes |
| Observations | 833 | 498 | 833 | 498 |
| $R^2$ | 0.092 | 0.118 | 0.071 | 0.111 |

Estimates from OLS regressions. RBM is relative benchmark. ABM is absolute benchmark. Pt. 1 (part 1) refers to decisions rounds 1-6, pt.2 refers to rounds 7-10. Controls include: age, sex, risk preference, self-reported personality traits and financial product preferences. Clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[35] We do not control in our analysis for whether or not subjects used the advice that they purchased, since only N=7 subjects disregarded advice they purchased. Key results are robust to omitting these observations.

Hypothesis 2 states that the relative benchmark treatment leads to higher payoffs and better decisions than the absolute benchmark treatment. To test this hypothesis, we conduct the same set of regression analyses restricted to the subset of subjects in either of the two benchmark treatments. Again we compare the earlier and later rounds of the experiment. As indicated in table 6.7, column 4, we find that in comparison to absolute benchmarks, the relative benchmarks improved net payoffs.

Table 6.6 Treatment effects on decision quality

| | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|
| | Choice best | | Choice worst | | No choice | |
| | Pt. 1 | Pt. 2 | Pt. 1 | Pt. 2 | Pt. 1 | Pt. 2 |
| Advice (A) | 0.183** | 0.261*** | -0.217*** | -0.197*** | -0.046 | -0.119** |
| | (0.090) | (0.092) | (0.080) | (0.066) | (0.054) | (0.057) |
| RBM | -0.008 | 0.124 | -0.020 | -0.128** | -0.048 | -0.064 |
| | (0.085) | (0.089) | (0.063) | (0.057) | (0.052) | (0.050) |
| ABM | -0.027 | -0.016 | -0.031 | -0.085* | -0.040 | -0.012 |
| | (0.087) | (0.086) | (0.061) | (0.051) | (0.053) | (0.039) |
| A*RBM | -0.043 | -0.240* | 0.087 | 0.101 | -0.069 | -0.033 |
| | (0.116) | (0.126) | (0.103) | (0.100) | (0.070) | (0.080) |
| A*ABM | 0.040 | -0.082 | 0.035 | 0.123 | -0.027 | 0.038 |
| | (0.121) | (0.133) | (0.107) | (0.090) | (0.080) | (0.072) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 833 | 498 | 833 | 498 | 948 | 540 |
| Pseudo $R^2$ | 0.069 | 0.074 | 0.117 | 0.170 | 0.076 | 0.185 |

Marginal effects from probit regression. RBM is relative benchmark. ABM is absolute benchmark. Pt. 1 (part 1) refers to rounds 1-6, pt.2 refers to rounds 7-10. Controls include: age, sex, risk preference, self-reported personality traits and financial product preferences. Clustered robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The previous result can be attributed to the fact that, in comparison to the ABM treatment, the RBM treatment reduced the number of decision tasks in which were no choice was made. This effect was statistically significant during the latter stages of the experiment, with a reduction of approximately 10 percentage points,

as outlined in the result in table 6.8, column 6. Relative to the ABM treatment, the RBM treatment does not have a significant effect on the probability of subjects selecting the best or worst options.

Table 6.7 Treatment effects relative vs. absolute benchmarks

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Payoff | | Payoff (net) | |
| | Pt. 1 | Pt. 2 | Pt. 1 | Pt. 2 |
| Advice | 1.421*** | 0.482 | 0.667 | -0.204 |
| | (0.468) | (0.445) | (0.451) | (0.455) |
| RBM | -0.410 | 0.640 | -0.347 | 0.778* |
| | (0.466) | (0.471) | (0.442) | (0.460) |
| Controls | Yes | Yes | Yes | Yes |
| Observations | 568 | 337 | 568 | 337 |
| $R^2$ | 0.084 | 0.125 | 0.075 | 0.158 |

Estimates from OLS regressions. RBM is relative benchmark. Pt. 1 (part 1) refers to rounds 1-6, pt.2 refers to rounds 7-10. Controls include: age, sex, risk preference, self-reported personality traits and financial product preferences. Clustered robust standard errors in parentheses,
*** p<0.01, ** p<0.05, * p<0.1

Table 6.8 Treatment effects relative vs. absolute benchmarks

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Choice best | | Choice worst | | No choice | |
| | Pt. 1 | Pt. 2 | Pt. 1 | Pt. 2 | Pt. 1 | Pt. 2 |
| Advice | 0.160*** | 0.037 | -0.157*** | -0.043 | -0.038 | -0.050* |
| | (0.050) | (0.055) | (0.038) | (0.030) | (0.034) | (0.030) |
| RBM | -0.016 | 0.047 | 0.035 | -0.045 | -0.032 | -0.107*** |
| | (0.051) | (0.055) | (0.041) | (0.034) | (0.034) | (0.032) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 568 | 337 | 568 | 337 | 636 | 362 |
| (Pseudo) $R^2$ | 0.072 | 0.096 | 0.088 | 0.164 | 0.156 | 0.322 |

Marginal effects from probit regressions. RBM is relative benchmark. Pt. 1 (part 1) refers to rounds 1-6, pt.2 refers to rounds 7-10. Controls include: age, sex, risk preference, personality traits and product preferences. Clustered robust standard errors in parentheses,
*** p<0.01, **p<0.05, *p<0.1

We now turn to the hypothesis that benchmarks lower demand for advice. To test this, we estimate a probit model in which we regress the dummy variable which indicates whether subjects purchased advice (1=purchased advice, 0=did not purchase advice) on the benchmark treatments. Results are presented in table 6.9. We find that absolute benchmarks have a marginally positive effect on demand for advice (column 1), whereas relative benchmarks do not have such an effect (column 2). However, when we compare the effect of the absolute to relative benchmarks we find that the difference is not statistically significant (column 3). Overall we thus find no significant effect of benchmarks on demand for advice.

Table 6.9 Treatment effects relative vs. absolute benchmarks on advice demand

| | (1) Absolute vs. none | (2) Relative vs. none | (3) Absolute vs. Relative |
|---|---|---|---|
| Benchmark | 0.140* | 0.111 | 0.080 |
| | (0.073) | (0.111) | (0.062) |
| Controls | Yes | Yes | Yes |
| Pseudo $R^2$ | 0.319 | 0.288 | 0.345 |
| Observations | 500 | 510 | 510 |

Marginal effects of probit regression. Dependent variable is if subjects purchased advice (1=yes) Controls for round (pre. vs post round 6), age, sex, risk attitude, self-reported personality and product preferences. Clustered robust standard errors in parentheses,***p<0.01, **p<0.05, *p<0.1

## 6.5 Discussion

In this experiment we investigate how different kinds of choice architectures affect the quality of decision-making in complex choices. Subjects are presented a series of decision-tasks where they are asked to select the cheapest of four products that each consist of five cost-components. In addition, a time constraint of 30 seconds is imposed to increase the cognitive load. This setup serves to simulate complex choices that are ubiquitous in economic life, for example selecting a complex financial product such as a mortgage, which often has numerous cost-components and small print conditions.

In our experiment we compare the effect of two types of interventions. In the benchmarks treatments subjects are informed about "market averages" of each attribute, where the values for the various cost-components are expressed either in absolute deviation from or relative to the "market average" set to one hundred. In the advice treatments subjects can receive information regarding the sub-optimality of some of the options at a fixed cost.

Our results indicate that absolute benchmarks do not affect quality of decision-making. In contrast, relative benchmarks do improve decision quality. This result builds on previous studies that show that information processing and decision-making can be steered by modifying attribute salience (Jarvenpaa, 1990; Lurie and Mason, 2007; Mandel and Johnson, 2002). The findings in our study suggest that relative benchmarks improve decision quality by making optimal and suboptimal cost-components more salient, enabling subjects to rank products in terms of their respective number of optimal and suboptimal cost-components. In particular, our findings align with Samek et al. (2016) who find that allowing users to see the ranking of all features of a product improves the quality of decision making in complex choices. In a similar vein, our results align with Ericson and Starc, (2016) who show that the standardization of health insurance plans leads consumers to more accurately differentiate between choice options, but also that changes in the choice set are complementary to changes in the information interface.

Furthermore, we find that relative benchmarks not only affect decision quality, but also decision quantity, i.e. the probability that any of the four products is selected within the imposed time constraint. In line with results about decision quality, we find that the probability of a decision being made at all increases as products become more dissimilar in terms of the number of optimal and suboptimal attributes. This finding is line with previous studies that show that choice architecture can counter the problem of "choice inertia" without limiting the size of the choice set (Besedeš et al., 2015b). In light of the policy debate on

benchmarks and standard products our results allow for the tentative conclusion that their potential positive effect will depend on the specificities of the presentation and on product complexity and variety. Providing or imposing a (relative) benchmark is most likely to be effective in markets where a wide range of products with widely different attributes are being sold (i.e. phone contracts, health and pension plans and mortgages).

Our results furthermore indicate that the option to buy information improves the quality of decision making. This finding suggests that financial advice – in the absence of misalignment of incentives between advisor and client – can be beneficial for consumers.

In addition, we test how benchmarks affected the demand for advice. This research question is inspired by the fact that consumers that face complex financial choices, such as purchasing a mortgage, commonly have the option of purchasing advice from financial experts. Our results indicate that the provision of relative benchmarks does not affect the demand for costly information whilst absolute benchmarks have a marginally significant *positive* effect on the demand for advice.

The results regarding the positive effect of the costly advice treatment should be interpreted with caution. In particular, it is to be noted that our operationalization of the costly information treatments differs from real-life advice purchasing in two important aspects. In our treatments, subjects were given the option to pay in order to have two of the three suboptimal options being revealed as such. In contrast to real-life advice purchasing, subjects had to make their purchase decision before the actual choice set was revealed. Moreover, this information was not provided by another person, but directly by the computer. As such, agency problems and trust between advisor and advisee should not play a role in our experiment whereas such factors have been shown to significantly affect demand for advice in real markets (Bonaccio and Dalal, 2006; Inderst and Ottaviani, 2012). Still, by offering trustworthy advice of a known quality (two

options will be eliminated), we have established that such advice would improve decision making across the board and if anything is complementary, not substitute to (relative) benchmarks.

In sum, our study shows that decision-making in complex choices can be improved through benchmarks if these are presented in relative terms. Furthermore, we show that such benchmarks have limited effect on demand for advice. This being the first study on benchmarks in the context of complex choices, further research is warranted to replicate our results in different contexts and subject pools, explore the role of contextual factors in this process, and to shed more light on the psychological mechanisms by which relative benchmarks aid decision-making. Moreover, to increase external validity, several subsequent steps could be taken. In a first step, our lab experiment could be repeated with a subject pool of representative financial services consumers. In case our findings are confirmed, a field experiment may be conducted to focus on complex financial decision making without time pressure, allowing for heterogeneous preferences and risk attitudes, and possibly also introducing the element of trust in financial advice.

# Appendix 6.I Experiment screens

**Figure A1 -** *Sample choice screen CONTROL*

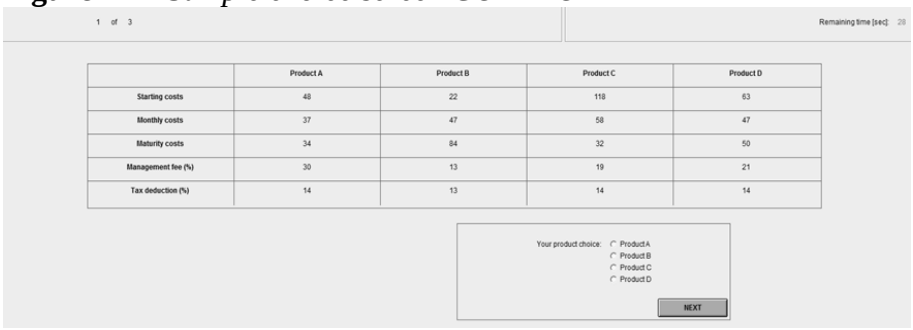



**Figure A2 -** *Sample choice screen ADVICE*

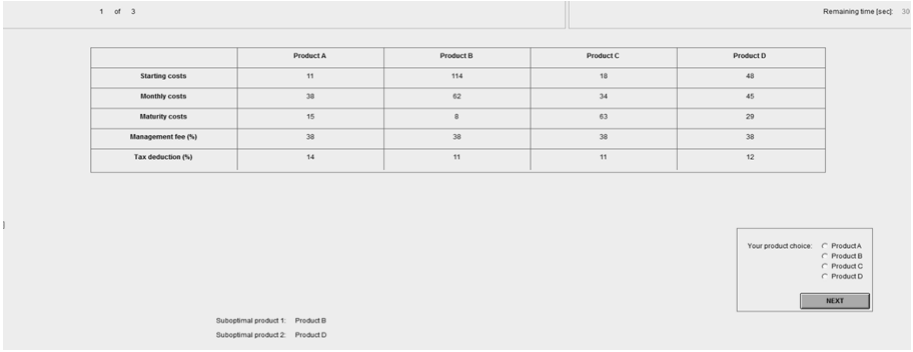



**Figure A3 -** *Sample choice screen ABSOLUTE AVERAGE*

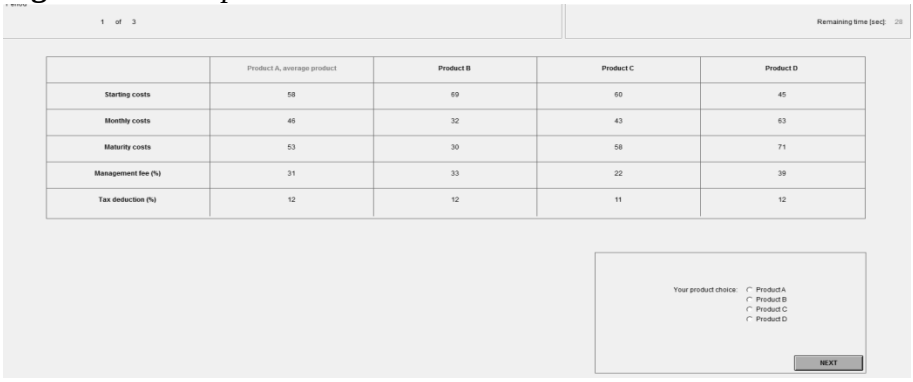


**Figure A4 -** *Sample choice screen ADVICE ABSOLUTE AVERAGE*

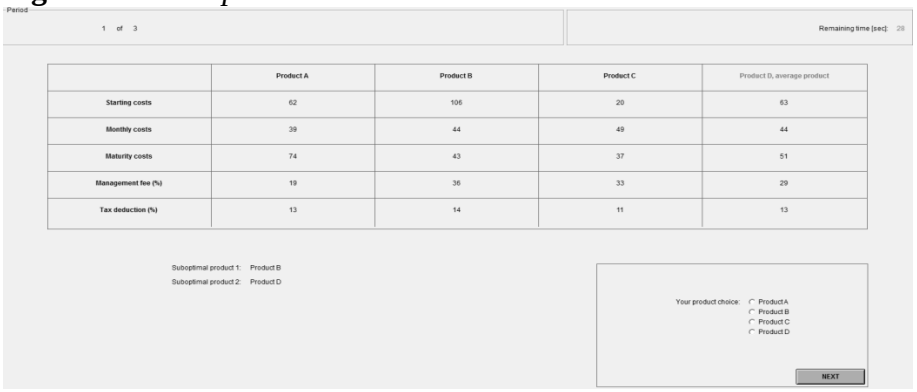



**Figure A5 -** *Sample choice screen RELATIVE AVERAGE*

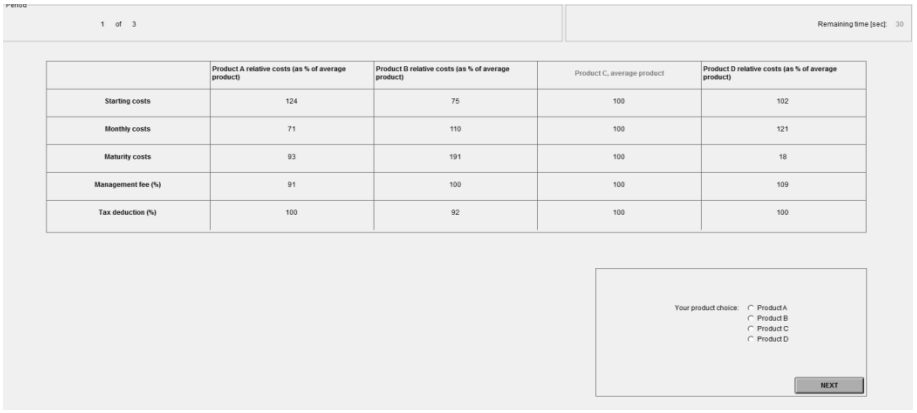



**Figure A6 -** *Sample choice screen ADVICE RELATIVE AVERAGE*

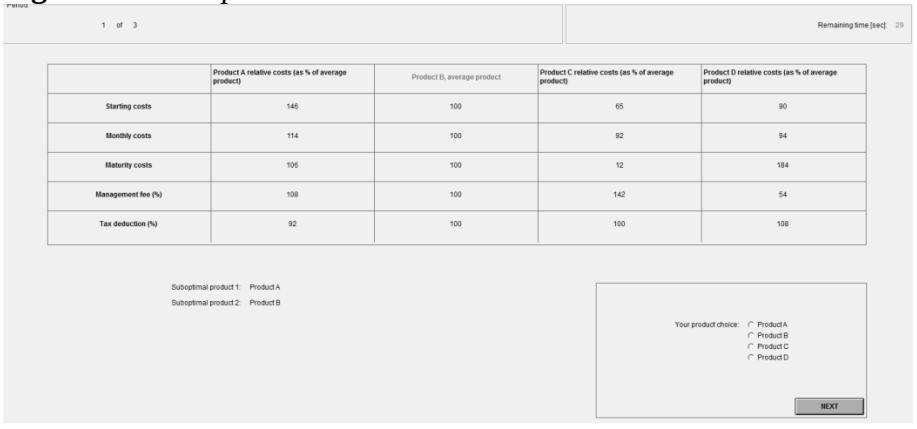
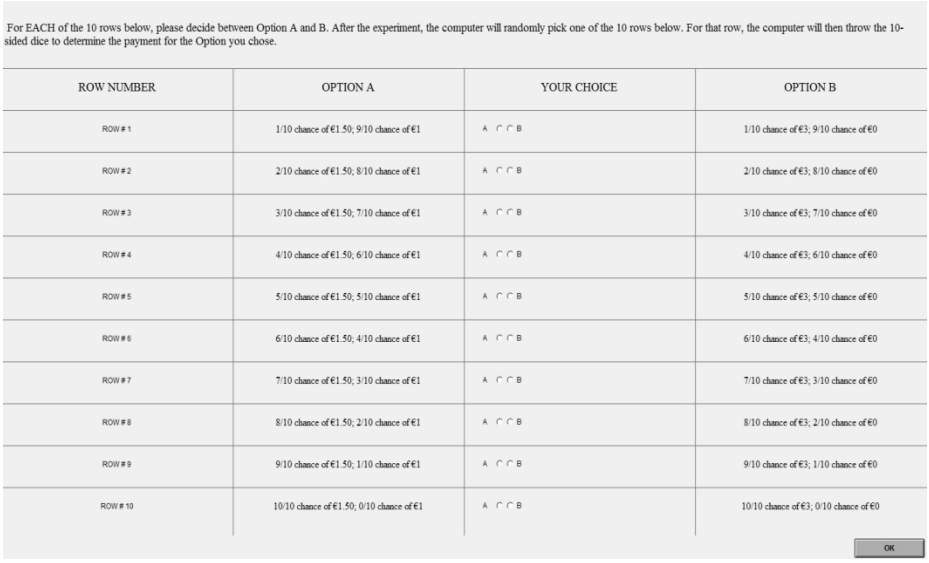


**Figure A7 -** *Choice screen Lottery*



For EACH of the 10 rows below, please decide between Option A and B. After the experiment, the computer will randomly pick one of the 10 rows below. For that row, the computer will then throw the 10-sided dice to determine the payment for the Option you chose.

| ROW NUMBER | OPTION A | YOUR CHOICE | OPTION B |
|---|---|---|---|
| ROW # 1 | 1/10 chance of €1.50; 9/10 chance of €1 | A ○ ○ B | 1/10 chance of €3; 9/10 chance of €0 |
| ROW # 2 | 2/10 chance of €1.50; 8/10 chance of €1 | A ○ ○ B | 2/10 chance of €3; 8/10 chance of €0 |
| ROW # 3 | 3/10 chance of €1.50; 7/10 chance of €1 | A ○ ○ B | 3/10 chance of €3; 7/10 chance of €0 |
| ROW # 4 | 4/10 chance of €1.50; 6/10 chance of €1 | A ○ ○ B | 4/10 chance of €3; 6/10 chance of €0 |
| ROW # 5 | 5/10 chance of €1.50; 5/10 chance of €1 | A ○ ○ B | 5/10 chance of €3; 5/10 chance of €0 |
| ROW # 6 | 6/10 chance of €1.50; 4/10 chance of €1 | A ○ ○ B | 6/10 chance of €3; 4/10 chance of €0 |
| ROW # 7 | 7/10 chance of €1.50; 3/10 chance of €1 | A ○ ○ B | 7/10 chance of €3; 3/10 chance of €0 |
| ROW # 8 | 8/10 chance of €1.50; 2/10 chance of €1 | A ○ ○ B | 8/10 chance of €3; 2/10 chance of €0 |
| ROW # 9 | 9/10 chance of €1.50; 1/10 chance of €1 | A ○ ○ B | 9/10 chance of €3; 1/10 chance of €0 |
| ROW # 10 | 10/10 chance of €1.50; 0/10 chance of €1 | A ○ ○ B | 10/10 chance of €3; 0/10 chance of €0 |

OK

# Appendix 6.2 Instructions

*No advice treatments*

**Welcome to this experiment!**
PLEASE READ THE FOLLOWING INSTRUCTIONS AND THE INSTRUCTIONS ON YOUR SCREEN VERY CAREFULLY AS IT WILL AFFECT YOUR PERFORMANCE AND PROFITS IN THIS EXPERIMENT.

**The experiment**
The experiment consists of three parts. In the first part you have to repeatedly choose a product. During the second part of the experiment you are asked to choose repeatedly in which lottery you want to take part. The last part of the experiment is a questionnaire.

**Part 1: Choosing the optimal financial product**
In the first part of this experiment you are presented a table with four different products (product A, B, C and D) and your goal is to select the optimal product. The optimal product is the product with the lowest total costs.

Suppose the different products are financial products and that the maturity time of each product is one year. This means that you buy the product for a period of one year. The total costs of the products are not given, however you are presented with four different sub costs and a tax deduction of each product:

1. *Starting costs:* you have to pay these costs once when you buy the product.
2. *Monthly costs:* these are monthly costs which you have to pay every month for the duration of one year.
3. *Maturity costs:* costs that have to be paid at the end of your contract.
4. *Management fee (presented as a percentage of starting costs):* fee is paid once
5. *Tax deduction (presented as a percentage of monthly costs):* a tax saving once a year

These costs will be presented in a table similar to Table 1.

**Table 1**

| | Product A | Product B | Product C | Product D |
|---|---|---|---|---|
| Start costs | | | | |
| Monthly costs | | | | |
| Maturity costs | | | | |
| Management fee (%) | | | | |
| Tax deduction (%) | | | | |

Knowing only the sub costs of all four products, it is your task to determine within 30 seconds which product is the optimal product. The time limit is presented in the upper right corner of your screen. If you do not choose a product within this time constraint, the computer will automatically choose the product with the highest total costs.

***Payoff of Part 1***

Your payoff depends on how well you make your decisions. You will be informed about how well you made your decisions only at the end of this part of the experiment. You are given an initial endowment of €8 and profits will be added to this when you choose optimal products and money will be deducted if you choose suboptimal products. At the end of the experiment the computer will randomly select one decision round and your decision in that specific round determines your payoff.

| **Decision** | **Payoff** |
|---|---|
| Optimal | + €5 |
| Second best | + €2.50 |
| Second worst | - €2.50 |
| Worst | - €5 |

***Screen instructions***

This hand-out sheet provides you with general information for the experiment. More specific information will be shown on your screen at the start of the experiment. Read these specific instructions very carefully as they influence your investment decisions.

**Part 2: Lottery**

In this part of the experiment 10 pairs of lotteries are presented to you and you have to choose in which lottery you want to take part in, lottery A or B. There is no time constraint in this part of the experiment.

***Payoff of Part 2***

The payoff you will receive in this part of the experiment is an extra payoff on top of what you already earned in the first part. The payoff depends on the lottery you participate in and on the outcome of the lottery. The outcome of the lottery is randomly decided by the computer. Also the payoff period is randomly determined. At the end of the experiment the computer will randomly select one of the pairs of lotteries and your decision for that specific pair determines your payoff. The possible payoffs are €3, €1.50, €1 and €0.

**Part 3: Questionnaire**

At the end of the experiment there will be a questionnaire for you to fill in. Please take your time and fill in this questionnaire truthfully. In the meantime we will prepare your payments from the previous parts of the experiment.

Good luck!

*Advice treatments*

**Welcome to this experiment!**

PLEASE READ THE FOLLOWING INSTRUCTIONS AND THE INSTRUCTIONS ON YOUR SCREEN VERY CAREFULLY AS IT WILL AFFECT YOUR PERFORMANCE AND PROFITS IN THIS EXPERIMENT.

**The experiment**

The experiment consists of three parts. In the first part you have to repeatedly choose a product. During the second part of the experiment you are asked to choose repeatedly in which lottery you want to take part. The last part of the experiment is a questionnaire.

**Part 1: Choosing the optimal financial product**

In the first part of this experiment you are presented a table with four different products (product A, B, C and D) and your goal is to select the optimal product. The optimal product is the product with the lowest total costs.

Suppose the different products are financial products and that the maturity time of each product is one year. This means that you buy the product for a period of one year. The total costs of the products are not given, however you are presented with four different sub costs and a tax deduction of each product:

1. *Starting costs:* you have to pay these costs once when you buy the product.
2. *Monthly costs:* these are monthly costs which you have to pay every month for the duration of one year.
3. *Maturity costs:* costs that have to be paid at the end of your contract.
4. *Management fee (presented as a percentage of starting costs):* fee is paid once
5. *Tax deduction (presented as a percentage of monthly costs):* a tax saving once a year

These costs will be presented in a table similar to Table 1.

**Table 1**

| | Product A | Product B | Product C | Product D |
|---|---|---|---|---|
| Start costs | | | | |
| Monthly costs | | | | |
| Maturity costs | | | | |
| Management fee (%) | | | | |
| Tax deduction (%) | | | | |

Knowing only the sub costs of all four products, it is your task to determine within 30 seconds which product is the optimal product. The time limit is presented in the upper right corner of your screen. If you do not choose a product within this time constraint, the computer will automatically choose the product with the highest total costs.

***Buying advice***

Before the above table is shown you will be offered the opportunity to buy advice. The price of advice is €2.50 and this will be deducted from the payoff if you decide to buy advice. Buying advice will significantly increase your chances of choosing the optimal product, as the least optimal product is automatically eliminated and also another suboptimal product is indicated. This leaves you with only two options. Please note that advice is bought for each round separately. Buying advice in round 1 will only provide you with advice for round 1. The choice to buy advice in round 2 will again be presented prior to round 2.

***Payoff of Part 1***

Your payoff depends on how well you make your decisions. You will be informed about how well you made your decisions only at the end of this part of the experiment. You are given an initial endowment of €8 and profits will be added to this when you choose optimal products and money will be deducted if you choose suboptimal products. At the end of the experiment the computer will randomly select one decision round and your decision in that specific round determines your payoff.

| **Decision** | **Payoff (without advice)** | **Decision** | **Payoff (with advice)** |
|---|---|---|---|
| Optimal | + €5 | Optimal | + €2.50 |
| Second best | + €2.50 | Second best | + €0 |
| Second worst | - €2.50 | Second worst | - €5 |
| Worst | - €5 | Worst | - €7.50 |

***Screen instructions***

This hand-out sheet provides you with general information for the experiment. More specific information will be shown on your screen at the start of the experiment. Read these specific instructions very carefully as they influence your investment decisions.

**Part 2: Lottery**

In this part of the experiment 10 pairs of lotteries are presented to you and you have to choose in which lottery you want to take part in, lottery A or B. There is no time constraint in this part of the experiment.

***Payoff of Part 2***

The payoff you will receive in this part of the experiment is an extra payoff on top of what you already earned in the first part. The payoff depends on the lottery you participate in and on the outcome of the lottery. The outcome of the lottery is randomly decided by the computer. Also the payoff period is randomly determined. At the end of the experiment the computer will randomly select one of the pairs of lotteries and your decision for that specific pair determines your payoff. The possible payoffs are €3, €1.50, €1 and €0.

**Part 3: Questionnaire**

At the end of the experiment there will be a questionnaire for you to fill in. Please take your time and fill in this questionnaire truthfully. In the meantime we will prepare your payments from the previous parts of the experiment.

Good luck!

# Appendix 6.3 Survey Instruments

To measure preferences for financial products we used the following set of questions as developed by the Dutch Authority for Financial Markets (AFM)[36]. Respondents are asked to indicate whether they indicate more with the statements on the left or the right of the screen, using a seven-point scale.

| | | |
|---|---|---|
| I acquire a lot of information | ooooooo | I try to limit the amount of information |
| I spend a lot of time on it | ooooooo | I try to do it as fast as possible |
| I consider all alternatives | ooooooo | I consider only a limited number of alternatives |
| I research as much as possible on my own | ooooooo | I prefer that others do the research for me |
| I trust advisors or intermediaries easily | ooooooo | I do not blindly trust the advice of advisors or agents |
| I talk a lot about it with friends and family | ooooooo | I do not talk about it with friends and family |
| I search until I have found the best product | ooooooo | I stop searching when I find a satisfying product |
| I am willing to make a bet | ooooooo | I play as much as possible on safe |
| I like to try new products | ooooooo | I stay with the familiar products |
| I prefer a simple product | ooooooo | It may also be a more complicated product |

To measures personality traits we used a ten-item version of Big Five Inventory, where subjects asked to indicate their agreement with the following statements (Rammstedt and John, 2007)[37]:

| I see myself as someone who.. | disagree strongly | disagree | disagree a little | neither disagree nor agree | agree a litle | agree | agree strongly |
|---|---|---|---|---|---|---|---|
| ..is generally trusting | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..tends to be lazy | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..is relaxed, handles stress well | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..has few artistic interests | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..is outgoing sociable | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..tends to find fault with others | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..does a thorough job | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..gets nervous easily | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ..has an active imagination | (1) | (2) | (3) | (4) | (5) | (6) | (7) |

[36] AFM (2004, December). Rapport: Kennismaking met de financiële consument. Retrieved from: https://www.afm.nl/nl-nl/nieuws/rapporten/2005/kennismaking-met-financiele-consument

[37] Due to a software problem, the results from the first BFI-10 survey question where not stored.

# Chapter 7
# Conclusion

## 7.1 Overview

Why do humans behave selfishly in some cases and selflessly in other cases? These understanding of these so-called "social preferences" is relevant to many economic phenomena such as financial fraud and charitable donations. This dissertation presents five experimental studies about the drivers of social preferences, with a particular focus on bounded rationality. In particular, these studies focus on how social preferences are shaped by culture, beliefs and choice architecture. The studies presented in this dissertation apply lab-, online- and field-experiments, which are increasingly recognized as complementary methods for studying the mechanisms of economic behavior (Charness and Fehr, 2015; Levitt and List, 2009). The studies use data on knowledge, beliefs and attitudes as measured with surveys and vignettes, as well as behavioral data obtained through incentived economic games and field observations. Where possible, different data sources are triangulated to test robustness and investigate the interplay between knowledge, beliefs, attitudes and behaviors (Jick, 1979). Subjects in these experiments include university students, bank employees, entrepreneurs and professionals in charitable organizations. Inspired by the vision of a unified science of behavior (Gintis, 2007) and employing methods from economics and psychology, this dissertation aims to contribute to a model of behavior which accounts for our inherent sociality and bounded rationality (Thaler, 2017).

## 7.2 Main conclusions

**Ethics interventions and bank employee behavior**

*Chapter 2* presents a study about the effect of an interactive ethics program on the behavior of front-office bank employees vis-à-vis their clients. Applying a field experiment and a novel mystery shopping method the study yield three main conclusions. First, in the majority of cases the bank employees did not provide advice which placed the clients' interest central. This indicates that despite the fact that banks and governments have taken various policy measures, the financial sector is still facing a challenge with placing the client first. Second, the program did not have a statistically significant effect on employee behavior vis-à-vis the clients. This result is robust to several model specifications and holds both in terms of actual observed advice and in terms of the mystery shoppers subjective perception of the bank employees' attitude towards the client. Furthermore, the ex-post survey indicates that the program did not have significant effect on employees beliefs and attitudes regarding ethical behavior. Third, when comparing shops with high vs. low variance in pre-existing levels of client-focused behavior it is found that the program actually had a significantly negative effect for the former group, while for the group the effect is moderately positive – particularly among shops with lower average levels of pre-existing ethicality. These results suggest that the program, whilst having the potential to promote an ethical culture, may have inadvertently provided a platform for less ethical employees to exert a negative influence on their colleagues. In sum, these results suggests details matter in developing ethics programs for the financial sector. Further testing of other approaches, which is discussed in the next section, is thus warranted.

**Professional expertise and biased fund allocation**

*Chapter 3* presents a study about professional expertise and the allocation of funds to charitable projects. Lab experiments with students show that subjects typically prefer to support projects that save larger relative numbers of lives, even if alternative projects can save larger absolute numbers of lives. We ask ourselves whether this bias to support projects that help large relative but small absolute numbers of lives is also displayed by people with expertise working in the charitable and social sector. This questions is relevant because many charitable funds are not allocated to the projects that could save most lives, and this discrepancy might be partly caused by the reference group bias in professionals that allocate these funds. Using a survey experiment with vignettes, we find that students and professionals are equally susceptible to this bias. Furthermore, we find that this bias is reduced when projects are evaluated not separately but in pairs. These result suggest that the reference group bias is not very sensitive to professional experience. In terms of practical relevance, this study suggests that joint- rather than separate evaluation of projects can contribute to funds being more effectively targeted.

**Free will belief and altruism**

*Chapter 4* presents a study about charitable giving and the role of people's belief in free will. This study builds on previous research which indicates that when belief is undermined people are more likely to cheat for personal gain. This study shows that that reducing of belief in free will indeed leads to reduced donations. However, the study also shows that this effect does not occur for religious subjects. This result is linked to the fact that religious people associate with norms that provide a moral imperative to help the poor. Consequently, it is argued that this subgroup is less susceptible to the effects of reduced belief in free will. In sum, this study extends the literature about the role of free will beliefs in altruistic behavior and suggests these effects depend on religiosity and social norms.

**Ambiguity and choosing for others**

*Chapter 5* presents a study about decisions that people make for others, and the role of ambiguity in such "other-decisions". The subjects, IT entrepreneurs in Bangladesh were presented with a description of two business trainings, which if followed would result either in certain yet small salary increase or a large salary increase with unknown probability, i.e. the ambiguous case. Some subjects were asked to make these choices for themselves while others chose on behalf of a colleague. The results indicate that attitudes towards ambiguity were not significantly different for self- vs. other-choices. However, we do find significantly lower ambiguity aversion for other-choices in the subset of subjects that believed the winning probability was below 50% for the ambiguous case. In sum, the findings suggest that uptake of new knowledge technologies with uncertain benefits may be greater when such decisions are delegated to others.

**Complex decisions and choice architecture**

*Chapter 6* presents an experiment about so-called complex choices, where people face multiple options that each consist of multiple attributes. Such complex choices are common in economic life, for example when consumers decide between mortgages and health insurance policies. In this lab experiments subjects are incentivized to select – under time pressure – the cheapest product from a list of products that each consist of multiple cost components. The treatment consists of subjects being presented with so-called "benchmark products" which provide a market average for each cost component. In one benchmark treatment attribute values were presented in absolute terms, while in the "relative benchmark" attribute values were expressed in relative terms to the market average which was set at 100. The results indicate that the relative but not the absolute benchmark leads to improved decisions. In turn, this suggests that a simple and low-cost information framing intervention can help consumers make improved decisions in complex choices.

## 7.3 Suggestions for future research

**Prosociality and bounded rationality**

This dissertation provides a number of suggestions for future research that pertain to the topic prosociality and how this is influenced by bounded rationality and cognitive biases. With respect to chapter 2 about bank employees behavior vis-à-vis clients, it would be interesting to use a similar audit methodology to investigate the effectiveness of ethics reminders. Whereas the specific context of this study resulted in low compliance with the so-called "ethical client greeting", whereby employees were instructed to use a particular slogan when greeting clients, simple "moral reminders" can be effective in promoting integrity (Shu et al., 2012). Various other types of reminders may have potential in the context of the banking sector. For example, employees may be requested to use a word with moral connotations when logging into their computer. A related strategy could be to send daily moral reminders to employees' phone or show pop-up messages on their computer screen. Such an approach has been shown to be effective in promoting empathy and prosocial behavior (Konrath et al., 2015). As for group meetings, while these seem to have potential for promoting ethical behavior, there seems to be a risk of "contagion" of unethical norms. To counter this it may be useful to experiment with group sessions that are led by vocal employees that have a record for their commitment towards clients and who can inspire less client-oriented peers to adjust their behavior.

Another related line for future research pertains to the role of cognitive biases in the allocation of charitable funds. In this context, one potentially interesting extension is to investigate experts' susceptibility to other biases such as the aversion to high overhead costs. Although our findings and previous studies – such as Kahneman and Klein (2009) – suggest that expertise is not associated with reduced susceptibility to various cognitive biases, there are also studies that do identify differences between novice and expert decision-making (Kaustia and

Knüpfer, 2008). Thus it remains an open questions whether professionals in the charitable sector are susceptible to other cognitive biases and to what extent this stands in the way of more effective allocation of funds to social projects. Furthermore, the finding that joint evaluation significantly reduced susceptibility to the reference group bias – as well as other findings, such as the result that subtle differences in framing can reduce donors' irrational aversion to overhead costs (Gneezy et al., 2014) – suggests that small changes in the choice architecture can promote more effective fund allocation. It would be interesting to explore other applications of this principle, for example the portioning the choice set in smaller subsets, which has been shown to improve decision making in complex consumer choices (Besedeš et al., 2015).

**Prosociality and beliefs**

Next, this dissertation investigates the implications of beliefs on prosocial behaviors, in particular in chapter 4 which concerns how belief in free will influences charitable donations. A potential avenue for future research is to investigate whether the interaction of free will beliefs and religion-based adherence to helping norms also extends to secular philosophies that promote helping norms, such as August Comte's "religion of humanity" which holds altruism as a central pillar (Comte, 1852). Another open question concerns the duration of the behavioral effects induced through priming free will disbelief. Whereas our experiment measured altruism directly after subjects were exposed to the reading task it remains unclear whether changes in beliefs and subsequent behaviors are longer-lasting. Furthermore, it might be interesting to investigate the behavioral effect of more intensive and frequent exposure to ideas that could undermine the belief in free will, for example in the case of students participating in a course about neuroscience.

In addition, this study also raises the question how beliefs about free will influence people's attribution of responsibility to others? Whereas our result and

the previous literature has focused on the way in which free will disbelief can provide people with a "motivated reason" to justify their own selfish tendencies, it could also be argued that free will disbelief influences considerations about the responsibility and agency of others. Whereas a person with strong belief in will and personal agency might conclude that a homeless person has only himself to blame for his predicament, someone whose belief in free will and agency has been undermined may consider this homeless person a victim of circumstances. It could be interesting to explore the behavioral effects of exposure to messages that highlight how the agency and free will of "victims" is constrained by forces outside of their control. Such a study would tie into a broader research agenda about the role of "just world beliefs" in shaping redistributive and political preferences (Bénabou and Tirole, 2006) and how such beliefs and preferences differ across cultures (Frank et al., 2015).

**Bounded rationality and individual decisions**

Finally, there are a number of suggestions for future research pertaining to bounded rationality and decisions where there are no direct tradeoffs between the self and others. Regarding chapter 5, one avenue for future research would be to investigate self-other differences in the context of incentivized decisions. There is already one recent incentived lab-experiment that shows no significant self-other differences in ambiguity attitudes (König-Kersting and Trautmann, 2016b), and extensions into relevant field settings would make a welcome contribution to this knowledge base. Furthermore, it would be interesting to explore the role of perceived social distance towards the "other" in self-other choices in the domain of risk and uncertainty. Whereas our study found no difference in ambiguity attitudes between choices made for hypothetical others that were socially closer (a befriended colleague) vs. more distant (an unknown colleague), various other studies suggest that social distance is associated with reduced levels of risk aversion (Sun et al., 2017); impatience (Ziegler and Tunney, 2012) and loss

aversion (Andersson et al., 2013). It is suggested that when choosing for others that are further removed socially, emotions such impatience and fear play a lesser role compared to self-choices, which in turn leads to decisions that are more rational in terms of maximizing expected utility.

Finally, the experiment presented in chapter 6 raises various questions. The main finding that relative but not absolute benchmarks improve decision quality suggests that subjects were steered towards using the heuristic of between-product comparisons of attributes, rather than calculating costs for each separate products, as such comparisons are arguably easier in the relative framing. However, there may also be choice environments where such a heuristic is suboptimal, for example when certain cost components constitute a disproportionate fraction of total product cost. In such situations, a heuristic of counting per product the number optimally priced attributes may result in suboptimal decisions. Experiments with variation in cost function parametrization can shed more light on this question. Another avenue for future research would be to replicate this study with a sample of non-student subjects. Since numerical and financial literacy arguably plays an important role in complex choices such as in our experiment, the benefits of choice architecture interventions such as the relative benchmark treatment may be greater in the general population. Field extensions, for example in the domain of mortgages, could then be a next step in investigating this question.

## 7.4 Policy implications for private and public sector

The studies in this dissertation have a number of tentative implications for economic policy. The word tentative is used here because human behavior is strongly context-dependent, a point which is discussed in more detail below. As such, the results of empirical studies which consider behaviors in a particular context should not be extrapolated to other settings without caution. This being

said, the results of this dissertation have several implications for practitioners and policymakers.

**Putting behavioral assumptions to the experimental test**

First, the overall findings fit into the increasingly agreed-upon view that humans are neither fully rational nor fully selfish. This in turn implies that policymakers involved in modelling and predicting economic processes at the macro-level would benefit from accounting for these deviations from the classical assumptions of a fully rationality and self-interested homo economicus, which is still a central model for many policymakers.

Next, this dissertation support the increasingly popular notion that field experimentation is a valuable tool for policymakers and practitioners that aim to influence human behavior. In recent years, various governments have started more systematic efforts to conduct behavioral experiments in order to improve policy effectiveness, with the notable example of the British "Behavioral Insights Team" whose experiments have resulted in improved policies in the domains of health, consumer finance, education and energy sustainability. The Netherlands is following a similar trend, with a network of behavioral scientists working within- and with ministries to experimentally test and improve policies in various domains (Dur, 2017).

The rationale for this "evidence-based policy" approach is that many observational research fails to provide accurate estimates of the causal effect of policies due to unobservable differences between people exposed and not exposed to a policy. Such observable differences can be particularly relevant in the behavioral domain, for example if more motivated people self-select into a certain policy. This endogeneity problem can be solved by random assignment, as applied in lab- and field experiments, and sometimes by quasi-experimental methods that exploit variation around a precise cut-off, for example in education policies where

students with grades above a threshold are eligible for scholarships to study abroad (Oosterbeek and Webbink, 2011).

In recent years, field experiments firms have also become more popular in firms (Bandiera et al., 2011; Delfgaauw et al., 2013; Levitt and Neckermann, 2014), which increasingly wish to understand the causal effects of their activities (Maas and Liket, 2011).

However, given the inherent context specificity of field experiments policymakers should not take results from individual studies as ultimate truths. For example, while our field experiment indicates that an ethics program has limited effects on the behavior of bank employees, it may well be that in another context – for example a context with less within-group variation – such a program could in fact be more effective. It thus seems appropriate for policymakers and practitioners to closely work together to develop, test and improve policies and interventions up to the smallest details. Furthermore, by conducting multiple experiments with similar interventions across different contexts a better understanding can be gained of the robustness of the effect of such policies. A notable example here pertains to policies to promote access to financial services for people in low-income countries. It was long thought that the provision of loans to low-income groups in developing countries would greatly reduce poverty, for example by helping people to start up their own business. The promise of this so-called "microfinance" movement, for which Muhammad Yunus received the 2016 Nobel prize, was put to the test in a number of field experiments across different countries and borrower groups. A recent review of six of these studies concluded that – despite initial enthusiasm from policymakers, practitioners and academics – the microfinance programs only showed moderate effects (Banerjee et al., 2015). This finding in turn can help policymakers in allocating scarce resources to development programs. Similarly, it would be valuable to conduct multiple field experiments around interventions to promote ethical behavior in the financial

sector, such that policymakers and practitioners learn which type of interventions are most promising.

On a related note, it is also worth stressing here that people do not all behave and respond to policies and interventions in a similar manner. As Albert Camus said "*we are all special cases*". The results from this dissertation align with this basic idea: different subgroups responded differently to programs in the field (see chapter 2) and treatments in the lab (see chapters 4 and 5). At the same time, experimentation with different groups of subjects can also help in discrediting common-sense ideas such as the notion that professionals are less susceptible to certain cognitive biases (see chapter 3). This all goes to show that policymakers would benefit from testing their cherished ideas, policies and preconceptions, whilst maintaining a critical perspective regarding the universality of findings from individual studies. While in the context of well-controlled physics experiments Einstein rightly pointed out that *"...[no] amount of experimentation can ever prove me right; a single experiment can prove me wrong*", it may take a few more experiments in different settings, with different subject populations, to firmly reject the rational for economic policies. All the same, the bottom line is that the evidence should have the final word, and policymakers should reject or adjust policies which are shown – by experiment – to be ineffective.

**Promoting prosocial behavior**

Next, this dissertation provides a number of concrete policy implications regarding the promotion of prosocial behavior in various segments of the economy. The results from chapter 2 suggest that misconduct in the financial sector, in particular behavior whereby clients' interests are not placed central, is still prevalent, despite the various policy measures that have been taken in recent years. This should motivate practitioners to test other measures to promote such client-focused behavior. Since the ethics training program tested in our experiment had limited success, and was even contra-productive for part of the bank shops, it is desirable

to test other approaches which are less sensitive to contagion of social norms by employees that are less concerned about the client. An example of such an alternative policy worth testing is the use of frequent ethical reminders that can be delivered to employees in a personal manner, for example through their mobile phone. The findings in chapter 3 indicate that charitable funds may be more effectively allocated if donors and professionals in charitable organizations evaluate projects not in isolation but rather in a comparative manner. This way, it becomes easier for decision-makers to identify which projects yield the greatest benefits in absolute terms. Finally, the results from chapter 4 suggest that beliefs about agency and free will shape people's willingness to help others. For policymakers and organizations that wish to promote support for certain forms of redistribution, for example development aid and charitable donations, it may be worthwhile to experiment with information and education interventions that highlight the limited agency or free will of victims, since this may promote support for such policies.

**Accounting for bounded rationality**

Subsequently, the results from chapter 5 provide less clear-cut policy recommendations, since self-other differences in ambiguity attitudes are only observed for a subset of subjects. However, since effects are observed for a considerable subset, and because delegated choices that involve uncertainty are common in many domains with policy relevance – for example in the case of teachers and parents making choices for their children – it would be worth to conduct further experiments in this domain. For example, in line with the social distance hypothesis, parents may be more inclined than teachers to stimulate their children to pursue a study which provides more certainty in terms of career perspectives, even though the expected benefit may be smaller. Finally, the results from chapter 6 suggest that policymakers may wish to promote among consumers the use of benchmark tools that allow for easier comparison of product features in

the context of choices between large numbers of complex products, such as mortgages. However, caution is warranted here because the benefit of such tools may depend on the weights of the cost components. Either way, it would be recommendable to test such tools before large-scale implementation.

## 7.5 Epilogue

A nice thing about applying the scientific method to the study of social preferences and bounded rationality is that findings apply not just to others, but also to myself. Why do I sometimes behave quite selfishly yet help complete strangers at other times? Why do I keep postponing to switch to a better health insurance policy? Why do I give rational advice to my friends yet make silly mistakes myself? All these questions fall in the domain of behavioral economics, and I feel fortunate and thankful for having had the luxury to spend over four years working together with inspiring and helpful people to gain a better understanding of such matters.

In the spirit of the ancient aphorism "know thyself" the writing of this dissertation has made me increasingly aware of my own bounded rationality and cognitive biases, such as my forgetfulness and tendency to discount or even avoid information that is contrary to my beliefs. Luckily, I am also prone to occasional over-optimism and self-deception, so I can maintain a somewhat positive self-image despite all these shortcomings.

While there unfortunately still is much selfishness, stupidity and suffering in this world, the past years of research have given me hope that there is plenty of scope for influencing humans towards smarter and more social decisions. In contrast to what is sometimes assumed, our preferences and behaviors are not fixed and predestined, but dynamic products of evolution, culture, institutions, incentives, policies and interventions. It is my impression that economics and the social sciences are only just starting to unravel the fascinating interplay of these

factors, and that there are enough interesting problems to keep us busy for the foreseeable future.

Moreover, the bounds to our sociality and rationality are expanding rapidly through cultural evolution, scientific advances and the development of new forms of intelligence, which can help "mankind 1.0" in gaining a better understanding of itself and the world. How these developments will unfold – and how they will influence our beliefs, behaviors and the human condition – remains an open question far beyond the scope of this dissertation. For now, it is my hope that this dissertation can make a small contribution to a more selfless world in which people make better choices for themselves, and for others.

# Nederlandse samenvatting

Waarom gedragen mensen zich soms onbaatzuchtig en coöperatief, en soms juist zeer egoïstisch? Dit is een centrale vraag in de economie. Denk bijvoorbeeld aan een verschijnsel als corruptie, waarvan de totale kosten circa 5 procent bedragen van het "gross world product" (Organisation for Economic Co-operation and Development, 2014). Tegelijkertijd zijn er genoeg voorbeelden van onbaatzuchtigheid, zoals blijkt uit het feit dat Nederlanders in 2015 naar schatting 5.7 miljard euro doneerden aan goede doelen (Bekkers et al., 2017). Terwijl een verschijnsel als corruptie in lijn is met de klassiek economische assumptie van een door eigenbelang geleide mens, zijn andere verschijnselen zoals het doneren van geld aan vreemden in strijd met dit model van de "homo economicus".

Deels kunnen zulke schijnbaar onbaatzuchtige gedragingen gerijmd worden met het klassieke model door het verschijnsel van "warm-glow", waarbij mensen zelf een plezierig gevoel overhouden aan het helpen van anderen, zodat dit gedrag indirect toch als zelfzuchtig gezien kan worden (Andreoni, 1990). Hiernaast worden vanuit de evolutionaire biologie diverse vormen van altruïsme geduid als indirecte vormen van zelfzuchtigheid. Zo kan bijvoorbeeld het helpen van bloedverwanten onder bepaalde condities voordelen opleveren voor de eigen genen, en hiermee indirect toch weer als zelfzuchtig kwalificeren. Het helpen van een niet-familielid die later weer een wederdienst verleent kan aldus niet als puur altruïsme worden beschouwd (Hamilton, 1964; Trivers, 1971).

Echter, deze theorieën bieden slechts gedeeltelijke verklaringen. Zo geeft het warm-glow model geen antwoord op de vraag *waarom* mensen een plezierig gevoel overhouden aan het helpen van vreemden. En de evolutionaire modellen van "kin altruism" en "reciprocal altruism" verklaren niet waarom mensen, zelfs in een volledig anonieme context zonder positieve effecten op hun reputatie, substantiële hoeveelheden geld doneren aan vreemden.

Een meer volledige verklaring van het feit dat mensen soms egoïstisch en soms altruïstisch handelen vloeit voort uit de sociale context waarin de mensheid is geëvolueerd. In deze prehistorische context waren groepen met meer onderlinge samenwerking en altruïstisch gedrag vaak succesvoller dan groepen van egoïstische mensen, bijvoorbeeld bij het jagen op dieren of het strijden met andere groepen (Bowles and Gintis, 2011). Tegelijkertijd zijn er voor individuën binnen samenwerkende groepen nog steeds voordelen te behalen door zelfzuchtig te handelen. Onder zulke condities kan handhaving van de sociale normen worden gefaciliteerd door "altruïstisch bestraffen" waarbij mensen non-coöperatieve groepsgenoten bestraffen, ondanks het feit dat zij hiermee zelf kosten en risico's lopen (Fehr and Gächter, 2002). Een cultuur waarin scheefschaatsen wordt afgekeurd en bestraft biedt daarmee voordelen voor alle groepsleden.

Naast deze evolutionaire invloeden (Bowles and Gintis, 2011; Cesarini et al., 2009) wordt ons sociaal gedrag ook gevormd door de interactie van onze psychologie en onze hedendaagse omgeving. In klassieke economische theorie wordt aangenomen dat de mens rationele keuzes maakt en consistente preferenties heeft. Zo zou iemand die van ijs houdt meer bereid zijn te betalen voor een grote schep ijs dan voor een kleine schep. Echter, deze assumptie gaat niet altijd op omdat mensen beperkte cognitieve capaciteiten ("bounded rationality") hebben.

Zo blijkt bijvoorbeeld dat, wanneer de grotere schep ijs in zeer groot bakje wordt gedaan en de kleinere schep ijs in een zeer klein bakje, mensen vaak het kleine bakje prefereren. De interpretatie van deze "bias" is dat wij moelijker een inschatting kunnen maken van absolute hoeveelheden (hoeveel ijs een bakje bevat) dan van proporties (hoe vol een bakje zit). Gegevan dat we meer focusen op aspecten die gemakkelijker kunnen worden geëvalueerd kan dit leiden tot inconsistente keuzepatronen (Bartels, 2006). In een wereld waarin grote bakjes meer ijs bevatten kan zo'n vuistregel ("heuristic") functioneel zijn, maar in een wereld met doorgewinterde ijsverkopers kan dit leiden to economische irrationaliteit (Gigerenzer and Gaissmaier, 2011; Haselton et al., 2015).

De ontdekking van verschillende “cognitive biases” door psychologen en economen in de afgelopen decennia heeft de beperkingen van het klassieke model van de mens als rationele beslisser blootgelegd, en vormt de basis voor de gedragseconomie (Kahneman, 2003; Thaler, 2015). Deze psychologische kronkels, die in het verleden wellicht adaptief waren, zorgen ervoor dat mensen frequent inconsistente beslissingen nemen die suboptimaal zijn voor henzelf, zoals in de ijssalon, maar ook bij meer meer gewichtige beslissingen zoals het kiezen van een beleggingsproduct of hypotheek. Personen die worden gedreven door eigenbelang maken geregeld ook gebruik (of beter: misbruik) van de biases van anderen, zoals bijvoorbeeld bij adviseurs in de financiële sector (Mullainathan et al., 2012). Ook zijn mensen soms bewust irrationeel en inconsistent in de manier waarop ze bepaalde informatie verzamelen en interpreteren, om zo hun incorrecte overtuiging in stand te houden en hun zelfzuchtige preferenties te rechtvaardigen. Zo kiezen sommige klimaat-sceptici ervoor om enkel stukken te lezen waarin het broeikaseffect wordt ontkent, en kritische artikelen te ontwijken. Door het bewust ontwijken van deze informatie kunnen ze zichzelf bijvoorbeeld laten geloven dat ze sociaal handelen wanneer ze met een SUV boodschappen gaan doen(Sweeny et al., 2010).

Hiernaast kunnen sommige cognitieve biases ertoe leiden dat onbaatzuchtige mensen keuzes nemen die suboptimaal zijn voor degenen die ze willen helpen, bijvoorbeeld in het bij donaties aan goede doelen. Zo letten veel mensen bij het evalueren van een stichting meer op de overhead kosten (een proportie) dan op de uiteindelijke effecten die worden bereikt (een absolute waarde, bijvoorbeeld het aantal levens dat wordt gered). Het gevolg is dat met gedoneerde gelden niet het maximale aantal mensen wordt geholpen (Gneezy et al., 2014).

Voor een belangrijk deel zijn de inzichten over sociale preferenties en bounded rationality voortgevloeid uit de “experimentele revolutie” binnen de economie. Tot laat in de twintigste eeuw voerde theoretisch onderzoek de

boventoon en was het empirische onderzoek hoofdzakelijk observationeel, waardoor causaliteit vaak lastig kon worden onderscheiden van correlatie. Vanaf de jaren tachtig raakte de experimentele benadering geleidelijk in zwang binnen de economie (Hamermesh, 2013).

Met lab-experimenten worden onder zeer gecontroleerde condities de mechanismes van economisch gedrag onderzocht. Bij veldexperimenten is er minder controle, maar hier staat tegenover dat onderzoek kan worden gedaan in een realistische context die moeilijk is te repliceren in het lab. Ook weten subjecten in veldexperimenten vaak niet weten dat ze onderdeel zijn van een studie, waardoor sociaal wenselijk attitudes en gedragingen en andere vormen van "observation bias" kunnen worden vermeden. In recente jaren zijn veldexperimenten ook steeds populairder bij overheden, bedrijven en andere organisaties die willen begrijpen hoe zij burgers, medewerkers en consumenten kunnen "nudgen" richting bepaald gedrag (Dur, 2017; Thaler and Sunstein, 2008). Ook zijn er in recente jaren diverse experimentele tussenvormen ontstaan, zoals het "lab-in-the-field" experiment waarbij een lab-experiment met subjecten in andere populaties wordt uitgevoerd, zoals consumenten in ontwikkelingslanden. Deze verschillende experimentele methodes zijn complementair en bieden – wanneer gecombineerd – inzicht in de robuustheid van de causaliteit en de invloed van contextuele factoren (Stoop et al., 2012).

Deze proefschrift poogt met een vijftal veld- en lab-experimenten een bijdrage te leveren aan het antwoord op de vraag waarom mensen zich soms onbaatzuchtig en soms juist egoïstisch gedragen. Hierbij staat de rol van cultuur, cognitieve biases en bounded rationality centraal. Tevens wordt in de laatste twee hoofdstukken gekeken naar individuele keuzes. De contributies van deze studies zijn voornamelijk empirisch; hiernaast poogt het eerste hoofdstuk ook een methodologische bijdrage te leveren in de vorm van een nieuw instrument om klantgericht gedrag van werknemers in de bancaire sector te meten.

Hoofdstuk 2 (in samenwerking met Alain Cohn, Ernst Fehr, Karen Maas, Helen Toxopeus en Michel Marechal) is gericht op de vraag hoe klantgericht gedrag bij bankmedewerkers kan worden gestimuleerd. Deze studie bouwt voort op eerder experiment waaruit blijkt dat de cultuur bij banken oneerlijk en zelfzuchtig gedrag bij medewerkers in de hand werkt (Cohn et al., 2014). Middels een veldexperiment bij alle filialen van een grote commerciële bank wordt het effect getoetst van een gedragsinterventie waarbij bankmedewerkers wekelijks met elkaar spreken over ethische dilemma's die zij in hun werk tegenkomen. Het klantgericht gedrag van deze medewerkers word gemeten met een "mystery shopping" methode, waarbij acteurs zich voordoen als potentiële klanten en met gestandaardiseerde scripts advies van de medewerker over financiële producten inwinnen. Elk bankfiliaal is 12 keer bezocht, 6 keer voorafgaand aan het programma, en 6 keer tijdens de periode van circa 2 maanden waarin het programma werd geïmplementeerd. De mystery shopping scripts waren zo geconstrueerd dat het belang van de bankmedewerker (producten verkopen) haaks staat op het belang van deze vermeende klant; namelijk om geïnformeerd worden over het feit dat deze producten niet geschikt waren, bijvoorbeeld omdat de klant het product niet goed begreep. De belangrijkste bevinding is dat dit programma – gemiddeld genomen – geen statistisch significant effect had op het klantgericht gedrag en dat de meerderheid van bankmedewerkers in hun advies de klant niet centraal stellen. Ook blijkt dat het programma een negatief effect had bij filialen waar voorafgaand aan het programma meer variatie was in de mate van klantgerichtheid. Een mogelijke verklaring hiervoor is dat de groepsgesprekken onbedoeld fungeerden als platform voor de dat minder klantgerichte medewerkers om de klantgerichtheid van hun collega's te ondermijnen.

Hoofdstuk 3 (in samenwerking met Dinand Webbink) gaat over de zogenaamde "reference group bias". Op basis van eerdere lab-experimenten blijkt dat de beoordeling van charitatieve projecten sterker wordt beïnvloedt door het deel van de populatie die geholpen wordt dan door het absolute aantal mensen

(Bartels, 2006). Zo geven de meeste mensen een hogere waardering aan project A waarbij in een vluchtelingenkamp van 1.000 personen er 500 mensen worden gered (50%) van een levensbedreigende ziekte dan aan een project B met identieke kosten waarbij in een vluchtelingenkamp van 10.000 personen er 1.000 mensen (10%) worden gered. Dit is irrationeel voor iemand die zoveel mogelijk levens wil redden, omdat in project B voor gelijke kosten dubbel zoveel levens worden gered. Deze studies zijn echter uitgevoerd met studenten en het is de vraag of professionals bij goede doelen en overheidsinstellingen net zo gevoelig voor deze bias. We onderzoeken dit met een drietal experimenten met zowel studenten en experts. Subjecten krijgen opeenvolgend een aantal hypothetische projecten voorgelegd waarbij er variatie is op twee dimensies: (i) het aantal mensen dat gered kan worden en (ii) het totaal aantal mensen in de groep waarvan subgroep een onderdeel vormt (de "reference group"). De voornaamste bevinding van deze studie is dat beide groepen een vergelijkbare mate van gevoeligheid hebben voor deze bias. Hiernaast onderzoeken we of de gevoeligheid voor deze bias kan worden verminderd door mensen projecten te laten vergelijken, in plaats van deze sequentieel en los van elkaar te evalueren. De intuïtie hierbij is dat men de absolute aantallen begunstigden in de projecten onderling kan vergelijken, waardoor dit aspect makkelijker kan worden geëvalueerd en vervolgens meer gewicht krijgt in de uiteindelijke waardering. Voor zowel studenten en experts leidt deze "joint evaluation" modus inderdaad tot een verminderde gevoeligheid voor de reference group bias. Donaties en projectgelden zouden dus waarschijnlijk effectiever kunnen worden ingezet wanneer mensen verschillende opties met elkaar vergelijken.

Hoofdstuk 4 (in samenwerking met John Protzko, Kellie Liket en Vera Schölmerich) betreft het verschijnsel van "motivated reasoning" en de manier waarop zelfzuchtigheid wordt beïnvloedt door ideeën over het bestaan van vrije wil. Deze studie is geïnspireerd op een experiment waaruit blijkt dat mensen vaker bereid zijn te liegen voor financiëel gewin nadat ze een tekst hebben gelezen

waarin wordt geclaimd dat vrije wil niet bestaat. In deze tekst, een artikel uit het populair wetenschappelijk blad "New Scientist" wordt neurowetenschappelijk onderzoek aangehaald waaruit zou blijken dat alle keuzes die wij maken worden bepaald door onze hersenen en omgevingsfactoren. Subjecten die deze tekst hadden gelezen (de treatment groep) waren vervolgens vaker bereid om te liegen voor persoonlijk gewin dan subjecten die een neutrale tekst hadden gelezen over energie technologie; de controle groep (Vohs and Schooler, 2008). De interpretatie van dit resultaat is dat ongeloof in vrije wil mensen een excuus geeft om hun zelfzuchtigheid te rechtvaardigen. Voortbouwend hierop vroegen wij onszelf af of ongeloof in vrije wil ook altruïstisch gedrag zou ondermijnen. Om dit te testen voerden wij via het online platform Amazon mTurk een experiment uit bij N=108 subjecten in de Verenigde Staten. Bij de treatment groep werd het geloof in vrije wil significant gereduceerd door het lezen van het New Scientist artikel. Vervolgens speelden subjecten 24 "binary dictator games" waarbij ze een geldbedrag op twee manieren konden verdelen tussen zichzelf en een arm persoon in een ontwikkelingsland. Hierbij was er telkens deze keuze tussen een meer en minder zelfzuchtige verdeling (bijvoorbeeld alles zelf houden vs. de helft weggeven). In lijn met onze verwachting blijkt dat subjecten in de treatment groep vaker de meer zelfzuchtige optie kiezen. Echter, voor religieuze subjecten trad dit effect niet op. Een mogelijke verklaring hiervoor is dat voor deze subjecten er vanuit hun religie een sterkere sociale norm is om geld te geven aan de armen, en dat het geloof in vrije wil geen directe effecten heeft op het conformeren aan deze norm. In andere woorden, als altruïsme geen keuze is maar een sociale verplichting speelt vrije wil – of in elk geval, het geloof daarin - een minder belangrijke rol.

Hoofdstuk 5 (in samenwerking met Karen Maas en Haki Pamuk) gaat over de rol van onzekerheid in keuzes die mensen maken voor anderen; hierbij kan bijvoorbeeld worden gedacht aan een financieel manager die voor haar klant beslist in welke aandelen te investeren. In de economie wordt een onderscheid

gemaakt tussen risico en onzekerheid (Knight, 2012). In het geval van risico zijn winsten en kansen bekend, bijvoorbeeld als iemand geblinddoekt 1 knikker mag pakken uit een schaal met daarin 50 rode knikkers (ter waarde van 0€) en 50 blauwe knikkers (ter waarde van 50€). Bij onzekerheid zijn winsten of kansen onbekend, bijvoorbeeld als onbekend is hoeveel van de 100 knikkers rood vs. blauw zijn. Wanneer ze voor zichzelf kiezen zijn de meeste mensen zowel risico-avers en onzekerheids-avers: ze prefereren zekere winst (25€) boven 1 willekeurige knikker uit de "risico schaal", en ook boven 1 willekeurige knikker uit de "onzekerheids-schaal". Echter, uit een aantal recente experimenten blijkt dat risico-aversie verschillend is bij keuzes die mensen maken voor anderen. In sommige studies vindt men verlaagde risico-aversie in "other choices" (Chakravarty et al., 2011; Polman, 2012) terwijl in andere studies juist een tegenovergesteld patroon naar voren komt( Füllbrunn and Luhan, 2015; Kvaløy et al., 2014). Deze verschillen worden onder andere toegeschreven aan de verantwoordelijkheid die mensen voelen voor anderen, en de risico-preferenties die zij op andere projecteren. Op basis van deze resultaten stelden wij onszelf de vraag of mensen in het maken van keuzes voor anderen ook andere preferenties hebben omtrent onzekerheid . Om dit te onderzoeken voerden wij een "lab-in-the-field" experiment uit met N=371 ondernemers in Bangladesh. Deze ondernemers kregen een aantal vignetten gepresenteerd, waarin diverse trainingsprogramma's werd beschreven. In verband met het feit dat er binnen de Islamitische cultuur in Bangladesh een taboe is rondom spelletjes waarbij geld kan worden gewonnen maakten wij in deze context gebruik van hypothetische scenario's. Bij programma A had deelname een zekere winst (een salarisverhoging); bij programma B was deze winst groter maar risicovol (50% kans), en bij programma C was de winst ook groter maar onzeker (onbekende kans). Om een inschatting te maken van de mate van risico- en onzekerheidsaversie werd subjecten gevraagd een voorkeur aan te geven tussen programma's A vs. B, en tussen programma's A vs. C. Aan de hand van deze "preference elicitation" methode kon vervolgens ook een

inschatting gemaakt worden van de onzekerheids-aversie gecontroleerd voor de mate van risico-aversie (Charness and Gneezy, 2010). Sommige subjecten werd gevraagd deze keuzes te maken voor zichzelf, terwijl andere subjecten werd gevraagd dit te doen voor een collega. Uit de resultaten blijkt dat subjecten dezelfde mate van onzekerheids-aversie hebben in keuzes voor zichzelf en keuzes voor anderen. Deze bevinding komt overeen met de resultaten van een recent lab-experiment met studenten in een westerse universiteit. Echter, binnen de subgroep van subjecten bij een vragenlijst na afloop van het experiment aangaven de onzekere kans kleiner in te schatten dan 50% vinden wij een lagere mate van onzekerheids-aversie in keuzes voor anderen. Ten slotte blijkt uit dat subjecten gemiddeld genomen juist *meer* risico-avers zijn bij keuzes voor anderen.

Hoofdstuk 6 (in samenwerking met Mark Sanders en Stephanie Rosenkranz) betreft een lab-experiment over zogenaamde "complexe keuzes" tussen verschillende financiële producten die elk uit diverse kosten-componenten bestaan. Zulke complexiteit speelt bijvoorbeeld bij het kiezen van hypotheken of zorgverzekeringen, waar consumenten – geven hun beperkte rationaliteit vaak suboptimale keuzes maken (Schram and Sonnemans, 2011). Uit de literatuur blijkt dat mensen in zulke situaties geholpen kunnen worden om betere keuzes te maken door eenvoudige aanpassingen in de manier waarop de informatie wordt gepresenteerd (Besedeš et al., 2015; Samek et al., 2016). Wij testen in dit lab-experiment met 158 studenten het effect van een nieuwe "keuze hulp" in zulke situaties. Subjecten kregen een reeks matrices te zien van 4 producten bestaande uit 5 kosten-attributen. De totale kosten per product konden berekend worden op basis van deze attributen. Subjecten konden meer geld verdienen naarmate zij goedkopere producten selecteerde. Echter, gegeven de tijdsdruk waaronder deze berekening gemaakt moest worden was dit een complexe keuze. In de treatment groep kregen studenten een zogenaamd "benchmark product" te zien. Dit bestond uit de gemiddelden per attribuut. Er waren twee versies van deze treatment: in de ene groep werden deze gemiddelden in absolute termen uitgedrukt ("absolute

benchmark treatment"), en voor het andere deel werden de gemiddelden gelijk gesteld aan 100, en de waardes voor de overige producten uitgedrukt als percentage hiervan ("relative benchmark product"). Uit de resultaten blijkt dat de relatieve- maar niet de absolute benchmarks tot een verbetering van de keuzes leidde. Dit suggereert dat consumenten gebaat zijn bij instrumenten waarmee zij eenvoudiger zien hoe kosten van verschillende attributen zich verhouden tot het marktgemiddelde.

Samenvattend pogen deze studies nieuwe inzichten te geven in de oorzaken, gevolgen en interactie van sociale preferenties en beperkte rationaliteit in de context van economische keuzes. Het is mijn hoop dat dit onderzoek een kleine bijdrage levert aan een onbaatzuchtiger wereld waarin mensen betere keuzes maken voor zichzelf en voor anderen.

# English summary

Humans sometimes behave very selfishly, yet in other cases they prefer to help others. These "social preferences" in turn play a key role in many economic phenomena ranging from financial fraud to charitable donations. This dissertation presents a series of lab- and field-experiments about the drivers of social preferences, with a particular focus on bounded rationality, culture and beliefs. These experiments are conducted in a diverse set of populations, ranging from bank employees to entrepreneurs in a developing country to professionals in charitable organizations. The results indicate that humans differ considerably in their prosociality. Furthermore, well-intended "nudges" such as a group-level ethics training can be ineffective or even cause more selfish behavior. In addition, it is shown that people have a tendency to interpret information such to justify their selfish tendencies. However, the results also indicate that small changes to the manner in which information is presented can cause people to make better choices for themselves, and for others. As such, these studies support the notion that there is great value in putting conjectures and assumptions about preferences and behaviors in the economic domain to the experimental test.

# References


Abdellaoui, M., Baillon, A., Placido, L., Wakker, P.P., 2011. The Rich Domain of Uncertainty: Source Functions and Their Experimental Implementation. Am. Econ. Rev. 101, 695–723. https://doi.org/10.1257/aer.101.2.695

Abrantes-Metz, R.M., Kraten, M., Metz, A.D., Seow, G.S., 2012. Libor manipulation? J. Bank. Finance 36, 136–150. https://doi.org/10.1016/j.jbankfin.2011.06.014

Afshin et al., A., 2017. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. N. Engl. J. Med. 13–27.

Agnew, J.R., Szykman, L.R., 2005. Asset Allocation and Information Overload: The Influence of Information Display, Asset Choice, and Investor Experience. J. Behav. Finance 6, 57–70. https://doi.org/10.1207/s15427579jpfm0602_2

Alexander, R.D., 1987. The Biology of Moral Systems. Transaction Publishers.

Alpizar, F., Carlsson, F., Johansson-Stenman, O., 2008. Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica. J. Public Econ. 92, 1047–1060. https://doi.org/10.1016/j.jpubeco.2007.11.004

Andersson, O., Holm, H.J., Tyran, J.-R., Wengström, E., 2013. Risking other people's money: Experimental evidence on bonus schemes, competition, and altruism (Working Paper No. 989). IFN Working Paper.

Andreoni, J., 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. Econ. J. 100, 464–477. https://doi.org/10.2307/2234133

Aquinas, T., 1981. Summa theologiae, trans. Engl. Dominic. N. Y. Christ. Class.

Aristotle, Kennedy, G.A., 2006. On Rhetoric: A Theory of Civic Discourse. Oup Usa.

Aristotle, M., Book, H., 1933. 1045a 8-10 Aristotle in 23 Volumes, Vols. 17, 18, translated by Hugh Tredennick. Lond. William Heine Mann Ltd.

Arnett, J.J., 2008. The neglected 95%: Why American psychology needs to become less American. Am. Psychol. 63, 602–614. https://doi.org/10.1037/0003-066X.63.7.602

Arvedlund, E., 2009. Too Good to Be True: The Rise and Fall of Bernie Madoff. Portfolio.

Ayyagari, M., Beck, T., Demirguc-Kunt, A., 2007. Small and Medium Enterprises Across the Globe. Small Bus. Econ. 29, 415–434. https://doi.org/10.1007/s11187-006-9002-5

Bandiera, O., Barankay, I., Rasul, I., 2011. Field Experiments with Firms. J. Econ. Perspect. 25, 63–82. https://doi.org/10.1257/jep.25.3.63

Banerjee, A., Karlan, D., Zinman, J., 2015. Six Randomized Evaluations of Microcredit: Introduction and Further Steps. Am. Econ. J. Appl. Econ. 7, 1–21. https://doi.org/10.1257/app.20140287

Barham, B.L., Chavas, J.-P., Fitz, D., Salas, V.R., Schechter, L., 2014. The roles of risk and ambiguity in technology adoption. J. Econ. Behav. Organ. 97, 204–218. https://doi.org/10.1016/j.jebo.2013.06.014
Barkow, J.H., Cosmides, L., Tooby, J., 1995. The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press.
Baron, J., 1997. Confusion of Relative and Absolute Risk in Valuation. J. Risk Uncertain. 14, 301–309. https://doi.org/10.1023/A:1007796310463
Bartels, D.M., 2006. Proportion dominance: The generality and variability of favoring relative savings over absolute savings. Organ. Behav. Hum. Decis. Process. 100, 76–95. https://doi.org/10.1016/j.obhdp.2005.10.004
Batson, C.D., Schoenrade, P.A., Pych, V., 1985. Brotherly love or self-concern? Behavioural consequences of religion. Adv. Psychol. Relig. 11, 185–208.
Baumeister, R.F., Heatherton, T.F., Tice, D.M., 1994. Losing control: How and why people fail at self-regulation. Academic Press, San Diego, CA, US.
Baumeister, R.F., Masicampo, E.J., DeWall, C.N., 2009. Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness. Pers. Soc. Psychol. Bull. 35, 260–268. https://doi.org/10.1177/0146167208327217
Baumeister, R.F., Monroe, A.E., 2014. Chapter One - Recent Research on Free Will: Conceptualizations, Beliefs, and Processes, in: Zanna, J.M.O. and M.P. (Ed.), Advances in Experimental Social Psychology. Academic Press, pp. 1–52.
Baumeister, R.F., Vohs, K.D., 2007. Self-Regulation, Ego Depletion, and Motivation. Soc. Personal. Psychol. Compass 1, 115–128. https://doi.org/10.1111/j.1751-9004.2007.00001.x
Beck, T., Demirguc-Kunt, A., 2006. Small and medium-size enterprises: Access to finance as a growth constraint. J. Bank. Finance 30, 2931–2943. https://doi.org/10.1016/j.jbankfin.2006.05.009
Becker, G.M., Degroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behav. Sci. 9, 226–232. https://doi.org/10.1002/bs.3830090304
Bekkers, R., Schuyt, T., Gouwenberg, B., 2017. Geven in Nederland 2017.
Bekkers, R., Wiepking, P., 2011. Who gives? A literature review of predictors of charitable giving Part One: Religion, education, age and socialisation. Volunt. Sect. Rev. 2, 337–365. https://doi.org/10.1332/204080511X6087712
Bekkers, R., Wiepking, P., 2010. Accuracy of self-reports on donations to charitable organizations. Qual. Quant. 45, 1369–1383. https://doi.org/10.1007/s11135-010-9341-9
Ben Zur, H., Breznitz, S.J., 1981. The effect of time pressure on risky choice behavior. Acta Psychol. (Amst.) 47, 89–104. https://doi.org/10.1016/0001-6918(81)90001-9

Bénabou, R., Tirole, J., 2006. Belief in a Just World and Redistributive Politics. Q. J. Econ. 121, 699–746. https://doi.org/10.1162/qjec.2006.121.2.699
Benjamin, D.J., Cesarini, D., Chabris, C.F., Glaeser, E.L., Laibson, D.I., Guðnason, V., Harris, T.B., Launer, L.J., Purcell, S., Smith, A.V., Johannesson, M., Magnusson, P.K.E., Christakis, N.A., Atwood, C.S., Hebert, B.M., Freese, J., Hauser, R.M., Hauser, T.S., Hultman, C.M., Lichtenstein, P., Beauchamp, J.P., Grankvist, A., 2012. The Promises and Pitfalls of Genoeconomics.
Benson, P.L., Donahue, M.J., Erickson, J.A., 1989. Adolescence and religion: A review of the literature from 1970 to 1986. Res. Soc. Sci. Study Relig. 1, 153–81.
Bernhard, H., Fischbacher, U., Fehr, E., 2006. Parochial altruism in humans. Nature 442, 912–915. https://doi.org/10.1038/nature04981
Bertrand, M., Mullainathan, S., 2004. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. Am. Econ. Rev. 94, 991–1013. https://doi.org/10.1257/0002828042002561
Besedeš, T., Deck, C., Sarangi, S., Shor, M., 2015. Reducing Choice Overload without Reducing Choices. Rev. Econ. Stat. 97, 793–802. https://doi.org/10.1162/REST_a_00506
Besedeš, T., Deck, C., Sarangi, S., Shor, M., 2012. Decision-making strategies and performance among seniors. J. Econ. Behav. Organ. 81, 524–533. https://doi.org/10.1016/j.jebo.2011.07.016
Besedeš, T., Deck, C., Sarangi, S., Shor, M., 2010. Age Effects and Heuristics in Decision Making. Rev. Econ. Stat. 94, 580–595. https://doi.org/10.1162/REST_a_00174
Bettman, J.R., Kakkar, P., 1977. Effects of Information Presentation Format on Consumer Information Acquisition Strategies. J. Consum. Res. 3, 233–240. https://doi.org/10.1086/208672
Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., Roberts, J., 2013. Does Management Matter? Evidence from India. Q. J. Econ. 128, 1–51. https://doi.org/10.1093/qje/qjs044
Bloom, N., Van Reenen, J., 2010. Why Do Management Practices Differ across Firms and Countries? J. Econ. Perspect. 24, 203–224. https://doi.org/10.1257/jep.24.1.203
Boatright, J.R., 2013. Swearing to be Virtuous: The Prospects of a Banker's Oath. Rev. Soc. Econ. 71, 140–165. https://doi.org/10.1080/00346764.2013.800305
Bobek, D.D., Roberts, R.W., Sweeney, J.T., 2007. The Social Norms of Tax Compliance: Evidence from Australia, Singapore, and the United States. J. Bus. Ethics 74, 49–64. https://doi.org/10.1007/s10551-006-9219-x
Bonaccio, S., Dalal, R.S., 2006. Advice taking and decision-making: An integrative literature review, and implications for the organizational

sciences. Organ. Behav. Hum. Decis. Process. 101, 127–151. https://doi.org/10.1016/j.obhdp.2006.07.001
Bossaerts, P., Ghirardato, P., Guarnaschelli, S., Zame, W.R., 2010. Ambiguity in Asset Markets: Theory and Experiment. Rev. Financ. Stud. 23, 1325–1359. https://doi.org/10.1093/rfs/hhp106
Bourget, D., Chalmers, D.J., 2013. What do philosophers believe? Philos. Stud. 170, 465–500. https://doi.org/10.1007/s11098-013-0259-7
Bowles, S., 2016. The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens. Yale University Press.
Bowles, S., Gintis, H., 2011. A Cooperative Species: Human Reciprocity and Its Evolution. Princeton University Press.
Brañas-Garza, P., Bucheli, M., Espinosa, M.P., García-Muñoz, T., 2013. Moral Cleansing and Moral Licenses: Experimental Evidence. Econ. Amp Philos.
Bromme, R., Rambow, R., Nückles, M., 2001. Expertise and estimating what other people know: the influence of professional experience and type of knowledge. J. Exp. Psychol. Appl. 7, 317–330. https://doi.org/10.1037/1076-898X.7.4.317
Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon's Mechanical Turk A New Source of Inexpensive, Yet High-Quality, Data? Perspect. Psychol. Sci. 6, 3–5. https://doi.org/10.1177/1745691610393980
Burmeister-Lamp, K., Lévesque, M., Schade, C., 2012. Are entrepreneurs influenced by risk attitude, regulatory focus or both? An experiment on entrepreneurs' time allocation. J. Bus. Ventur. 27, 456–476. https://doi.org/10.1016/j.jbusvent.2011.12.001
Burt, C., Strongman, K., 2005. Use of images in charity advertising: Improving donations and compliance rates. Int. J. Organ. Behav. 8, 571–580.
Cadena, X., Schoar, A., 2011. Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments (Working Paper No. 17020). National Bureau of Economic Research. https://doi.org/10.3386/w17020
Calcagno, R., Monticone, C., 2015. Financial literacy and the demand for financial advice. J. Bank. Finance 50, 363–380. https://doi.org/10.1016/j.jbankfin.2014.03.013
Caliendo, M., Fossen, F.M., Kritikos, A.S., 2009. Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey. Small Bus. Econ. 32, 153–167. https://doi.org/10.1007/s11187-007-9078-6
Callen, M., Isaqzadeh, M., Long, J.D., Sprenger, C., 2014. Violence and Risk Preference: Experimental Evidence from Afghanistan. Am. Econ. Rev. 104, 123–148. https://doi.org/10.1257/aer.104.1.123
Camerer, C.F., Johnson, E.J., 1991. The process-performance paradox in expert judgment - How can experts know so much and predict so badly?, in: Ericsson, K.A., Smith, J. (Eds.), Toward a General Theory of Expertise:

Prospects and Limits. Cambridge University Press, Cambridge, pp. 195–217.
Casler, K., Bickel, L., Hackett, E., 2013. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. Comput. Hum. Behav. 29, 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009
Caviola, L., Faulmüller, N., Everett, J.A.C., Savulescu, J., Kahane, G., 2014. The evaluability bias in charitable giving: Saving administration costs or saving lives? Judgm. Decis. Mak. 9, 303–316.
Cesarini, D., Dawes, C.T., Johannesson, M., Lichtenstein, P., Wallace, B., 2009. Genetic Variation in Preferences for Giving and Risk Taking. Q. J. Econ. 124, 809–842. https://doi.org/10.1162/qjec.2009.124.2.809
Chakravarty, S., Harrison, G.W., Haruvy, E.E., Rutström, E.E., 2011. Are You Risk Averse over Other People's Money? South. Econ. J. 77, 901–913. https://doi.org/10.4284/0038-4038-77.4.901
Charness, G., Dufwenberg, M., 2010. Bare promises: An experiment. Econ. Lett. 107, 281–283. https://doi.org/10.1016/j.econlet.2010.02.009
Charness, G., Fehr, E., 2015. From the lab to the real world. Science 350, 512–513. https://doi.org/10.1126/science.aad4343
Charness, G., Gneezy, U., 2010. Portfolio Choice and Risk Attitudes: An Experiment. Econ. Inq. 48, 133–146. https://doi.org/10.1111/j.1465-7295.2009.00219.x
Charness, G., Rabin, M., 2002. Understanding Social Preferences with Simple Tests. Q. J. Econ. 117, 817–869. https://doi.org/10.1162/003355302760193904
Chernev, A., Böckenholt, U., Goodman, J., 2015. Choice overload: A conceptual review and meta-analysis. J. Consum. Psychol. 2, 333–358. https://doi.org/10.1016/j.jcps.2014.08.002
Choi, J.-K., Bowles, S., 2007. The Coevolution of Parochial Altruism and War. Science 318, 636–640. https://doi.org/10.1126/science.1144237
Cohn, A., Fehr, E., Maréchal, M.A., 2014. Business culture and dishonesty in the banking industry. Nature 516, 86–89. https://doi.org/10.1038/nature13977
Comte, A., 1852. Système de politique positive: Statique sociale. Carilian-Goeury et V. Delmont.
Cramer, J.S., Hartog, J., Jonker, N., Van Praag, C.M., 2002. Low risk aversion encourages the choice for entrepreneurship: an empirical test of a truism. J. Econ. Behav. Organ., Psychological Aspects of Economic Behavior 48, 29–36. https://doi.org/10.1016/S0167-2681(01)00222-0
Croson, R., Sundali, J., 2005. The Gambler's Fallacy and the Hot Hand: Empirical Data from Casinos. J. Risk Uncertain. 30, 195–209. https://doi.org/10.1007/s11166-005-1153-2

Crump, M.J.C., McDonnell, J.V., Gureckis, T.M., 2013. Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research. PLOS ONE 8, e57410. https://doi.org/10.1371/journal.pone.0057410
Dale, A., Strauss, A., 2009. Don't Forget to Vote: Text Message Reminders as a Mobilization Tool. Am. J. Polit. Sci. 53, 787–804. https://doi.org/10.2307/20647951
Danilov, A., Biemann, T., Kring, T., Sliwka, D., 2013. The dark side of team incentives: Experimental evidence on advice quality from financial service professionals. J. Econ. Behav. Organ. 93, 266–272. https://doi.org/10.1016/j.jebo.2013.03.012
Darwin, C., 1859. On the Origin of Species by Means of Natural Selection: Or the Preservation of Favoured Races in the Struggle for Life. Watts.
Deelstra, Y., Nooteboom, S.G., Kohlmann, H.R., van den Berg, J., Innanen, S., 2003. Using knowledge for decision-making purposes in the context of large projects in The Netherlands. Environ. Impact Assess. Rev. 23, 517–541. https://doi.org/10.1016/S0195-9255(03)00070-2
Delfgaauw, J., Dur, R., Sol, J., Verbeke, W., 2013. Tournament Incentives in the Field: Gender Differences in the Workplace. J. Labor Econ. 31, 305–326. https://doi.org/10.1086/667996
Dennett, D.C., 2015. Elbow Room: The Varieties of Free Will Worth Wanting. MIT Press.
Dijksterhuis, A., Bos, M.W., Nordgren, L.F., Baaren, R.B. van, 2006. On Making the Right Choice: The Deliberation-Without-Attention Effect. Science 311, 1005–1007. https://doi.org/10.1126/science.1121629
Dreu, C.K.W.D., Dussel, D.B., Velden, F.S.T., 2015. In intergroup conflict, self-sacrifice is stronger among pro-social individuals, and parochial altruism emerges especially among cognitively taxed individuals. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.00572
Dreu, C.K.W.D., Greer, L.L., Handgraaf, M.J.J., Shalvi, S., Kleef, G.A.V., Baas, M., Velden, F.S.T., Dijk, E.V., Feith, S.W.W., 2010. The Neuropeptide Oxytocin Regulates Parochial Altruism in Intergroup Conflict Among Humans. Science 328, 1408–1411. https://doi.org/10.1126/science.1189047
Duhaime, E.P., 2015. Is the call to prayer a call to cooperate? A field experiment on the impact of religious salience on prosocial behavior. Judgm. Decis. Mak. 10, 593–596.
Dur, R., 2017. Gedragseconomie en beleid KVS Preadviezen 2017.
Egan, M., Matvos, G., Seru, A., 2016. The Market for Financial Adviser Misconduct (Working Paper No. 22050). National Bureau of Economic Research. https://doi.org/10.3386/w22050
Ekelund, J., Johansson, E., Järvelin, M.-R., Lichtermann, D., 2005. Self-employment and risk aversion—evidence from psychological test data. Labour Econ. 12, 649–659. https://doi.org/10.1016/j.labeco.2004.02.009

Ellsberg, D., 1961. Risk, Ambiguity, and the Savage Axioms. Q. J. Econ. 75, 643–669. https://doi.org/10.2307/1884324
Elston, J.A., Audretsch, D.B., 2011. Financing the entrepreneurial decision: an empirical approach using experimental data on risk attitudes. Small Bus. Econ. 36, 209–222. https://doi.org/10.1007/s11187-009-9210-x
Engel, C., 2011a. Dictator games: a meta study. Exp. Econ. 14, 583–610. https://doi.org/10.1007/s10683-011-9283-7
Engel, C., 2011b. Dictator games: a meta study. Exp. Econ. 14, 583–610. https://doi.org/10.1007/s10683-011-9283-7
Engle, W.J.C., Escobal, J., Laszlo, S.C., 2011. Ambiguity Aversion and Portfolio Choice in Small-Scale Peruvian Farming. BE J. Econ. Anal. Policy 11. https://doi.org/10.2202/1935-1682.2331
Engle-Warnick, J., Escobal, J., Laszlo, S., 2007. Ambiguity Aversion as a Predictor of Technology Choice: Experimental Evidence from Peru (SSRN Scholarly Paper No. ID 1077656). Social Science Research Network, Rochester, NY.
Epstein, K., 2006. Crisis mentality. Stanf. Soc. Innov. Rev. Spring.
Ericson, K.M.M., Starc, A., 2016. How product standardization affects choice: Evidence from the Massachusetts Health Insurance Exchange. J. Health Econ. 50, 71–85. https://doi.org/10.1016/j.jhealeco.2016.09.005
Falk, A., Fischbacher, U., 2006. A theory of reciprocity. Games Econ. Behav. 54, 293–315. https://doi.org/10.1016/j.geb.2005.03.001
Falk, A., Meier, S., Zehnder, C., 2013. Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples. J. Eur. Econ. Assoc. 11, 839–852. https://doi.org/10.1111/jeea.12019
Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. Evol. Hum. Behav. 25, 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4
Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137–140. https://doi.org/10.1038/415137a
Fehr, E., Schmidt, K.M., 1999. A Theory of Fairness, Competition, and Cooperation. Q. J. Econ. 114, 817–868. https://doi.org/10.1162/003355399556151
Fetherstonhaugh, D., Slovic, P., Johnson, S., Friedrich, J., 1997. Insensitivity to the Value of Human Life: A Study of Psychophysical Numbing. J. Risk Uncertain. 14, 283–300. https://doi.org/10.1023/A:1007744326393
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10, 171–178. https://doi.org/10.1007/s10683-006-9159-4
Fong, C.M., 2007. Evidence from an Experiment on Charity to Welfare Recipients: Reciprocity, Altruism and the Empathic Responsiveness Hypothesis*. Econ. J. 117, 1008–1024. https://doi.org/10.1111/j.1468-0297.2007.02076.x

Forbes, D.P., 2005. Are some entrepreneurs more overconfident than others? J. Bus. Ventur. 20, 623–640. https://doi.org/10.1016/j.jbusvent.2004.05.001
Forlani, D., Mullins, J.W., 2000. Perceived risks and choices in entrepreneurs' new venture decisions. J. Bus. Ventur. 15, 305–322. https://doi.org/10.1016/S0883-9026(98)00017-2
Fowler, J.H., 2005. Human cooperation: Second-order free-riding problem solved? Nature 437, E8–E8. https://doi.org/10.1038/nature04201
Fox, C.R., Tversky, A., 1995. Ambiguity Aversion and Comparative Ignorance. Q. J. Econ. 110, 585–603. https://doi.org/10.2307/2946693
Frank, D.H., Wertenbroch, K., Maddux, W.W., 2015. Performance pay or redistribution? Cultural differences in just-world beliefs and preferences for wage inequality. Organ. Behav. Hum. Decis. Process. Complete, 160–170. https://doi.org/10.1016/j.obhdp.2015.04.001
Friedrich, J., Barnes, P., Chapin, K., Dawson, I., Garst, V., Kerr, D., 1999. Psychophysical Numbing: When Lives Are Valued Less as the Lives at Risk Increase. J. Consum. Psychol., Ethical Trade-Offs in Consumer Decision Making 8, 277–299. https://doi.org/10.1207/s15327663jcp0803_05
Friesen, L., Earl, P.E., 2015. Multipart tariffs and bounded rationality: An experimental analysis of mobile phone plan choices. J. Econ. Behav. Organ. 116, 239–253. https://doi.org/10.1016/j.jebo.2015.04.016
Füllbrunn, S., Luhan, W.J., 2015. Am I My Peer's Keeper? Social Responsibility in Financial Decision Making (SSRN Scholarly Paper No. ID 2591379). Social Science Research Network, Rochester, NY.
Gaudecker, H.-M. von, Soest, A. van, Wengström, E., 2012. Experts in experiments. J. Risk Uncertain. 45, 159–190. https://doi.org/10.1007/s11166-012-9151-7
Gigerenzer, G., Gaissmaier, W., 2011. Heuristic Decision Making. Annu. Rev. Psychol. *62*, 451-82.https://doi.org/10.1146/annurev-ps-62-040111-200001
Gigerenzer, G., Selten, R., 2002. Bounded Rationality: The Adaptive Toolbox. MIT Press.
Gigerenzer, G., Todd, P.M., Group, A.R., 1999. Simple Heuristics That Make Us Smart. Oxford University Press.
Gino, F., 2008. Do we listen to advice just because we paid for it? The impact of advice cost on its use. Organ. Behav. Hum. Decis. Process. 107, 234–245. https://doi.org/10.1016/j.obhdp.2008.03.001
Gino, F., Ayal, S., Ariely, D., 2009. Contagion and Differentiation in Unethical Behavior: The Effect of One Bad Apple on the Barrel. Psychol. Sci. 20, 393–398. https://doi.org/10.1111/j.1467-9280.2009.02306.x
Gino, F., Moore, D.A., 2007. Effects of task difficulty on use of advice. J. Behav. Decis. Mak. 20, 21–35. https://doi.org/10.1002/bdm.539

Gino, F., Pierce, L., 2009. The abundance effect: Unethical behavior in the presence of wealth. Organ. Behav. Hum. Decis. Process. 109, 142–155. https://doi.org/10.1016/j.obhdp.2009.03.003
Gintis, H., 2007. Unifying the behavioral sciences II. Behav. Brain Sci. 30, 45–53. https://doi.org/10.1017/S0140525X0700088X
Gintis, H., Bowles, S., Boyd, R., Fehr, E., 2003. Explaining altruistic behavior in humans. Evol. Hum. Behav. 24, 153–172. https://doi.org/10.1016/S1090-5138(02)00157-5
Glimcher, P.W., Fehr, E., 2013. Neuroeconomics: Decision Making and the Brain. Academic Press.
Gloss, A., Pollack, J.M., Ward, M.K., 2017. A risky shift? An exploration of the measurement equivalence of entrepreneurial attitudes and entrepreneurial orientation across socioeconomic gradients. J. Bus. Ventur. Insights 7, 32–37. https://doi.org/10.1016/j.jbvi.2017.01.003
Gneezy, U., Keenan, E.A., Gneezy, A., 2014. Avoiding overhead aversion in charity. Science 346, 632–635. https://doi.org/10.1126/science.1253932
Goldsmith, D.J., 2000. Soliciting advice: The role of sequential placement in mitigating face threat. Commun. Monogr. 67, 1–19. https://doi.org/10.1080/03637750009376492
Goldsmith, D.J., Fitch, K., 1997. The Normative Context of Advice as Social Support. Hum. Commun. Res. 23, 454–476. https://doi.org/10.1111/j.1468-2958.1997.tb00406.x
Greifeneder, R., Scheibehenne, B., Kleber, N., 2010. Less may be more when choosing is difficult: Choice complexity and too much choice. Acta Psychol. (Amst.) 133, 45–50. https://doi.org/10.1016/j.actpsy.2009.08.005
Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1, 114–125. https://doi.org/10.1007/s40881-015-0004-4
Gunia, B.C., Wang, L., Huang, L., Wang, J., Murnighan, J.K., 2012. Contemplation and Conversation: Subtle Influences on Moral Decision Making. Acad. Manage. J. 55, 13–33. https://doi.org/10.5465/amj.2009.0873
Hackethal, A., Haliassos, M., Jappelli, T., 2012. Financial advisors: A case of babysitters? J. Bank. Finance 36, 509–524. https://doi.org/10.1016/j.jbankfin.2011.08.008
Hai Yap Teoh, See Liang Foo, 1997. Moderating effects of tolerance for ambiguity and risktaking propensity on the role conflict-perceived performance relationship: Evidence from singaporean entrepreneurs. J. Bus. Ventur. 12, 67–81. https://doi.org/10.1016/S0883-9026(96)00035-3
Hamermesh, D.S., 2013. Six Decades of Top Economics Publishing: Who and How? J. Econ. Lit. 51, 162–172. https://doi.org/10.1257/jel.51.1.162
Hamilton, W.D., 1964. The genetical evolution of social behaviour. II. J. Theor. Biol. 7, 17–52. https://doi.org/10.1016/0022-5193(64)90039-6

Harris, S., 2012. Free Will. Simon and Schuster.
Haselton, M.G., Nettle, D., Andrews, P.W., 2015. The Evolution of Cognitive Bias, in: Buss, D.M. (Ed.), The Handbook of Evolutionary Psychology. John Wiley & Sons, Inc., pp. 724–746. https://doi.org/10.1002/9780470939376.ch25
Heiss, F., Leive, A., McFadden, D., Winter, J., 2013. Plan selection in Medicare Part D: Evidence from administrative data. J. Health Econ. 32, 1325–1344. https://doi.org/10.1016/j.jhealeco.2013.06.006
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., 2001. In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. Am. Econ. Rev. 91, 73–78.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N.S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F.W., Patton, J.Q., Tracer, D., 2005. "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. Behav. Brain Sci. 28, 795–815. https://doi.org/10.1017/S0140525X05000142
Henrich, J., Heine, S.J., Norenzayan, A., 2010a. The weirdest people in the world? Behav. Brain Sci. 33, 61-83; discussion 83-135. https://doi.org/10.1017/S0140525X0999152X
Henrich, J., Heine, S.J., Norenzayan, A., 2010b. The weirdest people in the world? Behav. Brain Sci. 33, 61–83. https://doi.org/10.1017/S0140525X0999152X
Henrich, J., Heine, S.J., Norenzayan, A., 2010c. Most people are not WEIRD. Nature 466, 466029a. https://doi.org/10.1038/466029a
Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial Punishment Across Societies. Science 319, 1362–1367. https://doi.org/10.1126/science.1153808
Higuchi, Y., Nam, V.H., Sonobe, T., 2015. Sustained impacts of Kaizen training. J. Econ. Behav. Organ. 120, 189–206. https://doi.org/10.1016/j.jebo.2015.10.009
Hinds, P.J., 1999. The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. J. Exp. Psychol. Appl. 205–221.
Holm, H.J., Opper, S., Nee, V., 2013. Entrepreneurs Under Uncertainty: An Economic Experiment in China. Manag. Sci. 59, 1671–1687. https://doi.org/10.1287/mnsc.1120.1670
Holt, C.A., Laury, S., 2002. Risk Aversion and Incentive Effects (SSRN Scholarly Paper No. ID 893797). Social Science Research Network, Rochester, NY.
Horton, J.J., Rand, D.G., Zeckhauser, R.J., 2011. The online laboratory: conducting experiments in a real labor market. Exp. Econ. 14, 399–425. https://doi.org/10.1007/s10683-011-9273-9

Hsee, C.K., Zhang, J., 2010. General Evaluability Theory. Perspect. Psychol. Sci. 5, 343–355. https://doi.org/10.1177/1745691610374586
Huberman, G., Iyengar, S.S., Jiang, W., 2007. Defined Contribution Pension Plans: Determinants of Participation and Contributions Rates. J. Financ. Serv. Res. 31, 1–32. https://doi.org/10.1007/s10693-007-0003-6
Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., Platt, M.L., 2006. Neural Signatures of Economic Preferences for Risk and Ambiguity. Neuron 49, 765–775. https://doi.org/10.1016/j.neuron.2006.01.024
Hume, D., 1758. Essays, Moral, Political, and Literary [WWW Document]. URL http://www.econlib.org/library/LFBooks/Hume/hmMPL.html
Hung, A., Yoong, J., 2010. Asking for Help: Survey and Experimental Evidence on Financial Advice and Behavior Change (SSRN Scholarly Paper No. ID 1532993). Social Science Research Network, Rochester, NY.
Inderst, R., Ottaviani, M., 2012. Financial Advice. J. Econ. Lit. 50, 494–512. https://doi.org/10.1257/jel.50.2.494
Inderst, R., Ottaviani, M., 2009. Misselling through Agents. Am. Econ. Rev. 99, 883–908. https://doi.org/10.1257/aer.99.3.883
Iyengar, S., Kamenica, E., 2006. Choice Overload and Simplicity Seeking 87.
Iyengar, S.S., Lepper, M.R., 2000. When Choice is Demotivating: Can One Desire Too Much of a Good Thing? J. Pers. Soc. Psychol. 79, 995–1006.
Jarvenpaa, S.L., 1990. Graphic displays in decision making — the visual salience effect. J. Behav. Decis. Mak. 3, 247–262. https://doi.org/10.1002/bdm.3960030403
Jenni, K., Loewenstein, G., 1997. Explaining the Identifiable Victim Effect. J. Risk Uncertain. 14, 235–257. https://doi.org/10.1023/A:1007740225484
Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. Admin. Sci Quart. 235–257. *24*(4), 602-611
Jonas, E., Schulz-hardt, S., Frey, D., Thelen, N., 2001. Confirmation Bias in Sequential Information Search After Preliminary Decisions: An Expansion of Dissonance Theoretical Research on Selective Exposure to Information. J. Pers. Soc. Psychol. 80, 557–571.
Jones, M., Sugden, R., 2001. Positive Confirmation Bias In The Acquisition Of Information. Theory Decis. 50, 59–99.
Kahneman, D., 2003a. Maps of Bounded Rationality: Psychology for Behavioral Economics. Am. Econ. Rev. 93, 1449–1475.
Kahneman, D., Klein, G., 2009. Conditions for intuitive expertise: a failure to disagree. Am. Psychol. 64, 515–526. https://doi.org/10.1037/a0016755
Kahneman, D., Knetsch, J.L., Thaler, R.H., 1991. Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. J. Econ. Perspect. 5, 193–206. https://doi.org/10.2307/1942711
Kahneman, D., Tversky, A., 1982. The Psychology of Preferences. Sci. Am. 246, 160–173. https://doi.org/10.1038/scientificamerican0182-160

Kan, K., Tsai, W.-D., 2006. Entrepreneurship and Risk Aversion. Small Bus. Econ. 26, 465–474. https://doi.org/10.1007/s11187-005-5603-7
Karlan, D., McConnell, M., Mullainathan, S., Zinman, J., 2016. Getting to the Top of Mind: How Reminders Increase Saving. Manag. Sci. 62, 3393–3411. https://doi.org/10.1287/mnsc.2015.2296
Karlan, D., Valdivia, M., 2010. Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions. Rev. Econ. Stat. 93, 510–527. https://doi.org/10.1162/REST_a_00074
Kaustia, M., Knüpfer, S., 2008a. Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. J. Finance 63, 2679–2702. https://doi.org/10.1111/j.1540-6261.2008.01411.x
Kaustia, M., Knüpfer, S., 2008b. Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. J. Finance 63, 2679–2702. https://doi.org/10.1111/j.1540-6261.2008.01411.x
Kirsch, I.S., And Others, 1993. Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey. U.
Kish-Gephart, J.J., Harrison, D.A., Treviño, L.K., 2010. Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. J. Appl. Psychol. 95, 1–31. https://doi.org/10.1037/a0017103
Knight, F.H., 2012. Risk, Uncertainty and Profit. Courier Corporation.
König-Kersting, C., Trautmann, S.T., 2016a. Ambiguity attitudes in decisions for others. Econ. Lett. 146, 126–129. https://doi.org/10.1016/j.econlet.2016.07.036
Konrath, S., Falk, E., Fuhrel-Forbis, A., Liu, M., Swain, J., Tolman, R., Cunningham, R., Walton, M., 2015. Can Text Messages Increase Empathy and Prosocial Behavior? The Development and Initial Validation of Text to Connect. PLoS ONE 10, e0137585. https://doi.org/10.1371/journal.pone.0137585
Koudstaal, M., Sloof, R., van Praag, M., 2015. Risk, Uncertainty, and Entrepreneurship: Evidence from a Lab-in-the-Field Experiment. Manag. Sci. 62, 2897–2915. https://doi.org/10.1287/mnsc.2015.2249
Krajbich, I., Bartling, B., Hare, T., Fehr, E., 2015. Rethinking fast and slow based on a critique of reaction-time reverse inference. Nat. Commun. 6. https://doi.org/10.1038/ncomms8455
Krueger, F., Hoffman, M., Walter, H., Grafman, J., 2014. An fMRI investigation of the effects of belief in free will on third-party punishment. Soc. Cogn. Affect. Neurosci. 9, 1143–1149. https://doi.org/10.1093/scan/nst092
Kruger, J., Dunning, D., 1999. Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-assessments. J. Pers. Soc. Psychol. 77, 1121–1134.
Kunda, Z., 1990. The case for motivated reasoning. Psychol. Bull. 480–498.

Kushnir, T., 2012. Developing a concept of choice. Adv. Child Dev. Behav. 43, 193–218.
Kvaløy, O., Eriksen, K., Luzuriaga, M., 2014. Risk-taking with Other People's Money (No. 2014/21), UiS Working Papers in Economics and Finance. University of Stavanger.
Levitt, S.D., List, J.A., 2009. Field experiments in economics: The past, the present, and the future. Eur. Econ. Rev. 53, 1–18. https://doi.org/10.1016/j.euroecorev.2008.12.001
Levitt, S.D., Neckermann, S., 2014. What field experiments have and have not taught us about managing workers. Oxf. Rev. Econ. Policy 30, 639–657. https://doi.org/10.1093/oxrep/grv003
Libet, B., 1993. Unconscious cerebral initiative and the role of conscious will in voluntary action, in: Neurophysiology of Consciousness, Contemporary Neuroscientists. Birkhäuser Boston, pp. 269–306. https://doi.org/10.1007/978-1-4612-0355-1_16
Lo, A.W., 2015. The Gordon Gekko Effect: The Role of Culture in the Financial Industry (Working Paper No. 21267). National Bureau of Economic Research. https://doi.org/10.3386/w21267
Lurie, N.H., Mason, C.H., 2007. Visual Representation: Implications for Decision Making. J. Mark. 71, 160–177. https://doi.org/10.1509/jmkg.71.1.160
Lusardi, A., Mitchell, O.S., 2011. Financial Literacy and Planning: Implications for Retirement Wellbeing (Working Paper No. 17078). National Bureau of Economic Research. https://doi.org/10.3386/w17078
Maas, K., Liket, K., 2011. Talk the Walk: Measuring the Impact of Strategic Philanthropy. J. Bus. Ethics 100, 445–464. https://doi.org/10.1007/s10551-010-0690-z
MacKenzie, M.J., Vohs, K.D., Baumeister, R.F., 2014. You Didn't Have to Do That Belief in Free Will Promotes Gratitude. Pers. Soc. Psychol. Bull. 0146167214549322. https://doi.org/10.1177/0146167214549322
Malthus, T.R., 1798. An Essay on the Principle of Population: The First Edition (1798) with Introduction and Bibliography. W. Pickering.
Mandel, N., Johnson, E.J., 2002. When Web Pages Influence Choice: Effects of Visual Primes on Experts and Novices. J. Consum. Res. 29, 235–245. https://doi.org/10.1086/341573
Mandeville, B., 1714. The Fable of the Bees.
Marteau, T.M., 1989. Framing of information: Its influence upon decisions of doctors and patients. Br. J. Soc. Psychol. 28, 89–94. https://doi.org/10.1111/j.2044-8309.1989.tb00849.x
Mazar, N., Amir, O., Ariely, D., 2008. The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. J. Mark. Res. 45, 633–644.
McKay, R., Efferson, C., Whitehouse, H., Fehr, E., 2011. Wrath of God: religious primes and punishment. Proc. R. Soc. Lond. B Biol. Sci. 278, 1858–1863. https://doi.org/10.1098/rspb.2010.2125

McKenzie, D., Woodruff, C., 2014. What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World? World Bank Res. Obs. 29, 48–82. https://doi.org/10.1093/wbro/lkt007
Mencius, 2003. . Chinese University Press.
Milinski, M., Semmann, D., Krambeck, H.-J., 2002. Donors to charity gain in both indirect reciprocity and political reputation. Proc. R. Soc. B Biol. Sci. 269, 881–883. https://doi.org/10.1098/rspb.2002.1964
Monroe, A.E., Brady, G., Malle, B.F., 2016. This Isn't the Free Will Worth Looking For General Free Will Beliefs Do Not Influence Moral Judgments, Agent-Specific Choice Ascriptions Do. Soc. Psychol. Personal. Sci. 1948550616667616. https://doi.org/10.1177/1948550616667616
Montague, P.R., 2008. Free will. Curr. Biol. 18, R584–R585. https://doi.org/10.1016/j.cub.2008.04.053
Mullainathan, S., Noeth, M., Schoar, A., 2012. The Market for Financial Advice: An Audit Study (Working Paper No. 17929). National Bureau of Economic Research. https://doi.org/10.3386/w17929
Mullainathan, S., Thaler, R.H., 2000. Behavioral Economics (Working Paper No. 7948). National Bureau of Economic Research. https://doi.org/10.3386/w7948
Nahmias, E., Shepard, J., Reuter, S., 2014. It's OK if 'my brain made me do it': People's intuitions about free will and neuroscientific prediction. Cognition 133, 502–516. https://doi.org/10.1016/j.cognition.2014.07.009
Neumark, D., Bank, R.J., Nort, V., D, K., 1996. Sex Discrimination in Restaurant Hiring: An Audit Study. Q. J. Econ. 111, 915–941. https://doi.org/10.2307/2946676
Nickerson, R.S., 1998. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Rev. Gen. Psychol. 2, 175–220.
Northcraft, G.B., Neale, M.A., 1987. Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. Organ. Behav. Hum. Decis. Process. 39, 84–97. https://doi.org/10.1016/0749-5978(87)90046-X
Nowak, M.A., Sigmund, K., 2005. Evolution of Indirect Reciprocity [WWW Document]. URL http://pure.iiasa.ac.at/7763/ (accessed 9.27.17).
Oosterbeek, H., Webbink, D., 2011. Does Studying Abroad Induce a Brain Drain? Economica 78, 347–366. https://doi.org/10.1111/j.1468-0335.2009.00818.x
Organisation for Economic Co-operation and Development, 2014. CleanGovBiz: Integrity in practice.
Paolacci, G., Chandler, J., Ipeirotis, P.G., 2010. Running experiments on Amazon Mechanical Turk. Judgm. Decis. Mak. 5, 411–419.
Pareek, S., 2014. The 73-Year Old Librarian Who Has Been Donating Every Rupee He Earned To The Poor For 30 Years! TheBetterIndia.com.

Payne, J.W., Bettman, J.R., Johnson, E.J., 1988. Adaptive Strategy Selection in Decision Making. J. Exp. Psychol. Learn. Mem. Cogn. 14, 534–552.
Peters, E., Dieckmann, N.F., Västfjäll, D., Mertz, C.K., Slovic, P., Hibbard, J.H., 2009. Bringing meaning to numbers: the impact of evaluative categories on decisions. J. Exp. Psychol. Appl. 15, 213–227. https://doi.org/10.1037/a0016978
Peterson, R.A., 2001. On the Use of College Students in Social Science Research: Insights from a Second-Order Meta-analysis. J. Consum. Res. 28, 450–461. https://doi.org/10.1086/323732
Phillips, J.W., Putnam, R.D., Campbell, D.E., Garrett, S.R., 2011. American Grace: How Religion Divides and Unites Us. JSTOR.
Pierce, L., Balasubramanian, P., 2015. Behavioral field evidence on psychological and social factors in dishonesty and misconduct. Curr. Opin. Psychol., Morality and ethics 6, 70–76. https://doi.org/10.1016/j.copsyc.2015.04.002
Pierce, L., Snyder, J., 2008. Ethical Spillovers in Firms: Evidence from Vehicle Emissions Testing. Manag. Sci. 54, 1891–1903. https://doi.org/10.1287/mnsc.1080.0927
Pollmann, M.M.H., Potters, J., Trautmann, S.T., 2014. Risk taking by agents: The role of ex-ante and ex-post accountability. Econ. Lett. 123, 387–390. https://doi.org/10.1016/j.econlet.2014.04.004
Polman, E., 2012. Self–other decision making and loss aversion. Organ. Behav. Hum. Decis. Process. 119, 141–150. https://doi.org/10.1016/j.obhdp.2012.06.005
Polman, E., 2010. Differences Between Making Decisions For The Self Versus For Others: A Reversal Of The Choice Overload Effect.
Pronin, E., Kugler, M.B., 2010. People believe they have more free will than others. Proc. Natl. Acad. Sci. 107, 22469–22474. https://doi.org/10.1073/pnas.1012046108
Raihani, N.J., Bshary, R., 2012. A positive effect of flowers rather than eye images in a large-scale, cross-cultural dictator game. Proc R Soc B rspb20120758. https://doi.org/10.1098/rspb.2012.0758
Raihani, N.J., Mace, R., Lamba, S., 2013. The Effect of $1, $5 and $10 Stakes in an Online Dictator Game. PLOS ONE 8, e73131. https://doi.org/10.1371/journal.pone.0073131
Rammstedt, B., John, O.P., 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. J. Res. Personal. 41, 203–212. https://doi.org/10.1016/j.jrp.2006.02.001
Reyna, V.F., Brainerd, C.J., 2007. The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. Learn. Individ. Differ. 17, 147–159. https://doi.org/10.1016/j.lindif.2007.03.010

Reynolds, D.B., Joseph, J., Sherwood, R., 2011. Risky Shift Versus Cautious Shift: Determining Differences In Risk Taking Between Private And Public Management Decision-Making. J. Bus. Econ. Res. JBER 7. https://doi.org/10.19030/jber.v7i1.2251
Ritov, I., Baron, J., 2011. Joint presentation reduces the effect of emotion on evaluation of public actions. Cogn. Emot. 25, 657–675. https://doi.org/10.1080/02699931.2010.512512
Ritov, I., Baron, J., 1990. Reluctance to vaccinate: Omission bias and ambiguity. J. Behav. Decis. Mak. 3, 263–277. https://doi.org/10.1002/bdm.3960030404
Ritzema, R.J., 1979. Religiosity and altruism: Faith without works? J. Psychol. Theol. 7, 105–113.
Robb, C.A., Babiarz, P., Woodyard, A., 2012. The Demand for Financial Professionals' Advice: The Role of Financial Knowledge, Satisfaction, and Confidence. Financ. Serv. Rev. 21, 291.
Robson, P.J.A., Bennett, R.J., 2000. SME Growth: The Relationship with Business Advice and External Collaboration. Small Bus. Econ. 15, 193–208. https://doi.org/10.1023/A:1008129012953
Samek, A., Hur, I., Kim, S.-H., Yi, J.S., 2016. An experimental study of the decision process with interactive technology. J. Econ. Behav. Organ. 130, 20–32. https://doi.org/10.1016/j.jebo.2016.06.004
Sargeant, A., 1999. Charitable Giving: Towards a Model of Donor Behaviour. J. Mark. Manag. 15, 215–238. https://doi.org/10.1362/026725799784870351
Sarkissian, H., Chatterjee, A., De Brigard, F., Knobe, J., Nichols, S., Sirker, S., 2010. Is Belief in Free Will a Cultural Universal? Mind Lang. 25, 346–358. https://doi.org/10.1111/j.1468-0017.2010.01393.x
Schopenhauer, A., 1839. Prize Essay on the Freedom of the Will, ed. G. Zoller. Cambridge: Cambridge University Press.
Schram, A., Sonnemans, J., 2011. How individuals choose health insurance: An experimental analysis. Eur. Econ. Rev. 55, 799–819. https://doi.org/10.1016/j.euroecorev.2011.01.001
Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2007. The Constructive, Destructive, and Reconstructive Power of Social Norms. Psychol. Sci. 18, 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x
Schulz, J.F., Fischbacher, U., Thöni, C., Utikal, V., 2014. Affect and fairness: Dictator games under cognitive load. J. Econ. Psychol., From Dual Processes to Multiple Selves: Implications for Economic Behavior 41, 77–87. https://doi.org/10.1016/j.joep.2012.08.007
Shanteau, J., 1992. Competence in experts: The role of task characteristics. Organ. Behav. Hum. Decis. Process., Experts and Expert Systems 53, 252–266. https://doi.org/10.1016/0749-5978(92)90064-E
Shariff, A.F., Greene, J.D., Karremans, J.C., Luguri, J.B., Clark, C.J., Schooler, J.W., Baumeister, R.F., Vohs, K.D., 2014. Free Will and Punishment A

Mechanistic View of Human Nature Reduces Retribution. Psychol. Sci. 0956797614534693. https://doi.org/10.1177/0956797614534693
Shariff, A.F., Norenzayan, A., 2007. God Is Watching You: Priming God Concepts Increases Prosocial Behavior in an Anonymous Economic Game. Psychol. Sci. 18, 803–809.
Sharman, J.C., 2010. Shopping for Anonymous Shell Companies: An Audit Study of Anonymity and Crime in the International Financial System. J. Econ. Perspect. 24, 127–140. https://doi.org/10.1257/jep.24.4.127
Shu, L.L., Mazar, N., Gino, F., Ariely, D., Bazerman, M.H., 2012. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. Proc. Natl. Acad. Sci. 109, 15197–15200. https://doi.org/10.1073/pnas.1209746109
Simon, H.A., 1982. Models of Bounded Rationality: Empirically grounded economic reason. MIT Press.
Slovic, P., 2007. "If I look at the mass I will never act": Psychic numbing and genocide. Judgm. Decis. Mak. 2, 79–95.
Slovic, P., Finucane, M.L., Peters, E., MacGregor, D.G., 2007. The affect heuristic. Eur. J. Oper. Res. 177, 1333–1352. https://doi.org/10.1016/j.ejor.2005.04.006
Small, D.A., Loewenstein, G., Slovic, P., 2007. Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. Organ. Behav. Hum. Decis. Process. 102, 143–153. https://doi.org/10.1016/j.obhdp.2006.01.005
Smilansky, S., 2000. Free will and illusion. OUP Oxford.
Smith, A., 1759. The Theory of Moral Sentiments.
Soll, J.B., Keeney, R.L., Larrick, R.P., 2013. Consumer Misunderstanding of Credit Card Use, Payments, and Debt: Causes and Solutions. J. Public Policy Mark. 32, 66–81. https://doi.org/10.1509/jppm.11.061
Stillman, T.F., Baumeister, R.F., Vohs, K.D., Lambert, N.M., Fincham, F.D., Brewer, L.E., 2010. Personal Philosophy and Personnel Achievement: Belief in Free Will Predicts Better Job Performance. Soc. Psychol. Personal. Sci. 1, 43–50. https://doi.org/10.1177/1948550609351600
Stoop, J., Noussair, C.N., van Soest, D., 2012. From the Lab to the Field: Cooperation among Fishermen. J. Polit. Econ. 120, 1027–1056. https://doi.org/10.1086/669253
Stutzer, A., Goette, L., Zehnder, M., 2011. Active Decisions and Prosocial Behaviour: A Field Experiment on Blood Donation. Econ. J. 121, F476–F493.
Sun, Q., Liu, Y., Zhang, H., Lu, J., 2017. Increased social distance makes people more risk-neutral. J. Soc. Psychol. 157, 502–512. https://doi.org/10.1080/00224545.2016.1242471

Sweeny, K., Melnyk, D., Miller, W., Shepperd, J.A., 2010. Information Avoidance: Who, What, When, and Why. Rev. Gen. Psychol. 14, 340–353. https://doi.org/10.1037/a0021288
Thaler, R., 2017. From Cashews to Nudges: The Evolution of Behavioral Economics.
Thaler, R.H., 2015. Misbehaving: The Making of Behavioral Economics. W.W. Norton.
Thaler, R.H., 1988. Anomalies: The Ultimatum Game. J. Econ. Perspect. 2, 195–206. https://doi.org/10.1257/jep.2.4.195
Thaler, R.H., Sunstein, C.R., 2008. Nudge: Improving Decisions about Health, Wealth, and Happiness. Yale University Press.
Thaler, R.H., Sunstein, C.R., 2003. Libertarian Paternalism. Am. Econ. Rev. 93, 175–179.
Thaler, R.H., Sunstein, C.R., Balz, J.P., 2014. Choice Architecture (SSRN Scholarly Paper No. ID 2536504). Social Science Research Network, Rochester, NY.
Trautmann, S.T., van de Kuilen, G., 2015. Ambiguity Attitudes, in: Keren, G., Wu, G. (Eds.), The Wiley Blackwell Handbook of Judgment and Decision Making. John Wiley & Sons, Ltd, Chichester, UK, pp. 89–116. https://doi.org/10.1002/9781118468333.ch3
Trautmann, S.T., Vieider, F.M., Wakker, P.P., 2008. Causes of ambiguity aversion: Known versus unknown preferences. J. Risk Uncertain. 36, 225–243. https://doi.org/10.1007/s11166-008-9038-9
Treviño, L.K., Nieuwenboer, N.A. den, Kish-Gephart, J.J., 2014. (Un)Ethical Behavior in Organizations. Annu. Rev. Psychol. 65, 635–660. https://doi.org/10.1146/annurev-psych-113011-143745
Trivers, R.L., 1971. The Evolution of Reciprocal Altruism. Q. Rev. Biol. 46, 35–57. https://doi.org/10.1086/406755
Trope, Y., Liberman, N., 2010. Construal-level Theory of Psychological Distance. Psychol. Rev. 117, 440–463.
Twenge, J.M., Zhang, L., Im, C., 2004. It's Beyond My Control: A Cross-Temporal Meta-Analysis of Increasing Externality in Locus of Control, 1960-2002. Personal. Soc. Psychol. Rev. 8, 308–319. https://doi.org/10.1207/s15327957pspr0803_5
Vanberg, C., 2008. Why Do People Keep Their Promises? An Experimental Test of Two Explanations1. Econometrica 76, 1467–1480. https://doi.org/10.3982/ECTA7673
Vohs, K.D., Heatherton, T.F., 2000. Self-Regulatory Failure: A Resource-Depletion Approach. Psychol. Sci. 11, 249–254. https://doi.org/10.1111/1467-9280.00250
Vohs, K.D., Schooler, J.W., 2008. The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating. Psychol. Sci. 19, 49–54. https://doi.org/10.1111/j.1467-9280.2008.02045.x

Vuletić, D., 2015. How Effective are Reminders and Frames in Incentivizing Blood Donations? (SSRN Scholarly Paper No. ID 2700608). Social Science Research Network, Rochester, NY.
Watson, J.D., Crick, F.H., 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171, 737–738.
Wiss, J., Andersson, D., Slovic, P., Västfjäll, D., Tinghög, G., 2015. The influence of identifiability and singularity in moral decision making. Judgm. Decis. Mak. Tallahass. 10, 492–502.
World Bank, 2015. World Development Report [WWW Document]. URL http://www.worldbank.org/en/publication/wdr2015 (accessed 9.28.17).
Yaniv, I., 2004a. Receiving other people's advice: Influence and benefit. Organ. Behav. Hum. Decis. Process. 93, 1–13. https://doi.org/10.1016/j.obhdp.2003.08.002
Yaniv, I., 2004b. The Benefit of Additional Opinions. Curr. Dir. Psychol. Sci. 13, 75–78. https://doi.org/10.1111/j.0963-7214.2004.00278.x
Yaniv, I., Kleinberger, E., 2000. Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation. Organ. Behav. Hum. Decis. Process. 83, 260–281. https://doi.org/10.1006/obhd.2000.2909
Youniss, J., McLellan, J.A., Yates, M., 1999. Religion, community service, and identity in American youth. J. Adolesc. 22, 243–253. https://doi.org/10.1006/jado.1999.0214
Ziegler, F.V., Tunney, R.J., 2012. Decisions for others become less impulsive the further away they are on the family tree., PLoS ONE 7, e49479–e49479. https://doi.org/10.1371/journal.pone.0049479, 10.1371/journal.pone.0049479
Zikmund-Fisher, B.J., Sarr, B., Fagerlin, A., Ubel, P.A., 2006. A Matter of Perspective: Choosing for Others Differs from Choosing for Yourself in Making Treatment Decisions. J. Gen. Intern. Med. 21, 618–622. https://doi.org/10.1111/j.1525-1497.2006.00410.x

# Curriculum Vitae

Job Harms was born in 1985 in Amsterdam, the Netherlands. He obtained a BSc in Liberal Arts and Sciences at University College Utrecht in 2006. He then pursued a MSc in Economics at VU University Amsterdam. He conducted field research in Kenya in partnership with the International Labor Organization, obtaining his degree in 2010. Between 2011-13 he worked for Triodos Facet, focusing on financial services for consumers in developing countries. In 2013 he started his doctoral research at the Erasmus School of Economics. He is currently a postdoctoral researcher at the Faculty for Behavioral and Social Sciences at Groningen University.

# Portfolio

**PhD training**

| | | |
|---|---|---|
| ***Courses*** | | |
| | Experimental economics, Tinbergen Institute summer course | 2015 |
| | Experimental design in ePRIME, Erasmus Institute of Psychology | 2014 |
| | Policy evaluation seminars, Erasmus School of Economics | 2014 |
| | Data analysis in R, ERIM summer course | 2014 |
| ***Seminars and workshops*** | | |
| | Arne Ryde Workshop, Lund University "Better Bankers: An Ethics Intervention in the Field" | 2017 |
| | Erasmus University Workshop on Economics & Law: "Free will belief and altruism" | 2017 |
| ***Conference presentations*** | | |
| | Science of Philanthropy Initiative, University of Chicago , "Free will belief and altruism" | 2017 |
| | IABS Conference, University of Amsterdam; "Better Bankers: An Ethics Intervention in the Field" | 2017 |
| | Economic Science Association, Hebrew University of Jerusalem; "Reference Points and Advice in Complex Choices" | 2016 |
| | Economic Science Association, Heidelberg University; "Self-other differences in ambiguity aversion" | 2015 |
| ***Teaching*** | | |
| | Teaching assistant, EFR Involve research project | 2015-17 |

# The ERIM PhD Series

The ERIM PhD Series contains PhD dissertations in the field of Research in Management defended at Erasmus University Rotterdam and supervised by senior researchers affiliated to the Erasmus Research Institute of Management (ERIM). All dissertations in the ERIM PhD Series are available in full text through the ERIM Electronic Series Portal: http://repub.eur.nl/pub. ERIM is the joint research institute of the Rotterdam School of Management (RSM) and the Erasmus School of Economics at the Erasmus University Rotterdam (EUR).

**Dissertations in the last four years**

Abbink, E.J., *Crew Management in Passenger Rail Transport*,
Promotors: Prof. L.G. Kroon & Prof. A.P.M. Wagelmans, EPS-2014-325-LIS,
http://repub.eur.nl/pub/76927

Acar, O.A., *Crowdsourcing for Innovation: Unpacking Motivational, Knowledge and Relational Mechanisms of Innovative Behavior in Crowdsourcing Platforms*,
Promotor: Prof. J.C.M. van den Ende, EPS-2014-321-LIS, http://repub.eur.nl/pub/76076

Akemu, O., *Corporate Responses to Social Issues: Essays in Social Entrepreneurship and Corporate Social Responsibility*, Promotors: Prof. G.M. Whiteman & Dr S.P. Kennedy,
EPS-2017-392-ORG, https://repub.eur.nl/pub/95768

Akin Ates, M., *Purchasing and Supply Management at the Purchase Category Level: Strategy, structure and performance*, Promotors: Prof. J.Y.F. Wynstra & Dr E.M.
van Raaij, EPS-2014-300-LIS, http://repub.eur.nl/pub/50283

Alexander, L., *People, Politics, and Innovation: A Process Perspective*,
Promotors: Prof. H.G. Barkema & Prof. D.L. van Knippenberg, EPS-2014-331-S&E,
http://repub.eur.nl/pub/77209

Alexiou, A. *Management of Emerging Technologies and the Learning Organization: Lessons from the Cloud and Serious Games Technology*, Promotors: Prof. S.J. Magala,
Prof. M.C. Schippers and Dr I. Oshri, EPS-2016-404-ORG, http://repub.eur.nl/pub/93818

Almeida e Santos Nogueira, R.J. de, *Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty*, Promotors: Prof. U. Kaymak &
Prof. J.M.C. Sousa, EPS-2014-310-LIS, http://repub.eur.nl/pub/51560

Alserda, G.A.G., *Choices in Pension Management*, Promotors: Prof. S.G. van der Lecq &
Dr O.W. Steenbeek, EPS-2017-432-F&A, https://repub.eur.nl/pub/103496

Benschop, N, *Biases in Project Escalation: Names, frames & construal levels*,

Promotors: Prof. K.I.M. Rhode, Prof. H.R. Commandeur, Prof. M. Keil & Dr A.L.P. Nuijten, EPS-2015-375-S&E, http://repub.eur.nl/pub/79408

Berg, W.E. van den, *Understanding Salesforce Behavior using Genetic Association Studies*, Promotor: Prof. W.J.M.I. Verbeke, EPS-2014-311-MKT, http://repub.eur.nl/pub/51440

Beusichem, H.C. van, *Firms and Financial Markets: Empirical Studies on the Informational Value of Dividends, Governance and Financial Reporting*, Promotors: Prof. A. de Jong & Dr G. Westerhuis, EPS-2016-378-F&A, http://repub.eur.nl/pub/93079

Bliek, R. de, *Empirical Studies on the Economic Impact of Trust*, Promotor: Prof. J. Veenman & Prof. Ph.H.B.F. Franses, EPS-2015-324-ORG, http://repub.eur.nl/pub/78159

Boons, M., *Working Together Alone in the Online Crowd: The Effects of Social Motivations and Individual Knowledge Backgrounds on the Participation and Performance of Members of Online Crowdsourcing Platforms*, Promotors: Prof. H.G. Barkema & Dr D.A. Stam, EPS-2014-306-S&E, http://repub.eur.nl/pub/50711

Bouman, P., *Passengers, Crowding and Complexity: Models for Passenger Oriented Public Transport*, Prof. L.G. Kroon, Prof. A. Schöbel & Prof. P.H.M. Vervest, EPS-2017-420-LIS, https://repub.eur.nl/

Brazys, J., *Aggregated Marcoeconomic News and Price Discovery*, Promotor: Prof. W.F.C. Verschoor, EPS-2015-351-F&A, http://repub.eur.nl/pub/78243

Bunderen, L. van, *Tug-of-War: Why and when teams get embroiled in power struggles*, Promotors: Prof. D.L. van Knippenberg & Dr. L. Greer, EPS-2018-446-ORG, https://repub.eur.nl/pub/105346

Burg, G.J.J. van den, *Algorithms for Multiclass Classification and Regularized Regression*, Promotors: Prof. P.J.F. Groenen & Dr. A. Alfons, EPS-2018-442-MKT, https://repub.eur.nl/pub/103929

Cancurtaran, P., *Essays on Accelerated Product Development*, Promotors: Prof. F. Langerak & Prof. G.H. van Bruggen, EPS-2014-317-MKT, http://repub.eur.nl/pub/76074

Chammas, G., *Portfolio concentration*, Promotor: Prof. J. Spronk, EPS-2017-410-F&E, https://repub.eur.nl/pub/94975

Cranenburgh, K.C. van, *Money or Ethics: Multinational corporations and religious organisations operating in an era of corporate responsibility*, Prof. L.C.P.M. Meijs, Prof. R.J.M. van Tulder & Dr D. Arenas, EPS-2016-385-ORG, http://repub.eur.nl/pub/93104

Consiglio, I., *Others: Essays on Interpersonal and Consumer Behavior*, Promotor: Prof. S.M.J. van Osselaer, EPS-2016-366-MKT, http://repub.eur.nl/pub/79820

Darnihamedani, P. *Individual Characteristics, Contextual Factors and Entrepreneurial Behavior*, Promotors: Prof. A.R. Thurik & S.J.A. Hessels, EPS-2016-360-S&E, http://repub.eur.nl/pub/93280

Dennerlein, T. *Empowering Leadership and Employees' Achievement Motivations: the Role of Self-Efficacy and Goal Orientations in the Empowering Leadership Process*, Promotors: Prof. D.L. van Knippenberg & Dr J. Dietz, EPS-2017-414-ORG, https://repub.eur.nl/pub/98438

Deng, W., *Social Capital and Diversification of Cooperatives*, Promotor: Prof. G.W.J. Hendrikse, EPS-2015-341-ORG, http://repub.eur.nl/pub/77449

Depecik, B.E., *Revitalizing brands and brand: Essays on Brand and Brand Portfolio Management Strategies*, Promotors: Prof. G.H. van Bruggen, Dr Y.M. van Everdingen and Dr M.B. Ataman, EPS-2016-406-MKT, http://repub.eur.nl/pub/93507

Duijzer, L.E., *Mathematical Optimization in Vaccine Allocation*, Promotors: Prof. R. Dekker & Dr W.L. van Jaarsveld, EPS-2017-430-LIS, https://repub.eur.nl/pub/101487

Duyvesteyn, J.G. *Empirical Studies on Sovereign Fixed Income Markets*, Promotors: Prof. P. Verwijmeren & Prof. M.P.E. Martens, EPS-2015-361-F&A, https://repub.eur.nl/pub/79033

Elemes, A., *Studies on Determinants and Consequences of Financial Reporting Quality*, Promotor: Prof. E. Peek, EPS-2015-354-F&A, https://repub.eur.nl/pub/79037

Ellen, S. ter, *Measurement, Dynamics, and Implications of Heterogeneous Beliefs in Financial Markets*, Promotor: Prof. W.F.C. Verschoor, EPS-2015-343-F&A, http://repub.eur.nl/pub/78191

Erlemann, C., *Gender and Leadership Aspiration: The Impact of the Organizational Environment*, Promotor: Prof. D.L. van Knippenberg, EPS-2016-376-ORG, http://repub.eur.nl/pub/79409

Eskenazi, P.I., *The Accountable Animal*, Promotor: Prof. F.G.H. Hartmann, EPS-2015-355-F&A, http://repub.eur.nl/pub/78300

Evangelidis, I., *Preference Construction under Prominence*, Promotor: Prof. S.M.J. van Osselaer, EPS-2015-340-MKT, http://repub.eur.nl/pub/78202

Faber, N., *Structuring Warehouse Management*, Promotors: Prof. M.B.M. de Koster & Prof. A. Smidts, EPS-2015-336-LIS, http://repub.eur.nl/pub/78603

Feng, Y., *The Effectiveness of Corporate Governance Mechanisms and Leadership Structure: Impacts on strategic change and firm performance*, Promotors: Prof. F.A.J. van den Bosch, Prof. H.W. Volberda & Dr J.S. Sidhu, EPS-2017-389-S&E, https://repub.eur.nl/pub/98470

Fernald, K., *The Waves of Biotechnological Innovation in Medicine: Interfirm Cooperation Effects and a Venture Capital Perspective*, Promotors: Prof. E. Claassen, Prof. H.P.G. Pennings & Prof. H.R. Commandeur, EPS-2015-371-S&E, http://hdl.handle.net/1765/79120

Fisch, C.O., *Patents and trademarks: Motivations, antecedents, and value in industrialized and emerging markets*, Promotors: Prof. J.H. Block, Prof. H.P.G. Pennings & Prof. A.R. Thurik, EPS-2016-397-S&E, http://repub.eur.nl/pub/94036

Fliers, P.T., *Essays on Financing and Performance: The role of firms, banks and board*, Promotors: Prof. A. de Jong & Prof. P.G.J. Roosenboom, EPS-2016-388-F&A, http://repub.eur.nl/pub/93019

Fourne, S.P., *Managing Organizational Tensions: A Multi-Level Perspective on Exploration, Exploitation and Ambidexterity*, Promotors: Prof. J.J.P. Jansen & Prof. S.J. Magala, EPS-2014-318-S&E, http://repub.eur.nl/pub/76075

Gaast, J.P. van der, *Stochastic Models for Order Picking Systems*, Promotors: Prof. M.B.M de Koster & Prof. I.J.B.F. Adan, EPS-2016-398-LIS, http://repub.eur.nl/pub/93222

Giurge, L., *A Test of Time; A temporal and dynamic approach to power and ethics*, Promotors: Prof. M.H. van Dijke & Prof. D. De Cremer, EPS-2017-412-ORG, https://repub.eur.nl/

Glorie, K.M., *Clearing Barter Exchange Markets: Kidney Exchange and Beyond*, Promotors: Prof. A.P.M. Wagelmans & Prof. J.J. van de Klundert, EPS-2014-329-LIS, http://repub.eur.nl/pub/77183

Gobena, L., *Towards Integrating Antecedents of Voluntary Tax Compliance*, Promotors: Prof. M.H. van Dijke & Dr P. Verboon, EPS-2017-436-ORG, https://repub.eur.nl/pub/103276

Groot, W.A., *Assessing Asset Pricing Anomalies*, Promotors: Prof. M.J.C.M. Verbeek & Prof. J.H. van Binsbergen, EPS-2017-437-F&A, https://repub.eur.nl/pub/103490

Hekimoglu, M., *Spare Parts Management of Aging Capital Products*, Promotor: Prof. R. Dekker, EPS-2015-368-LIS, http://repub.eur.nl/pub/79092

Hengelaar, G.A., *The Proactive Incumbent: Holy grail or hidden gem? Investigating whether the Dutch electricity sector can overcome the incumbent's curse and lead the sustainability transition*, Promotors: Prof. R.J. M. van Tulder & Dr K. Dittrich, EPS-2018-438-ORG, https://repub.eur.nl/pub/102953

Hogenboom, A.C., *Sentiment Analysis of Text Guided by Semantics and Structure*, Promotors: Prof. U. Kaymak & Prof. F.M.G. de Jong, EPS-2015-369-LIS, http://repub.eur.nl/pub/79034

Hogenboom, F.P., *Automated Detection of Financial Events in News Text*, Promotors: Prof. U. Kaymak & Prof. F.M.G. de Jong, EPS-2014-326-LIS, http://repub.eur.nl/pub/77237

Hollen, R.M.A., *Exploratory Studies into Strategies to Enhance Innovation-Driven International Competitiveness in a Port Context: Toward Ambidextrous Ports*, Promotors: Prof. F.A.J. Van Den Bosch & Prof. H.W. Volberda, EPS-2015-372-S&E, http://repub.eur.nl/pub/78881

Hout, D.H. van, *Measuring Meaningful Differences: Sensory Testing Based Decision Making in an Industrial Context; Applications of Signal Detection Theory and Thurstonian Modelling*, Promotors: Prof. P.J.F. Groenen & Prof. G.B. Dijksterhuis, EPS-2014-304-MKT, http://repub.eur.nl/pub/50387

Houwelingen, G.G. van, *Something To Rely On*, Promotors: Prof. D. de Cremer & Prof. M.H. van Dijke, EPS-2014-335-ORG, http://repub.eur.nl/pub/77320

Hurk, E. van der, *Passengers, Information, and Disruptions*, Promotors: Prof. L.G. Kroon & Prof. P.H.M. Vervest, EPS-2015-345-LIS, http://repub.eur.nl/pub/78275

Iseger, P. den, *Fourier and Laplace Transform Inversion with Applications in Finance*, Promotor: Prof. R. Dekker, EPS-2014-322-LIS, http://repub.eur.nl/pub/76954

Jacobs, B.J.D., *Marketing Analytics for High-Dimensional Assortments*, Promotors: Prof. A.C.D. Donkers & Prof. D. Fok, EPS-2017-445-MKT, https://repub.eur.nl/pub/103497

Kahlen, M. T., *Virtual Power Plants of Electric Vehicles in Sustainable Smart Electricity Markets*, Promotors: Prof. W. Ketter & Prof. A. Gupta, EPS-2017-431-LIS, https://repub.eur.nl/pub/100844

Kampen, S. van, *The Cross-sectional and Time-series Dynamics of Corporate Finance: Empirical evidence from financially constrained firms*, Promotors: Prof. L. Norden & Prof. P.G.J. Roosenboom, EPS-2018-440-F&A, https://repub.eur.nl/pub/105245

Keko. E, *Essays on Innovation Generation in Incumbent Firms*, Promotors: Prof. S. Stremersch & Dr N.M.A. Camacho, EPS-2017-419-MKT, https://repub.eur.nl/pub/100841

Khanagha, S., *Dynamic Capabilities for Managing Emerging Technologies*, Promotor: Prof. H.W. Volberda, EPS-2014-339-S&E, http://repub.eur.nl/pub/77319

Khattab, J., *Make Minorities Great Again: a contribution to workplace equity by identifying and addressing constraints and privileges*, Prof. D.L. van Knippenberg & Dr A. Nederveen Pieterse, EPS-2017-421-ORG, https://repub.eur.nl/pub/99311

Klooster, E. van 't, *Travel to Learn: the Influence of Cultural Distance on Competence Development in Educational Travel*, Promotors: Prof. F.M. Go & Prof. P.J. van Baalen, EPS-2014-312-MKT, http://repub.eur.nl/pub/51462

Koendjbiharie, S.R., *The Information-Based View on Business Network Performance: Revealing the Performance of Interorganizational Networks*, Promotors: Prof. H.W.G.M. van Heck & Prof. P.H.M. Vervest, EPS-2014-315-LIS, http://repub.eur.nl/pub/51751

Koning, M., *The Financial Reporting Environment: The Role of the Media, Regulators and Auditors*, Promotors: Prof. G.M.H. Mertens & Prof. P.G.J. Roosenboom, EPS-2014-330-F&A, http://repub.eur.nl/pub/77154

Konter, D.J., *Crossing Borders with HRM: An Inquiry of the Influence of Contextual Differences in the Adoption and Effectiveness of HRM*, Promotors: Prof. J. Paauwe, & Dr L.H. Hoeksema, EPS-2014-305-ORG, http://repub.eur.nl/pub/50388

Korkmaz, E., *Bridging Models and Business: Understanding Heterogeneity in Hidden Drivers of Customer Purchase Behavior*, Promotors: Prof. S.L. van de Velde & Prof. D. Fok, EPS-2014-316-LIS, http://repub.eur.nl/pub/76008

Krämer, R., *A license to mine? Community organizing against multinational corporations*, Promotors: Prof. R.J.M. van Tulder & Prof. G.M. Whiteman, EPS-2016-383-ORG, http://repub.eur.nl/pub/94072

Kroezen, J.J., *The Renewal of Mature Industries: An Examination of the Revival of the Dutch Beer Brewing Industry*, Promotor: Prof. P.P.M.A.R. Heugens, EPS-2014-333-S&E, http://repub.eur.nl/pub/77042

Kysucky, V., *Access to Finance in a Cros-Country Context*, Promotor: Prof. L. Norden, EPS-2015-350-F&A, http://repub.eur.nl/pub/78225

Lee, C.I.S.G., *Big Data in Management Research: Exploring New Avenues*, Promotors: Prof. S.J. Magala & Dr W.A. Felps, EPS-2016-365-ORG, http://repub.eur.nl/pub/79818

Legault-Tremblay, P.O., *Corporate Governance During Market Transition: Heterogeneous responses to Institution Tensions in China*, Promotor: Prof. B. Krug, EPS-2015-362-ORG, http://repub.eur.nl/pub/78649

Lenoir, A.S. *Are You Talking to Me? Addressing Consumers in a Globalised World*, Promotors: Prof. S. Puntoni & Prof. S.M.J. van Osselaer, EPS-2015-363-MKT, http://repub.eur.nl/pub/79036

Leunissen, J.M., *All Apologies: On the Willingness of Perpetrators to Apologize*, Promotors: Prof. D. de Cremer & Dr M. van Dijke, EPS-2014-301-ORG, http://repub.eur.nl/pub/50318

Li, D., *Supply Chain Contracting for After-sales Service and Product Support*, Promotor: Prof. M.B.M. de Koster, EPS-2015-347-LIS, http://repub.eur.nl/pub/78526

Li, Z., *Irrationality: What, Why and How*, Promotors: Prof. H. Bleichrodt, Prof. P.P. Wakker, & Prof. K.I.M. Rohde, EPS-2014-338-MKT, http://repub.eur.nl/pub/77205

Liu, N., *Behavioral Biases in Interpersonal Contexts*, Supervisors: Prof. A. Baillon & Prof. H. Bleichrodt, EPS-2017-408-MKT, https://repub.eur.nl/pub/95487

Liket, K., *Why 'Doing Good' is not Good Enough: Essays on Social Impact Measurement*, Promotors: Prof. H.R. Commandeur & Dr K.E.H. Maas, EPS-2014-307-STR, http://repub.eur.nl/pub/51130

Lu, Y., *Data-Driven Decision Making in Auction Markets*, Promotors: Prof. H.W.G.M. van Heck & Prof. W. Ketter, EPS-2014-314-LIS, http://repub.eur.nl/pub/51543

Ma, Y., *The Use of Advanced Transportation Monitoring Data for Official Statistics*, Promotors: Prof. L.G. Kroon & Dr J. van Dalen, EPS-2016-391-LIS, http://repub.eur.nl/pub/80174

Maira, E., *Consumers and Producers*, Promotors: Prof. S. Puntoni & Prof. C. Fuchs, EPS-2018-439-MKT, https://repub.eur.nl/pub/104387

Manders, B., *Implementation and Impact of ISO 9001*, Promotor: Prof. K. Blind, EPS-2014-337-LIS, http://repub.eur.nl/pub/77412

Mell, J.N., *Connecting Minds: On The Role of Metaknowledge in Knowledge Coordination*, Promotor: Prof. D.L. van Knippenberg, EPS-2015-359-ORG, http://hdl.handle.net/1765/78951

Meulen,van der, D., *The Distance Dilemma: the effect of flexible working practices on performance in the digital workplace*, Promotors: Prof. H.W.G.M. van Heck & Prof. P.J. van Baalen, EPS-2016-403-LIS, http://repub.eur.nl/pub/94033

Micheli, M.R., *Business Model Innovation: A Journey across Managers' Attention and Inter-Organizational Networks*, Promotor: Prof. J.J.P. Jansen, EPS-2015-344-S&E, http://repub.eur.nl/pub/78241

Moniz, A, *Textual Analysis of Intangible Information*, Promotors: Prof. C.B.M. van Riel, Prof. F.M.G de Jong & Dr G.A.J.M. Berens, EPS-2016-393-ORG, http://repub.eur.nl/pub/93001

Mulder, J. *Network design and robust scheduling in liner shipping*, Promotors: Prof. R. Dekker & Dr W.L. van Jaarsveld, EPS-2016-384-LIS, http://repub.eur.nl/pub/80258

Naumovska, I., *Socially Situated Financial Markets: A Neo-Behavioral Perspective on Firms, Investors and Practices*, Promotors: Prof. P.P.M.A.R. Heugens & Prof. A. de Jong, EPS-2014-319-S&E, http://repub.eur.nl/pub/76084

Neerijnen, P., *The Adaptive Organization: the socio-cognitive antecedents of ambidexterity and individual exploration*, Promotors: Prof. J.J.P. Jansen, P.P.M.A.R. Heugens & Dr T.J.M. Mom, EPS-2016-358-S&E, http://repub.eur.nl/pub/93274

Okbay, A., *Essays on Genetics and the Social Sciences*, Promotors: Prof. A.R. Thurik, Prof. Ph.D. Koellinger & Prof. P.J.F. Groenen, EPS-2017-413-S&E, https://repub.eur.nl/pub/95489

Oord, J.A. van, *Essays on Momentum Strategies in Finance*, Promotor: Prof. H.K. van Dijk, EPS-2016-380-F&A, http://repub.eur.nl/pub/80036

Peng, X., *Innovation, Member Sorting, and Evaluation of Agricultural Cooperatives*, Promotor: Prof. G.W.J. Hendriks, EPS-2017-409-ORG, https://repub.eur.nl/pub/94976

Pennings, C.L.P., *Advancements in Demand Forecasting: Methods and Behavior*, Promotors: Prof. L.G. Kroon, Prof. H.W.G.M. van Heck & Dr J. van Dalen, EPS-2016-400-LIS, http://repub.eur.nl/pub/94039

Peters, M., *Machine Learning Algorithms for Smart Electricity Markets*, Promotor: Prof. W. Ketter, EPS-2014-332-LIS, http://repub.eur.nl/pub/77413

Petruchenya, A., *Essays on Cooperatives: Emergence, Retained Earnings, and Market Shares*, Promotors: Prof. G.W.J. Hendriks & Dr Y. Zhang, EPS-2018-447-ORG, https://repub.eur.nl/pub/105243

Plessis, C. du, *Influencers: The Role of Social Influence in Marketing*, Promotors: Prof. S. Puntoni & Prof. S.T.L.R. Sweldens, EPS-2017-425-MKT, https://repub.eur.nl/pub/103265

Pocock, M., *Status Inequalities in Business Exchange Relations in Luxury Markets*, Promotors: Prof. C.B.M. van Riel & Dr G.A.J.M. Berens, EPS-2017-346-ORG, https://repub.eur.nl/pub/98647

Pozharliev, R., *Social Neuromarketing: The role of social context in measuring advertising effectiveness*, Promotors: Prof. W.J.M.I. Verbeke & Prof. J.W. van Strien, EPS-2017-402-MKT, https://repub.eur.nl/pub/95528

Protzner, S. *Mind the gap between demand and supply: A behavioral perspective on demand forecasting*, Promotors: Prof. S.L. van de Velde & Dr L. Rook, EPS-2015-364-LIS, *http*://repub.eur.nl/pub/79355

Pruijssers, J.K., *An Organizational Perspective on Auditor Conduct*, Promotors: Prof. J. van Oosterhout & Prof. P.P.M.A.R. Heugens, EPS-2015-342-S&E, http://repub.eur.nl/pub/78192

Riessen, B. van, *Optimal Transportation Plans and Portfolios for Synchromodal Container Networks*, Promotors: Prof. R. Dekker & Prof. R.R. Negenborn, EPS-2018-448-LIS, https://repub.eur.nl/pub/105248

Rietdijk, W.J.R. *The Use of Cognitive Factors for Explaining Entrepreneurship*, Promotors: Prof. A.R. Thurik & Prof. I.H.A. Franken, EPS-2015-356-S&E, http://repub.eur.nl/pub/79817

Rietveld, N., *Essays on the Intersection of Economics and Biology*, Promotors: Prof. A.R. Thurik, Prof. Ph.D. Koellinger, Prof. P.J.F. Groenen, & Prof. A. Hofman, EPS-2014-320-S&E, http://repub.eur.nl/pub/76907

Rösch, D. *Market Efficiency and Liquidity*, Promotor: Prof. M.A. van Dijk, EPS-2015-353-F&A, http://repub.eur.nl/pub/79121

Roza, L., *Employee Engagement in Corporate Social Responsibility: A collection of essays*, Promotor: Prof. L.C.P.M. Meijs, EPS-2016-396-ORG, http://repub.eur.nl/pub/93254

Schie, R. J. G. van, *Planning for Retirement: Save More or Retire Later?* Promotors: Prof. B. G. C. Dellaert & Prof. A.C.D. Donkers, EOS-2017-415-MKT, https://repub.eur.nl/pub/100846

Schoonees, P. *Methods for Modelling Response Styles*, Promotor: Prof. P.J.F. Groenen, EPS-2015-348-MKT, http://repub.eur.nl/pub/79327

Schouten, M.E., *The Ups and Downs of Hierarchy: the causes and consequences of hierarchy struggles and positional loss*, Promotors; Prof. D.L. van Knippenberg & Dr L.L. Greer, EPS-2016-386-ORG, http://repub.eur.nl/pub/80059

Smit, J. *Unlocking Business Model Innovation: A look through the keyhole at the inner workings of Business Model Innovation*, Promotor: Prof. H.G. Barkema, EPS-2016-399-S&E, http://repub.eur.nl/pub/93211

Sousa, M.J.C. de, *Servant Leadership to the Test: New Perspectives and Insights*, Promotors: Prof. D.L. van Knippenberg & Dr D. van Dierendonck, EPS-2014-313-ORG, http://repub.eur.nl/pub/51537

Staadt, J.L., *Leading Public Housing Organisation in a Problematic Situation: A Critical Soft Systems Methodology Approach*, Promotor: Prof. S.J. Magala, EPS-2014-308-ORG, http://repub.eur.nl/pub/50712

Straeter, L.M., *Interpersonal Consumer Decision Making*, Promotors: Prof. S.M.J. van Osselaer & Dr I.E. de Hooge, EPS-2017-423-MKT, https://repub.eur.nl/pub/100819

Subaşi, B., *Demographic Dissimilarity, Information Access and Individual Performance*, Promotors: Prof. D.L. van Knippenberg & Dr W.P. van Ginkel, EPS-2017-422-ORG, https://repub.eur.nl/pub/103495

Szatmari, B., *We are (all) the champions: The effect of status in the implementation of innovations*, Promotors: Prof. J.C.M & Dr D. Deichmann, EPS-2016-401-LIS, http://repub.eur.nl/pub/94633

Tuijl, E. van, *Upgrading across Organisational and Geographical Configurations*, Promotor: Prof. L. van den Berg, EPS-2015-349-S&E, http://repub.eur.nl/pub/78224

Tuncdogan, A., *Decision Making and Behavioral Strategy: The Role of Regulatory Focus in Corporate Innovation Processes*, Promotors: Prof. F.A.J. van den Bosch, Prof. H.W. Volberda, & Prof. T.J.M. Mom, EPS-2014-334-S&E, http://repub.eur.nl/pub/76978

Uijl, S. den, *The Emergence of De-facto Standards*, Promotor: Prof. K. Blind, EPS-2014-328-LIS, http://repub.eur.nl/pub/77382

Valogianni, K. *Sustainable Electric Vehicle Management using Coordinated Machine Learning*, Promotors: Prof. H.W.G.M. van Heck & Prof. W. Ketter, EPS-2016-387-LIS, http://repub.eur.nl/pub/93018

Vandic, D., *Intelligent Information Systems for Web Product Search*, Promotors: Prof. U. Kaymak & Dr Frasincar, EPS-2017-405-LIS, https://repub.eur.nl/pub/95490

Veelenturf, L.P., *Disruption Management in Passenger Railways: Models for Timetable, Rolling Stock and Crew Rescheduling*, Promotor: Prof. L.G. Kroon, EPS-2014-327-LIS, http://repub.eur.nl/pub/77155

Verbeek, R.W.M., *Essays on Empirical Asset Pricing*, Promotors: Prof. M.A. van Dijk & Dr M. Szymanowska, EPS-2017-441-F&A, https://repub.eur.nl/pub/102977

Vermeer, W., *Propagation in Networks:The impact of information processing at the actor level on system-wide propagation dynamics*, Promotor: Prof. P.H.M.Vervest, EPS-2015-373-LIS, http://repub.eur.nl/pub/79325

Versluis, I., *Prevention of the Portion Size Effect*, Promotors: Prof. Ph.H.B.F. Franses & Dr E.K. Papies, EPS-2016-382-MKT, http://repub.eur.nl/pub/79880

Vishwanathan, P., *Governing for Stakeholders: How Organizations May Create or Destroy Value for their Stakeholders*, Promotors: Prof. J. van Oosterhout & Prof. L.C.P.M. Meijs, EPS-2016-377-ORG, http://repub.eur.nl/pub/93016

Vlaming, R. de.,*Linear Mixed Models in Statistical Genetics*, Prof. A.R. Thurik, Prof. P.J.F. Groenen & Prof. Ph.D. Koellinger, EPS-2017-416-S&E, https://repub.eur.nl/pub/100428

Vries, H. de, *Evidence-Based Optimization in Humanitarian Logistics*, Promotors: Prof. A.P.M. Wagelmans & Prof. J.J. van de Klundert, EPS-2017-435-LIS, https://repub.eur.nl/pub/102771

Vries, J. de, *Behavioral Operations in Logistics*, Promotors: Prof. M.B.M de Koster & Prof. D.A. Stam, EPS-2015-374-LIS, http://repub.eur.nl/pub/79705

Wagenaar, J.C., *Practice Oriented Algorithmic Disruption Management in Passenger Railways*, Prof. L.G. Kroon & Prof. A.P.M. Wagelmans, EPS-2016-390-LIS, http://repub.eur.nl/pub/93177

Wang, P., *Innovations, status, and networks*, Promotors: Prof. J.J.P. Jansen & Dr V.J.A. van de Vrande, EPS-2016-381-S&E, http://repub.eur.nl/pub/93176

Wang, R., *Corporate Environmentalism in China*, Promotors: Prof. P.P.M.A.R Heugens & Dr F. Wijen, EPS-2017-417-S&E, https://repub.eur.nl/pub/99987

Wang, T., *Essays in Banking and Corporate Finance*, Promotors: Prof. L. Norden & Prof. P.G.J. Roosenboom, EPS-2015-352-F&A, http://repub.eur.nl/pub/78301

Wasesa, M., *Agent-based inter-organizational systems in advanced logistics operations*, Promotors: Prof. H.W.G.M van Heck, Prof. R.A. Zuidwijk & Dr A. W. Stam, EPS-2017-LIS-424, https://repub.eur.nl/pub/100527

Weenen, T.C., *On the Origin and Development of the Medical Nutrition Industry*, Promotors: Prof. H.R. Commandeur & Prof. H.J.H.M. Claassen, EPS-2014-309-S&E, http://repub.eur.nl/pub/51134

Wessels, C., *Flexible Working Practices: How Employees Can Reap the Benefits for Engagement and Performance*, Promotors: Prof. H.W.G.M. van Heck, Prof. P.J. van Baalen & Prof. M.C. Schippers, EPS-2017-418-LIS, https://repub.eur.nl/

Witte, C.T., *Bloody Business: Multinational investment in an increasingly conflict-afflicted world*, Promotors: Prof. H.P.G. Pennings, Prof. H.R. Commandeur & Dr M.J. Burger, EPS-2018-443-S&E, https://repub.eur.nl/pub/104027

Yang, S., *Information Aggregation Efficiency of Prediction Markets*, Promotor: Prof. H.W.G.M. van Heck, EPS-2014-323-LIS, http://repub.eur.nl/pub/77184

Yuan, Y., *The Emergence of Team Creativity: a social network perspective*, Promotors: Prof. D. L. van Knippenberg & Dr D. A. Stam, EPS-2017-434-ORG, https://repub.eur.nl/pub/100847

Ypsilantis, P., *The Design, Planning and Execution of Sustainable Intermodal Port-hinterland Transport Networks*, Promotors: Prof. R.A. Zuidwijk & Prof. L.G. Kroon, EPS-2016-395-LIS, http://repub.eur.nl/pub/94375

Yuferova, D. *Price Discovery, Liquidity Provision, and Low-Latency Trading*, Promotors: Prof. M.A. van Dijk & Dr D.G.J. Bongaerts, EPS-2016-379-F&A, http://repub.eur.nl/pub/93017

Zhang, Q., *Financing and Regulatory Frictions in Mergers and Acquisitions*, Promotors: Prof. P.G.J. Roosenboom & Prof. A. de Jong, EPS-2018-428-F&A, https://repub.eur.nl/pub/103871

Zuber, F.B., *Looking at the Others: Studies on (un)ethical behavior and social relationships in organizations*, Promotor: Prof. S.P. Kaptein, EPS-2016-394-ORG, http://repub.eur.nl/pub/94388

Job A. Harms was born in 1985 in Amsterdam, the Netherlands. He obtained a BSc in Liberal Arts and Sciences at University College Utrecht in 2006. He then pursued a MSc in Development Economics at the VU University Amsterdam. During his studies he conducted field research with the International Labor Organization, obtaining his degree in 2010. From 2011-2012 he worked for Triodos Facet. In 2013 he started his doctoral research at the Erasmus School of Economics. He is currently a postdoctoral researcher at the Faculty for Behavioral and Social Sciences at the University of Groningen.

**ERiM**

The Erasmus Research Institute of Management (ERIM) is the Research School (Onderzoekschool) in the field of management of the Erasmus University Rotterdam. The founding participants of ERIM are the Rotterdam School of Management (RSM), and the Erasmus School of Economics (ESE). ERIM was founded in 1999 and is officially accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW). The research undertaken by ERIM is focused on the management of the firm in its environment, its intra- and interfirm relations, and its business processes in their interdependent connections.

The objective of ERIM is to carry out first rate research in management, and to offer an advanced doctoral programme in Research in Management. Within ERIM, over three hundred senior researchers and PhD candidates are active in the different research programmes. From a variety of academic backgrounds and expertises, the ERIM community is united in striving for excellence and working at the forefront of creating new business knowledge.

**ERiM**

ERIM PhD Series
**Research in Management**

**Erasmus University Rotterdam (EUR)**
**Erasmus Research Institute of Management**
Mandeville (T) Building
Burgemeester Oudlaan 50
3062 PA Rotterdam, The Netherlands

P.O. Box 1738
3000 DR Rotterdam, The Netherlands
T +31 10 408 1182
E info@erim.eur.nl
W www.erim.eur.nl