Fairness Views in Social and Individual Decisions

by

Sinong Ma

Thesis

Submitted to the University of Warwick

in partial fulfillment of the requirements for the degree

of

Doctor of Philosophy

Warwick Business School

April 2017

THE UNIVERSITY OF WARWICK

Contents

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisors, Zvi Safra and Tigran Melkonyan, for introducing me to this interesting field, suggesting cutting-edge research problems and providing tremendous support for my study and research. Their enthusiasm and professional attitudes to research have greatly influenced and shaped my development, and their deep insights, patient guidance and consistent encouragement over the years have been invaluable. It has always been an honour to be their student.

I would like to express my thanks to members of the Behavioural Science group at Warwick Business School. This group has provided a friendly and ideal environment for collaborating and exchanging ideas. Special thanks go to Ed Gardiner, for connecting me with health institutions and making Chapter 5 of this thesis possible. I am also grateful to faculty members who have provided helpful comments on parts of this thesis and related presentations, including Jerker Denrell, Edi Karni, Nick Chater, Graham Loomes, Uzi Segal, Daniel Read, Andrea Isoni, Hossam Zeitoun and Dawn Eubanks.

Thanks also go to the intelligent postdocs and graduate students with whom I have had the pleasure to discuss my thoughts, including Marsha Kirichek, Jiwei Zheng, Ali Osseiran, Lisheng He, Zhihua Li, Rebecca McDonald, Evan Fradkin, Tim Vandendriessche, Caspar Von Schenck, Avri Bilovich, Hala Elbilbaisi, Asma Bagash, Juan Carpio, Karen Coletta, Umar Taj, Krishane Patel, Jess Whittlestone, Julia Kolodko, Hana Sysalova, Danae Arroyos Calvera, David Butler and Zhifan Zhou. I was particularly impressed by Marsha's expertise and ability while we worked together on the healthcare project with the NHS England. Jiwei used to share a flat with me in Tile Hill, and he has always been a reliable source of advice not only on empirical works but also on taekwondo skills.

In addition, I would like to express my appreciation to my friends from other groups, departments and institutions, including, among others too numerous to mention, Andrew Burlinson, Xiaoyu Li, Zizhong Yan, Gongtai Wang, Ellen He, Chao Wang, Dee Agyei, Zilong Zhang, Lauren Schrock, Xuwen Zhu, Fei Liu, Wentao Fu, Lin Zheng, Giuseppe Bova, Sonya Liu, Huikang Ying, Lucius Li, Lawrence Choo, Jiang Xue, Feng Song and Trixie Gadd. Thanks for being great friends and mentors.

Last but not least, I would like to thank my family for their understanding and support. Thank you to my parents who raise me up with great effort, provided me with an loving environment, and supported me in all my pursuits. And, most of all, thank you to my loving wife Jia, whose support and encouragement are so appreciated. Thank you.

Declarations

I declare that any material contained in this thesis has not been submitted for a degree to any other University.

Chapter 1

Introduction

Fairness and efficiency are two classical and connected topics in economics. They have become well known, perhaps due to Adam Smith's two influential works: *The Theory of Moral Sentiments* (1759), which highlights a concern for fairness concern as part of morality, and *The Wealth of Nations* (1776), which underlines a concern for efficiency. However, during the rapid development of economics, fairness has received disproportionately less attention than efficiency. As a result, many people, including some economists, have incorrectly understood that economics as a subject no longer cares about fairness.

The primary objective of this thesis is to dispel this misperception. We would argue here that, similar to efficiency, fairness is an important factor for both social and individual decisions, and sometimes its effect can be determined.

Written in a three-paper format, this thesis explores fairness from three different angles. These angles cover the broad areas of how theoretical economists model fairness in social choice theory[1], how the general public perceive distributive fairness, and how people implement their fairness norms in making real-life donations. This multidimensional exploration is believed to be crucial to a comprehensive understanding of fairness.

[1]Social choice theory deals with the aggregation of some measure of individual welfare into a collective measure (Sen, 2008).

Chapter 3 presents the first paper, which makes a theoretical contribution to social choice theory. We reconsider John Harsanyi's (1953, 1955, 1977b) utilitarian impartial observer theorem. In this theorem, Harsanyi argues that an impartial social decision can only be made by "the impartial observer" (who is behind a veil of ignorance and unknown about her identity in society) and this impartial observer follows a utilitarian principle (which credits the best policy as one that maximises the average expected utility). Departing from Harsanyi's individual-centred approach, we argue that, when societal decisions are at stake, postulates must not be drawn from individualistic behaviour. Rather, they should be based on societal norms. Hence, notions like societal fairness (Diamond, 1967; Sen 1977) should explicitly be the guiding principles. Continuing this line of thinking, we state and prove a utilitarian result that, rather than being based on the independence assumption, is based on the notion of procedural fairness and on symmetric treatment of societal and individual lotteries.

Chapter 4 presents the second paper, which contributes to an understanding of public perceptions of distributive fairness. The last few decades have witnessed an increasing proportion of all cancers are lifestyle-induced, a trend tends to continue. To assess policies used to address this challenge and to construct a more equitable healthcare system, it is important to gain a better understanding of how individual responsibility is viewed by individual society members. We have constructed a procedure to assess these preferences with respect to lifestyle-induced and hereditary cancers by eliciting donations to these two types of cancers. Lifestyle-induced cancers involve greater individual responsibility than hereditary cancers, as individuals can control more elements of lifestyle than heredity. The results of implementing our procedure via an online survey demonstrate that subjects take individual responsibility into account by donating about twice as much to cancers with less individual responsibility. Their choices are also affected by group identity, perceptions of cancer likelihood and social demographics.

Chapter 5 presents the third paper, which makes an empirical contributions to understanding how people imply fairness principles in everyday donations. The paper uses information on charitable contributions to cancer research in the United Kingdom to elicit information on fairness principles endorsed by donors. The latter face a choice between contributing to several hereditary and lifestyle-related cancers and their choices of how much to donate to different cancers reveal how they view luck vis-a-vis individual choices. We find that provision of information on lifestyle-related causes of cancer adversely affects contributions. In contrast, information on hereditary causes has a positive effect on donations. Thus, a non-negligible share of the donors lean toward choice egalitarianism, which conditions outcomes on the potential beneficiaries' choices, and this is mainly due to preferences of women who tend to strongly favor choice egalitarianism.

The remainder of this thesis is organised as follows. Chapter 2 presents a literature review. The final chapter of the thesis summarises the findings and draws conclusions.

Chapter 2

Literature Review

2.1 Background – Welfare Economics

Policy-choice problems occur frequently in almost every aspect of social life, including the distribution of resources and the construction of legal systems. Since each policy choice has its own specificities, it is rarely easy to select the optimum policy from various alternatives. In order to contribute to these policy-choice questions, economists and other social scientists have invested considerable effort in constructing procedures for assessing policies. One sub-branch – *social choice theory*[1] – has also been developed to allow for methodical discussion, in which a consequentialism view is adopted such that the criterion used for policy assessment is its effect on the wellbeing of social members.

Since this thesis is built also within the framework of welfare economics, it will be useful to clarify several key assumptions. The illustration of how these assumptions are

[1] One great thinker that must be mentioned is Kenneth Arrow. His notable contribution, known as Arrow's impossibility theory (1951), is perhaps the most influential one in social choice theory. We deliberately did not choose to review the literature based on Arrow's theorem because another fundamental theory, namely Harsanyi's (1953, 1955, 1977b) utilitarian impartial observer theorem, is more relevant to our research interest. Different to Arrow's framework, utility in Harsanyi's framework is assumed to be both cardinal and interpersonal comparable. These two assumptions will be explained in this section shortly, and they are also the reason that Harsanyi's theorem can escape the impossibility theory (Harsanyi 1979, p.303).

generated is by examining two essential steps in developing an appropriate procedure for policy assessment. The first step is to transfer the effect of the policy into a standardised measurement, since the relative effect of different policies cannot be evaluated without a comparable measurement. For example, if a society is facing a choice of whether or not to introduce a more aggressive tax policy so that more money can be allocated to public health, then, for each social member involved, the pain of paying more tax and the enjoyment of having more secure healthcare are measured on different dimensions and are thus incomparable. In order to allow for comparison, economists often use the term "utility" as a proxy for this standardised measurement. In welfare economics, the term utility refers specifically to policy effects on the wellbeing of individuals. A positive change in utility is allocated to each social member if a policy increases his or her wellbeing, while a negative change in utility is allocated if a policy harms his or her wellbeing. Policy assessments are then based solely on these standardised wellbeing measures, namely utilities.

The second, more challenging step is to aggregate information on all social members' wellbeing so that an overall judgement about the social desirability of a policy can be formed. Several normative ways have been suggested, which will be examined in detail shortly. However, another key assumption should be clarified here, which is that "utilities" in welfare economics are normally interpersonally comparable.[2] This is because policy evaluations are not always straightforward cases in which all social members are made better or worse off. In most cases, a policy may benefit some social members' wellbeing while simultaneously harming others.[3] Before a decision is made on this kind of question, a society needs to be able to compare one individual's loss with another's gain. In other words, the ability to balance the gains and losses allocated to different social members is an essential requirement

[2]Harsanyi (1955, p.316-320) repeatedly emphasised this assumption, which appeared unusual in his time. Researchers are now almost agreed that this kind of interpersonal comparison is meaningful, at least to some extent. The utilities in this thesis are interpersonally comparable.

[3]For example, in *Iphigenia in Aulis*, the Greek leader needed to decide whether to sacrifice his daughter so that the fleet could sail for Troy.

in conducting policy assessment. A less explicit assumption in a considerable part of the literature is that "utility" not only has an ordinal meaning but also a cardinal meaning.[4]

With interpersonally comparable utility built in, the quality of a policy can be uniquely captured by the aggregated utility value it brings to the whole society. Thus, under welfare economics, the policy-making problem is simplified to the establishment of a policy that brings maximal utilities to the society.[5]

2.2 Harsanyi's Work

This section examines the main difficulty remaining in policy assessment; that is, what normative principle should be followed when society aggregates information on all social members' wellbeing? In other words, how should the utility that a policy brings to society be maximised? No consensus has been reached within welfare economics; however, most discussions can be linked to Harsanyi's (1953, 1955, 1977b) "(preference) utilitarian principle". This principle suggests that the social welfare function should be an additive function of all individuals' wellbeing and "*the best policy should be the one which brings the largest arithmetic mean of individual wellbeing*". This principle has been both very influential and much disputed.

In this section, Harsanyi's work and his contribution will be outlined, highlighting where disputes have been raised and how they may possibly be solved. In order to standardise the terminologies, the term "individual" is used to refer to a social member involved in a social choice problem; the term "observer" is used to refer a social planner, policy maker or legislator.

[4]The difference between the ordinal and cardinal utility is the following: if policy A, B can produce society utility values of 10 and 5 respectively, then an ordinal utility means the numbers can only be used to generate an order between the social desirability of A,B, while the cardinal utility means that not only can it be said that A is more socially desirable than B, but also A is exactly twice as desirable as B. The use of cardinal utility can be seen in Harsanyi (1953, 1955, 1977b) and Fleming (1952), among others.

[5]This is analogous to utility maximisation in individual decision making.

Two main results from Harsanyi's work have made contributions to two classes of literature. The first is Harsanyi's "impartial observer theorem", which is linked to a long tradition of moral philosophy and aims to determine how to be a rational, impartial and sympathetic observer. The second is Harsanyi's "social aggregation theorem", which is linked to discussion of how individual preferences can be aggregated to produce social preferences. Although both results will be discussed below, the main focus of this thesis will be on the first result.

2.2.1 Impartial Observer Theorem

Harsanyi's work on the impartial observer theorem will be elaborated with reference to his answers to the following three questions:

- Question 1: In what situations can a choice made by an observer be called "moral"?
- Question 2: What behavioural patterns should individuals and observers hold?
- Question 3: What behavioural principle should an observer follow?

This approach is a reasonable way to proceed, as Harsanyi's work is indebted to the influence of three intellectual traditions and one analytical branch newly-developed in his time. Question 1 refers to the influence of the moral tradition of the sympathetic impartial observer that dates back to Smith (1759), and the moral tradition of pursuing universal principles that began with Kant (1785); Question 2 refers to the influence of the newly-developed analytical framework of rational behaviour (Von Neumann and Morgenstern, 1944; Marschak, 1950); and Question 3 refers to the influence of the utilitarian tradition (Bentham, 1789; Mill, 1863; Sidgwick, 1874; Edgeworth, 1881).[6]

[6]Harsanyi admitted these influences himself. See Harsanyi, 1977a, p.623-624 and Harsanyi, 1955, p.312-313.

- Question 1: In what situations can a choice made by an observer be called "moral"?

Largely influenced by Smith and Kant, Harsanyi pursued a universal criterion of moral choices.[7] He argued that a choice can be called "moral" only if a social planner achieves the highest degree of impersonality when making a decision; that is, the choice is made in complete ignorance of the planner's relative position. This idea was first published in 1953 (Harsanyi, 1953, p.434-435) and extended in 1955 (Harsanyi, 1955, p.316). To capture this idea, suppose someone makes the statement: "I would like to support the new tax policy, as it will improve my wellbeing". This may be a very reasonable choice from the speaker's point of view; however, it can hardly been regarded as a moral one, since this choice is obviously based on self-interest. It would become a moral choice if someone were to say: "I would like to support the new tax policy, even though I don't know if it will improve or harm my wellbeing".[8] In Harsanyi's terms, the first statement is a personal preference, while the second is a moral (or social) preference. A moral choice can only be derived from a moral preference, where the decision maker is unaware of his or her position in society.

This idea sheds light on how researchers have proposed that moral choices should be made. Vickrey (1945) had previously proposed a similar idea, but Harsanyi's work was carried out independently. Later, Rawls (1971) again independently proposed a similar model, called the "original position". Thus, researchers have to some extent agreed that a moral choice can be made only if the social planner is placed behind a veil of ignorance, without the knowledge that he or she can be anyone in society.

Harsanyi proposed that one should deprive oneself of personality and imagine oneself as an impartial observer. When facing a choice behind the veil of ignorance, an impartial observer may adopt all of an individual's preferences through the so-called "acceptance

[7] For a detailed discussion of what "universal" means here, see also Hare (1952, part III "Ought").

[8] A similar debate can be found in the literature, for example Harsanyi (1953, p.434) and Pattanaik (1968). In reality, people tend to make judgements on various problems based on their own emotionally affected and biased positions. However, there are occasions when most people will show the same moral preference. For examples, see Smith (1759).

principle" (Harsanyi, 1977b, p.52). That is to say, although impartial observers are unaware of their identity, they can still adopt one individual's preference if they imagine they would become that individual, and adopt another individual's preference if they imagine they would become that individual. This acceptance principle actually allows an observer to collect information about the preferences of all individuals involved in decision making.

Once the impartial observer is under a veil of ignorance and has collected all personal preference information under the principle of acceptance, the next question is what behavioural patterns this observer should have.

- Question 2: What behavioural patterns should individuals and observers hold?

Harsanyi appears to have been fascinated by axiomatic work on the rational behaviour theorem, which was newly developed at that time and attributable mainly to von Neumann and Morgenstern's work (1944). Von Neumann and Morgenstern suggested four postulates reflecting their idea of behavioural patterns for a rational individual, leading to the development of the expected utility theorem. This theorem suggests that, when faced with decision-making problems, an individual satisfying these four postulates will behave as if maximising the expected utility. These axiomatic requirements were similarly proposed by Marschak (1950), which were directly adopted by Harsanyi (1955, p.312-313). These postulates will be introduced according to their current refinements, rather than the original versions in Harsanyi's work, but the ideas behind these behavioural patterns remain the same. The four postulates are:

Postulate 1: Complete order

Postulate 1.1 Completeness

Postulate 1.2 Transitivity

Postulate 2: Continuity

Postulate 3: Nontriviality

Postulate 4: Independence

Harsanyi assumed that both individual and observer preferences should "naturally" satisfy these four postulates according to the rational behaviour theorem. Therefore, the first two axioms in Harsanyi's framework are individual rationality and social rationality.

Axiom A: **Individual Rationality** – All individuals in a society satisfy the four postulates; thus, they are all expected utility maximisers.

Axiom B: **Social Rationality** – The moral preferences of the observer satisfy the four rational behaviour postulates; thus, the observer is an expected utility maximiser.

Another of Harsanyi's axioms seeks to link individuals' preferences with the observer's preference. This axiom seems quite natural, requiring an observer to be indifferent between any two social situations if all individuals are indifferent.

Axiom C: **Individualism** – If all individuals are personally indifferent between two social situations, the observer should also be indifferent between these two social situations.

- Question 3: What behavioural principle should an observer follow?

The third question Harsanyi aimed to answer is what kind of social welfare criteria an observer should follow, given the above assumptions. He concluded that the three axioms would imply that the observer's social welfare function is a linear combination of all individual utilities, resulting in utilitarian criteria. It should be emphasised that Harsanyi did not take the utilitarian principle as an ideological doctrine, like the classical utilitarian school, but showed that the utilitarian principle can be derived from several pre-agreed axioms. Although some of these axioms may require reconsideration, Harsanyi's contribution in bringing analytical and axiomatic thinking to policy assessments was considerable. This work can be regarded as the starting point for modern analysis, enabling researchers to escape from a relatively implicit discussion of ideology and to exchange explicit opinions on the observer's behavioural pattern.

2.2.2 Social Aggregation Theorem

Hansanyi's (1955) second result is the aggregation theorem. No less important than the impartial observer theorem, this theorem focuses on two different ways of evaluating social policies. One is to evaluate from the individual perspective; that is, to collect the individual utilities of policies, and then aggregate them. The other is to evaluate from a social state-based perspective; that is, to evaluate the social utilities under each state, and then aggregate these evaluations. Harsanyi showed that, under his utilitarian principle, the social order of policies generated by these two approaches are identical; that is, the sum of the expected individual utilities should be equivalent to the expected social welfare for each social prospect. This result is normally known as Harsanyi's aggregation theorem.

Generalisation of Harsanyi's aggregation theorem can be seen in Blackorby, Donaldson and Weymark (1999), Blackborby, Donaldson and Mongin (2004) and Mongin and Pivato (2015). Blackorby, Donaldson and Weymark (1999) rephrase Harsanyi's theorem in terms of state-contingent alternatives (i.e. Savage's act space) rather than a classical lottery space, with a subjective expected utility theorem (but specifically assuming that the subjective measures are shared by all individuals). Blackborby, Donaldson and Mongin (2004) assume that the aggregation theorem holds and check what conditions are required. From their results, the social welfare function needs only to be weighted utilitarianism, which is weaker than utilitarianism, and only requires additively separable individual utilities. Mongin and Pivato (2015) use a group of monotonicity axioms, such that a fully separable space is constructed, showing that a similar result holds.

2.3 Criticism of Harsanyi's Work

Harsanyi's model has received considerable attention; however, it is apparent that some researchers have been unhappy with its utilitarian criteria, especially the aspect that social welfare function should be a linear function of individual utilities. While earlier researchers objected to the utilitarian criteria for being too fair,[9] later criticism of Harsanyi's theorem came mainly from three researchers who thought that this criterion was not fair enough: Diamond (1967), Rawls (1971) and Sen (1979). These criticisms all treated the lack of consideration of fairness in the utilitarian criterion as a serious flaw. However, of the three, only Diamond's criticism sought to improve Harsanyi's approach within welfare economics. As Harsanyi himself noted, "The most specific criticism (among the three) was Diamond's, who at least clearly recognized that my theory can be rejected only if one rejects one or more of its axioms" (Harsanyi, 1975a, p.314). Rawls' and Sen's criticisms were not constructive, but rather aimed to provide alternative theorems. Thus, Rawls' and Sen' disagreements did not lie explicitly in Harsanyi's framework but at a more fundamental level. Their criticism also partly took a non-welfarism perspective, which is a broader debate outwith the scope of this thesis. These criticisms are reviewed in the following sections.

2.3.1 Diamond's Criticism

Diamond's disagreement with Harsanyi was with regard to the certainty principle, which lies in the independence axiom of social rationality. Example 1 provides a better understanding of Diamond's argument.

Example 1: *Suppose a society, consisting of two individuals, Ann and Bob, is*

[9] For example, Robbins (1938) thought that the assumption advocated by utilitarianism that all individuals should be counted equally was too fair. He wrote, "I do not believe, and I never have believed, that in fact men are necessarily equal or should always be judged as such."

facing a policy-making problem of how to allocate an indivisible good. Assume, for both individuals, that obtaining the good represents a utility of 1, while obtaining nothing represents a utility of 0. Thus, the utility vector (0,1) can be used to capture the social state in which Bob gets the good and Ann gets nothing, while the utility vector (1,0) can be used to capture the social state in which Ann gets the good and Bob gets nothing. Two alternative policies, Policy 1 and Policy 2, are available. Policy 1 would allocate the good to Bob (yielding a utility vector (0,1)) and Ann (yielding a utility vector (1,0)) with equal probabilities, while Policy 2 would allocate the good to Bob for sure. If society regards these two individuals as equally important, would the observer be indifferent between Policy 1 and Policy 2?

Obviously, Harsanyi believed that society should be indifferent between the two policies, as they yield the same expected utility. However, Diamond's view was that Policy 1 is strictly preferable, since it gives Ann at least a "fair shake", while Policy 2 does not. Diamond summarised his objection as follows: "I am willing to accept the sure-thing principle for individual choice but not for social choice, since it seems reasonable for an individual to be concerned solely with final states while society is also interested in the process of choice" (Diamond, 1967, p.766). In other words, a fair procedure would bring additional utility for Diamond but not for Harsanyi.

2.3.2 Rawls' and Sen's Criticisms

Rawls sought to contribute to the traditional school of social contract, as represented by Locke, Rousseau and Kant (Rawls, 1971, viii; see also Harsanyi 1975b, p.594), and restricted his debate to a sense of justice. Thus, Harsanyi and Rawls' debate can be seen as an extension of the debate between two traditional schools of social choice: the utilitarian school and the social contract school. In Rawls' work, policy is agreed through a collective rationality (for example, Rawls, 1971, p.490); that is, a contract that all individuals would agree to when they are at the original position and unaware of their identities, rather than each individual being able to adopt other individuals' preferences through the acceptance principle and thus being able to make a social decision, as in Harsanyi's framework. Based on this fundamental difference, Rawls developed a different systematic account. His second principle, known as the "difference principle", directly contradicts Harsanyi's utilitarian principle. In economics, this principle is interpreted as the "maximin approach", where the quality of a policy is decided by the utility level received by the worst-off people. However, Rawls' criterion was initially applied to social primary goods rather than utilities.[10] If the difference between primary goods and utilities is waived, then Rawls' main disagreement with Harsanyi is a question of whether the probability should be used. As noted by Harsanyi (1977a, p.634):

> Yet the difference does not lie in the nature of the two models, which are based on almost identical qualitative assumptions. Rather, the difference lies in the decision rules, namely the maximin principle, which is fairly different in that Rawls avoids any use of numerical probabilities.

The difference between Rawls and Harsanyi lies in what knowledge the observer has when facing a choice problem behind a veil of ignorance. For Harsanyi, the observer is believed to have full information and is only unaware of his or her position; thus a prior can be formed to make a decision. For Rawls, the impartial observer is given even less information than in Harsanyi's case, and this information is insufficient to form a prior in decision making. Thus, the principle suggested by Rawls is the maximin approach.

Sen's disagreement with Harsanyi was grounded mainly in his concern that, given the complexity of the world, the welfare economics framework is unable to capture everything

[10] This difference is also noted by Sen (1979, p.205).

fully. Every principle, including the utilitarian principle, seems able to interpret only some legitimate applications (Sen, 1977, p.298). Sen's (1979) focus was the equity of the non-welfarism criterion of "basic capabilities",[11] rather than utilities or primary goods. For him, the concept of basic capabilities meant "a person being able to do certain basic things", which includes "the ability to move out ... the ability to meet one's nutritional requirements, the wherewithal to be clothed and sheltered, the power to participate in the social life of the community" (Sen, 1979, p.218). This work can be seen as a natural extension of Rawls' work, shifting attention from goods to "what goods do to human beings" (Sen, 1979, p.219).

2.4 Attitudes Towards the Criticism

Given these and many other criticisms, it appears that Harsanyi's utilitarian principle suffered from serving as a doctrine for solving all social choice problems. Indeed, everyone, including researchers, has their own opinion on how a social choice should be made, and there is unlikely to be a universal principle agreed by everyone. Furthermore in applying this theoretical principle to real-life scenarios, which are more murky and complex than those in theoretical discussions, such a universal principle is even less likely to exist. Thus, it is unsurprising to see that each individual principle, including the utilitarian principle, has restricted explanatory power and cannot capture every scenario.

However, this does not mean that the process of searching for such a universal principle should be stopped. On the contrary, it means that more research is necessary, which is consistent with trends in the development of the literature. Most researchers have correctly understood the initial purpose of the harsh criticisms. The criticisms mentioned above did not seek to deny Harsanyi's contribution completely, but rather to provide thoughts on other dimensions not captured in the utilitarian framework.[12] If some thoughts are believed to be

[11] The capability to pursue happiness is also mentioned by Robbins (1938, p.636).
[12] This can be seen, for example, in Rawls (1971) and Weymark (1991).

reasonable and should be incorporated into the discussion, then the next question is how can this be done? A typical approach, as seen in other sciences, is to build constitutively on already familiar elements[13]. Therefore, in the area of social choice problems, the literature has been built mainly on Harsanyi's framework. A utilitarian framework, rather than an alternative, has usually been chosen as the first choice for such a construction because it is a well-structured and mathematically elegant framework, which significantly reduces the potential risk of causing confusion.

2.5 Improvements Based on Harsanyi's Framework

This section examines how improvements have been developed over the last few decades. As discussed in Section 2.4, these have been made mainly by extending the domain in which Harsanyi's theorem can be applied by constantly relaxing or replacing his axioms and conditions. The focus of this thesis is on how to include the fairness concern in social welfare evaluations, a classical debate that also lies at the core of the criticisms mentioned in Section 2.3.

As shown above, the aspect of Harsanyi's utilitarianism theorem which has received most criticism is its indifference to the distribution of utilities; that is, an impartial observer under Harsanyi's theorem would not consider the fairness issue at all. This kind of impartial observer seems not to match real observations and is also ethically unacceptable to most researchers. It may be not necessary to pursue the other extreme, as Rawls did, in which a social observer is only concerned about fairness, but it would be reasonable to believe that the impartial observer cares about fairness, and that sometimes the effect of this consideration may be a determining factor in the decision. If the concern of fairness is indeed taken as an indivisible part of the impartial observer's preference, as observed by most social planners

[13]See Machina (1987, p.133) for an application of this approach in economics.

in daily life, then the question is how to introduce this concern to the observer's preference.

In Harsanyi's work, it can be seen that there are two different ways of calculating social welfare. The first is to obtain the expected utility for each individual, and then aggregate these expected individual utilities into an expected utility theorem; the second is to calculate the expected welfare under each possible state, and then aggregate all these states into an expected utility theorem. According to the aggregation theorem, these two ways will be identical under utilitarianism. However, this result will no longer hold if consideration of equality is introduced. More specifically, within Harsanyi's framework, introducing the equity consideration at a social level (while keeping other things unchanged) means that an impartial observer has to make a decision between "*equalisation of expected well-beings*" and the "*expected equalization of actual well-beings*" (Adler and Sanchirico, 2006). This is why there are two main directions for improving utilitarianism, namely the ex-ante and ex-post approaches.

We focus on the ex ante approach in this thesis, since generalisation of Harsanyi's utilitarianism has normally been conducted within this approach.[14] Also, our work presented in Chapter 3 is one of them. This generalisation has taken from at least two directions.

One is to bring heterogeneity to Harsanyi's framework, which can be attributed to the individual rationality axiom. In his original model, Harsanyi (1977b, p.51-52) required that impartial observers should have uniform "*imaginative empathy*". Similar requirements were also raised by Arrow (1977, p.159) using different terminologies. This requirement basically stipulates that, when impartial observers imagine being each and every individual, these extended preferences must coincide across all impartial observers. This requirement ensures that all impartial observer preferences are objective, thus reducing the social choice

[14]For readers who are interested in the ex post approach, noble discussions can be found in Broome (1978, 1984), Harel et al. (2005), Adler and Sanchirico (2006), Fleurbaey (2010), Grant et al. (2012a), and among others. For readers who are interested in insightful theories that combine ex ante and ex post approaches, see Machina (1989), Ben-Porath et al. (1997), Gajdos and Maurin (2004), Chew and Sagi (2012), Saito (2013), and among others.

problem to an individual choice problem.[15] However, this requirement seemed too demanding and unrealistic for authors such as Broome (1993) and Mongin (2001), who argued that impartial observer preferences should be extended from objective assessments to subjective assessments. This is analogous to Anscombe and Aumann's (1963) work on the rational behaviour theorem, from which Harsanyi initially developed his theorem.[16] For example, Broome (1993, p.65) wrote:

> But that [uniform imaginative empathy] is a fantasy. My position as an academic causes me to have particular values. Since those are my values, I cannot escape them, even when I am forming my preferences about lives in which I would not have those values.

In response to this concern, Grant et al. (2010) used an ex-ante approach to capture the concern about individual differences in attitudes to risk[17]. Returning to Example 1, it might be the case that Ann and Bob have different attitudes to risk, for example Bob is more risk loving than Ann. It would then be reasonable to imagine that the impartial observer, who could fully adopt Ann and Bob's preferences, would be concerned about this difference and would take it into consideration.

An interesting variant under this direction must be mentioned. Proposed by Karni and Safra (2000), it is now clear that Diamond's (1967) criticism can also be counteracted

[15]This is also observed by Rawls (1971, p.24), who wrote: "The principle of choice for an association of men is interpreted as an extension of the principle of choice for one man. Social justice is the principle of rational prudence applied to an aggregative conception of the group."

[16]Savage (1954, p.7), a great contributor to the rational behaviour theorem, wrote: "Rational behaviour assumes that all individuals in the same position, given the same information will make the same choice". He also stated his concern about this assumption: "personally, I doubt it can even be matched roughly with the reality" (Savage, 1954, p.7). In order to capture more real situations, Anscombe and Aumann (1963) extended the concept of objective probabilities (as in a dice game) to subjective probabilities (as in a horse game).

[17]This concern is remarked by Pattanaik (1968, p.1165-1166), who wrote: "what we are actually doing is to combine attitudes to risk of more than one person, and although the single individual's choice among risky prospects may satisfy the rules of consistency posited by the expected utility axioms, there is no reason to expect such consistency in the case where more than one person's attitudes are involved."

if individuals are modeled as *self-interest-seeking moral beings*[18]. These self-interest-seeking moral beings are different to individuals assumed in Harsanyi's work, since the former adopts the moral preference even under normal circumstances (not only behind a veil of ignorance).

Another direction in the generalisation of Harsanyi's work has been to technically reduce its strict requirement by restricting the domain in which impartial observer preferences are defined. In Harsanyi's original work, an observer is required to order all possible extended lotteries, which is quite demanding. Karni and Weymark (1998) restricted the domain from all possible extended lotteries to only impartial extended lotteries, a smaller domain considering only extended lotteries with an equal chance of being any individual in society. It was noted that situations in which different individuals might face different social alternative lotteries were allowed. Karni and Weymark (1998, p.327, Assumption A.5) further strengthened the acceptance principle and showed that an analogous result to Harsanyi's theorem could be obtained.

Safra and Weissengrin (2003) restricted the domain to constant extended lotteries – lotteries in which every individual in society faces the same social lottery. In this framework, impartial observers may imagine themselves to have different probabilities of becoming different individuals; that is, under each constant extended lottery, the impartial observer may have different identity lotteries. In addition, rather than strengthening the acceptance principle, as in Karni and Weymark (1998), they replaced the independence axiom by using a so-called "substitution" axiom (the latter is slight stronger than the former for non-convex sets, but equivalent for convex sets), and obtained the same result as Harsanyi.[19]

[18]This way of modelling follows the tradition of Hume (1740), in which possessing a moral sense is interpretted as inherent in human beings. This view is originally expressed in choice behaviors in Karni (1996) and then treated axiomatically in Karni and Safra (2002).

[19]See Grant et al. (2012b) for a third alternative in this direction. In this work, the domain was restricted to identity lottery/social alternative pairs, which is also strictly smaller than the original domain in Harsanyi's work.

2.6 Fairness in the Real World

Fairness concerns are of paramount importance, not only for normative studies of decisions behind a veil of ignorance (the focus of Sections 2.1 to 2.5), but also for various studies focusing on real-life decisions. Such concerns must be delicately considered, for instance in resolving social issues such as the siting of nuclear waste facilities (Oberholzer-Gee, Bohnet and Frey, 1997), unemployment due to unfair wages (Akerlof and Yellen, 1990), the distribution of healthcare resources (Cappelen and Norheim, 2005) and many other issues. Chapters 4 and 5 of this thesis will discuss concerns about justice in the context of charitable donations to cancer research. The primary goal of these two chapters is to understand how justice, alone or in tandem with other concerns (such as self-interest), affects donation behaviours.

In this thesis, fairness is viewed as a *given* motive that affects people's behaviours. Building on the traditional belief that people are motivated *exclusively* by self-interest, the fairness motive has been introduced into mainstream economics by several pioneering researchers. For instance, Rabin's (1993) incorporation of the fairness motive into game theory modelling inspired numerous subsequent studies; Yaari and Bar-Hillel (1984) and Kahneman, Knetsch and Thaler (1986) used surveys and rich vignettes to initiate empirical studies testing people's fairness views in real-life contexts; and Karni and Safra (2002) included moral value judgements (on procedural fairness) in axiomatisations of individual preference. As a result of these contributions and many others, Fehr and Schmidt (1999, p.817) stated almost two decades ago that "By now we have substantial evidence suggesting that fairness motives affect the behaviour of many people". Today, central debates around fairness (e.g. Cappelen et al., 2013; Brock et al., 2013; Cettolin and Riedl, 2016) have progressed from *whether* it affects people's behaviours, to *how* it affects them. The latter is the focus of this review.

When dealing with real-life decisions, a natural but important first question is: justice for *whom*? In other words, whose wishes for justice should be satisfied in social decisions such as resolving conflicts? Under a democratic system of governance, researchers commonly focus on justice for the *general public.* As Scitovsky (1986, p.3) states, "an important part of the economist's task is to find out how well the production and distribution of goods and services conform to the public's wishes. The first thing to ascertain in this connection is what the public's wishes are." In both branches of literature to which this thesis contributes – the construction of justice theories that are in line with generally accepted values, and the collection of empirical evidence that reveals the public's wishes on justice – the role of the general public is tacitly acknowledged.

Since the public's wishes for justice are wide-ranging, a second question must be clarified for this thesis is: justice for *what*? To answer this question, categorisation of the fairness literature is required. According to Konow (2003), the literature on fairness can be divided into four distinct "families": (1) the equality and need family (e.g. Rawls, 1971; Sen, 1979), which calls for equal satisfaction of needs; (2) the utilitarianism and welfare economics family (e.g. Harsanyi, 1977b), which is based on consequentialist ethics; (3) the equality and desert family (e.g. Nozick, 1974), which is constructed on the basis of proportionality and individual responsibility; and (4) the context family (e.g. Elster, 1992; Young, 1994), which emphasises that justice is context-dependent. Chapter 3 of this thesis will focus on the *utilitarianism and welfare economics* family, while Chapters 4 and 5 will refer to the *equality and desert* family.

As the utilitarianism and welfare economics family has already been discussed in Sections 2.1 to 2.5, the focus of this section is on the equality and desert family. Regarded as "the intellectual progeny of two philosophical traditions: the distributive justice theory of Aristotle and the natural law/desert theory of John Locke" (Konow, 2003, p.1206), this family aims to establish the impact of individual factors on just allocations. Individual factors with distinct individual responsibilities contribute to outcomes, and deserved allocations for each agent relate to these individual responsibilities. Hence, the key issue of debate is about factors for which each individual should be held accountable.

This thesis is interested in only two individual factors: luck and choice. Fairness principles refer to these two factors can be differentiated according to the degree to which individuals are held responsible. At one end of this spectrum of fairness principles is strict egalitarianism (Nielsen, 1985), which does not hold individuals responsible for any causes of inequality. According to the principle, social redistribution should be based solely on outcomes. At the opposite end of the spectrum is libertarianism (Nozick, 1974), which postulates that individuals should bear full responsibility for their circumstances, even if they are caused by bad luck. Some theories of distributive justice combine egalitarian principles with concerns for individual responsibility. One of the most notable among these is choice egalitarianism (Dworkin, 1981a, 1981b; Arneson, 1989; Cappelen et al., 2013), which holds people responsible for their choices but not for their luck.

Experimental evidence on these competing fairness principles regarding choice and luck has been collected by Konow (2000) and Cappelen et al. (2007, 2013).[20] These studies were carried out within the framework of two-stage dictator games, in which people's luck and choice are traceable.

Konow's (2000) study focuses solely on choice egalitarianism.[21] His main finding is that, when fairness is the only concern, splits by benevolent dictators, whose pay-offs are fixed and independent of their choices, follow choice egalitarianism exactly. That is, such splits are proportional to the agents' respective input levels when the produced resource is determined through the agents' choice, but are almost always equal when the produced resource is determined by random luck.

Rather than examining whether choice egalitarianism alone matters to participants, as in Konow (2000), Cappelen et.al. (2007) seek to check the prevalence among participants

[20] See also Frohlich et al. (2004), Cappelen et al. (2010) and Krawczyk (2010).

[21] For simplicity of terminology, "choice egalitarianism" is used to refer to what Konow (1996, 2000) called the "accountability principle". These two principles are identical in this context.

of the three fairness principles. They mimic a genuine investment environment, in which participants are able to choose the percentages of endowment they wanted to invest, while the rate of market return is determined by luck. According to their estimation, substantial proportions of participants uphold each of the fairness principles, with liberal egalitarianism, strict egalitarianism and libertarianism supported by 38.1, 43.5 and 18.4 percent respectively.

The main contribution of Cappelen et al. (2013) is their examination of differences in fairness views between third-party (impartial) spectators and stakeholders. According to their estimations, substantial proportions of both spectators and stakeholders uphold the three discussed fairness principles, and the distributions of these population shares are very similar for spectators and stakeholders. Based on these results, the authors conclude that spectators and stakeholders act as if they hold the same fairness views.

These previous laboratory experiments provide important insights into possible behaviour in real-life situations and guidance for theoretical developments. The potential for these predictions to be generalised to a wide range of domains hinges on combining theory with empirical evidence from naturally occurring environments (Winkler and Murphy, 1973; Harrison and List, 2004). Chapter 4 makes a first step in this direction in the context of understanding individuals' attitudes to different distributive justice principles.

Another gap in the literature is variety in people's perceptions. Although often assumed away in economics to avoid analytical complexity, this variety is itself an inseparable part of decision making (Simon, 1955). Chapter 5 contributes to this area by considering individual risk perceptions for the first time in the justice literature.

Chapter 3

Fairness and Utilitarianism without Independence

Abstract: In this work we reconsider Harsanyi's celebrated (1953, 1955, 1977b) utilitarian impartial observer theorem. Departing from Harsanyi's individual-centered approach, we argue that, when societal decisions are at stake, postulates must not be drawn from individualistic behavior. Rather, they should be based on societal norms. Hence, notions like societal fairness should explicitly be the guiding principles. Continuing this line of thinking, we state and prove a utilitarian result that, rather than being based on the independence assumption, is based on the notion of procedural fairness and on symmetric treatment of societal and individual lotteries.

"An axiomatic justification of utilitarianism would have more content to it if it started off at a place somewhat more distant from the ultimate destination" (Sen 1976, page 251)

3.1 Introduction

In this work we reconsider Harsanyi's celebrated (1953, 1955, 1977b) utilitarian impartial observer theorem. We propose an approach that puts more emphasis on procedural fairness and we offer a utilitarian result that does not use the independence assumption.

Harsanyi analyzed a society that needs to choose among alternate social policies, each of which is a probability distribution over a given set of social actions, where the latter associate outcomes with the society's members. Every social lottery ℓ induces a lottery ℓ_i on individual i. Individual i's preferences $\succcurlyeq_i$ are known and different individuals may possess distinct preferences.

To help determine the optimal social policy, Harsanyi suggested that every individual is endowed with social preferences. Individuals may develop these preferences by adopting the role of an impartial observer, thus disregarding their true identities and acting behind "a veil of ignorance". Therefore, the impartial observer can form her social preferences by imagining that she faces not only a lottery ℓ over social actions, but also a lottery γ over identities. Elements of γ can be interpreted as weights associated by the impartial observer with the different individuals. Then, the optimal social policy is determined by restricting attention to the equiprobable lottery $\gamma^e = \left(\frac{1}{n}, ..., \frac{1}{n}\right)$.

Harsanyi argued strongly for "Bayesian rationality". That is, he assumed that (among the other Bayesian postulates) all individuals satisfy the *independence assumption* of the expected utility theory, both at their personal and social preference layers. Harsanyi claimed that this "sound" axiom, together with the so-called *acceptance principle* (that an impartial observer fully adopts individual i's preferences if she imagines becoming that individual for sure), would force the impartial observer to be a (weighted) utilitarian. More formally, over all extended lotteries (γ, ℓ) in which the identity and the action lotteries are independently distributed, the impartial observer's preferences admit the following representation:

$$V(\gamma, \ell) = \sum_{i \in \mathcal{I}} \gamma_i U_i(\ell_i)$$

where γ_i is the probability of assuming person i's identity and $U_i(\ell_i) := \sum_x u_i(x)\ell_i(x)$ is person i's von Neumann-Morgenstern expected utility.

Like Harsanyi, most authors who derived modifications of the utilitarianism result within the impartial observer framework always assumed the independence axiom (see the works of Weymark (1991), Karni (1998) and Grant, Kajii, Polak and Safra (2010; henceforth GKPS)).[1] Notable exceptions within the related social aggregation framework are Blackorby, Donaldson and Mongin (2004) and Mongin and Pivato (2015).[2]

Interestingly, Harsanyi's entire emphasis on Bayesian rationality was based on an individual centered approach. Firstly, he assumed that rational individuals must satisfy the independence assumption and secondly, he claimed that society, by its need to be at least as rational as its members, must also satisfy independence (Harsanyi 1975). We disagree with Harsanyi on this. Instead we argue that when societal decision problems are at stake, postulates must not be drawn from individualistic behavior. Rather, they should be based on societal norms. Hence, when social preferences are formed, issues like societal fairness and equity should explicitly be the guiding principles.

In this work we focus on procedural fairness. This principle was first advocated

[1] A similar observation holds for most of the literature dealing with Harsanyi's social aggregation theorem. See Zhou (1997), Dhillon and Mertens (1999), Gilboa, Samet and Schmeidler (2004) and Fleurbaey and Mongin (2012).

[2] Unlike the other works (including the current one), these authors consider both *ex post* and *ex ante* analyses (and thus are able to employ Gorman's (1968) separability theorem).

by Diamond (1967) and was strongly supported by Sen (e.g., 1977). Its essence can be illustrated by the following example, which is an adoption of Diamond's example from the social aggregation framework to the impartial observer one. Consider a society that needs to decide on how to allocate an indivisible good between two individuals, 1 and 2, and let action a^i denotes allocating it to individual i. Suppose, as Diamond did, that $u_i(a^i) = 1$ for both i and $u_i(a^j) = 0$ for $i \neq j$ (that is, both individuals like the good, receive a utility of one unit from having it and zero otherwise). As was noted above assume that, when making a choice, the impartial observer considers the equiprobable identity lottery $\gamma^e = \left(\frac{1}{2}, \frac{1}{2}\right)$ (that is, she gives equal weights to the two individuals). Also assume that she evaluates all four outcomes in full agreement with the two individuals and adopts their utilities. The example can be described by the table

	a^1	a^2
1	1	0
2	0	1

where individuals 1 and 2 correspond to the rows, actions a^1 and a^2 correspond to the columns and the entries represent the impartial observer's utilities. The impartial observer has two policies at hand: Policy (1), which allocates the good to individual 1 (this policy is equivalent to choosing action a^1 and facing the first column of the table) and Policy (2), which allocates the good to one of the individuals, depending on the outcome of a toss of a fair coin (this policy is equivalent to the action lottery $\frac{1}{2}a^1 + \frac{1}{2}a^2$). The value of Policy (1) for Harsanyi's utilitarian observer is $\frac{1}{2} \times 1 + \frac{1}{2} \times 0 = \frac{1}{2}$, as is the value of Policy (2): $\frac{1}{2}\left(\frac{1}{2} \times 1 + \frac{1}{2} \times 0\right) + \frac{1}{2}\left(\frac{1}{2} \times 0 + \frac{1}{2} \times 1\right) = \frac{1}{2}$. Hence, the impartial observer is indifferent between the two policies.[3] However, Diamond and Sen argued that policy (2) provides both individuals with a "fair shake" and hence the impartial observer might prefer it.[4] This notion

[3]Note that the impartial observer is also indifferent between a^1 and a^2.

[4]A long list of real-life applications supporting Diamond's fairness consideration is provided by Elster (1989).

of procedural fairness is expressed in our work by the notion of (weak) *convexity* over action lotteries: if, given the identity equiprobable lottery γ^e, the observer is indifferent between two action lotteries ℓ and ℓ' (while two individuals disagree on their ranking) then their mixtures cannot be worse than them.[5]

Working in a framework in which the basic building blocks are two different types of lotteries, those over identities and those over actions, raises a natural question: should these types be treated similarly? Harsanyi, by construction, implicitly assumed that they should. Furthermore, in his own response to Diamond's concern about fairness, Harsanyi (1975) argued that even if randomizations were of value for promoting fairness (which he doubted), any explicit randomization is superfluous since "the great lottery of (pre-)life" may be viewed as having already given each child an equal chance of being each individual. That is, it does not matter whether a good is allocated by a (possibly imaginary) lottery over identities or by a (real) lottery over actions. Put it differently, Harsanyi argued that we need to be indifferent between "accidents of birth" (identity lotteries) and real "life chances" (action lotteries). On this issue we agree with Harsanyi and just make this assumption explicit. We call it *source indifference.*

Despite its innocuous appearance, the conjunction of this assumption with procedural fairness turns out to be rather forceful. More precisely, the main result of this work shows that (assuming impartiality) convexity, source indifference and a stronger notion of acceptance are necessary, and sufficient, for utilitarianism.

Since the independence axiom is not assumed here, this result is novel and quite unexpected. Paraphrasing Sen's quote, we believe that one could hardly find an axiomatic justification of utilitarianism that starts off at a place that is more distant from the ultimate destination than ours.

[5]Unlike Epstein and Segal (1992), we do not assume that such mixtures are *always* strictly preferred. This is in agreement with Sen (1977), who argued that mixtures are not always superior.

Lastly, our result carries an 'impossibility flavor': if societies are required to exhibit *strict* inclination towards procedural fairness, then source indifference cannot hold. Therefore, to accommodate views of authors like Diamond and Sen, the impartial observer must display preference for action lotteries over identity ones. We elaborate on this in the concluding section.

This work is organized as follows: Section 3.2 sets up the framework, Section 3.3 presents the assumptions, Section 3.4 states, and explains, the utilitarian result and Section 3.5 concludes. Finally, proofs are given in Section 3.6.

3.2 Setup and Notation

Let $\mathcal{X} = [x_{\min}, x_{\max}] \subset \mathbb{R}$ be a compact interval representing all possible outcomes and let $\triangle(\mathcal{X})$ denote the set of outcome lotteries, endowed with the weak convergence topology. With slight abuse of notation, we will let x denote the degenerate outcome lottery that assigns probability 1 to outcome x. Let T be a denumerable set of potential individual types, where each type $t \in T$ is characterized by a preference relation over $\triangle(\mathcal{X})$ that is complete, transitive, continuous (in that the weak upper and lower contour sets are closed), increases with respect to first-order stochastic-dominance and its asymmetric part is nonempty. The set of individuals under consideration is $\mathcal{I} = \cup_{t \in T} \mathcal{I}_t$, where $\mathcal{I}_t$ is a denumerable (infinite) set of type t individuals. In the sequel, individuals are denoted by i, j (and their preferences by $\succcurlyeq_i$, $\succcurlyeq_j$), without explicitly specifying their types. A society I is a finite subset of $\mathcal{I}$. Note that, even though we allow for societies in which some individuals are of the same type, these individuals may receive different outcomes and hence they need not be treated similarly. Also note that our framework departs from Harsanyi's in that, instead of working with one fixed finite society, we consider all finite subsets of $\mathcal{I}$.[6]

[6]Dealing with a large set of potential members of various societies is justified by our pursuit for a general rule, to be applied to all societies. Note that Harsanyi too (like other scholars) wanted his theory to be

A social policy, or an action, associates an outcome with every individual and hence is represented by a function $a : \mathcal{I} \rightarrow \mathcal{X}$. The set of all actions, endowed with the corresponding product topology, is denoted by $\mathcal{A}$ (two extreme actions, $a_{\max}$ and $a_{\min}$, defined by $a_{\max}(i) = x_{\max}$ and $a_{\min}(i) = x_{\min}$ for all i, respectively, will be used in the sequel). Let $\triangle(\mathcal{A})$ denote the set of simple lotteries (lotteries with finite support) over actions, with typical elements denoted by ℓ. With slight abuse of notation, we will let a denote the degenerate action lottery that assigns probability 1 to action a. A lottery $\ell \in \triangle(\mathcal{A})$ is sometimes written as $\ell = \sum_{a \in Supp(\ell)} \ell(a)\, a$.

Following Harsanyi, an observer imagines herself behind a veil of ignorance, uncertain about which identity she will assume in the given society. Let $\triangle(\mathcal{I})$ denote the set of simple identity lotteries on $\mathcal{I}$, where typical elements are denoted by γ (where γ_i is the probability assigned by the identity lottery γ to individual i). These lotteries represent the imaginary risks in the mind of the observer of being born as someone else. With slight abuse of notation, we will let i denote the degenerate identity lottery that assigns probability 1 to individual i. An imaginary lottery $\gamma \in \triangle(\mathcal{I})$ is sometimes written as $\gamma = \sum_{i \in Supp(\gamma)} \gamma_i i$. When the observer is faced with pairs of identity and action lotteries, it is assumed that they are independently distributed.

The observer is endowed with a preference relation $\succcurlyeq$ defined over the space of all product lotteries $\triangle(\mathcal{I}) \times \triangle(\mathcal{A})$. We assume throughout that $\succcurlyeq$ is complete, transitive, continuous and that its asymmetric part $\succ$ is nonempty. These assumptions imply that $\succcurlyeq$ admits a (nontrivial) continuous representation $V : \triangle(\mathcal{I}) \times \triangle(\mathcal{A}) \rightarrow \mathbb{R}$. That is, for any pair of product lotteries (γ, ℓ) and (γ', ℓ'), $(\gamma, \ell) \succcurlyeq (\gamma', \ell')$ if and only if $V(\gamma, \ell) \geqslant V(\gamma', \ell')$. Note that the observer might not be indifferent between getting some amount x under two different identities (this may happen, for example, if she values affirmative action policies).

applied to a large set of societies. The need for an infinite set of individuals is clarified in the proof of the theorem.

As a result, there exists no *objective* natural order over the set of basic identity-outcome pairs (i, x) and, therefore, monotonicity with respect to first-order stochastic-dominance relative to outcomes cannot be assumed. Instead, we require a weaker notion of monotonicity, based on the observer's *subjective* ranking over $\mathcal{I} \times \mathcal{X}$ ($\operatorname{Im} V$ stands for the image of V)

Definition 1. *Monotonicity*: For any pair of product lotteries (γ, a) and (γ', a'),

$$\sum_{\{i:V(i,a)\leqslant v\}} \gamma\,(i) \leqslant \sum_{\{i:V(i,a')\leqslant v\}} \gamma'\,(i) \quad \text{for all } v \in \operatorname{Im} V \quad \Rightarrow \quad (\gamma, a) \succcurlyeq (\gamma', a')$$

That is, a product lottery (γ, a) is preferred over another product lottery (γ', a') (both having degenerate action lotteries), if the probability of getting identity-action pairs with utilities not greater than v is always smaller under the first product lottery. Note that monotonicity is an ordinal condition that does not depend upon the choice of the numerical representation V.

For a given society I, let $\triangle(I)$ denote the set of identity lotteries over I.

Definition 2. *Utilitarianism*: The observer is a *utilitarian* if, for every society $I \subset \mathcal{I}$, her preferences restricted to $\triangle(I) \times \triangle(\mathcal{A})$ admit a representation of the form

$$V\,(\gamma, \ell) = \sum_{i \in I} \gamma_i U_i\,(\ell_i)$$

where $\ell_i \in \triangle(\mathcal{X})$ is the lottery faced by individual i (i.e., $\ell_i(x) = \sum_{\{a \in supp(l):a(i)=x\}} \ell\,(a)$) and $U_i\,(\ell_i) := \sum_{x \in \mathcal{X}} u_i\,(x)\,\ell_i\,(x)$ is an expected utility (EU) representation of $\succcurlyeq_i$.

As is well-known, the main behavioral property that characterizes EU preferences is *independence*:

Definition 3. *Independence*: Let $\tilde{\succcurlyeq}$ be a preference relation on $\triangle(\mathcal{X})$. Then, for all

$p, q, r \in \triangle(\mathcal{X})$ and for all $\beta \in [0,1]$,

$$p \tilde{\succsim} q \quad \Rightarrow \quad \beta p + (1-\beta)\, r \tilde{\succsim} \beta q + (1-\beta)\, r$$

3.3 Assumptions

We make the following assumptions on $\succsim$:

Axiom 1. *Impartiality*: For any two individuals $i, j \in \mathcal{I}$,

$$(1) \quad \text{for all } \ell \in \triangle(\mathcal{A}), \ \succsim_i = \succsim_j \ \text{ and } \ \ell_i = \ell_j \ \Rightarrow \ (i, \ell) \sim (j, \ell)$$

$$(2) \quad (i, a_{\max}) \sim (j, a_{\max}) \ \text{ and } \ (i, a_{\min}) \sim (j, a_{\min})$$

Part (1) of this axiom states that, given an action lottery ℓ, if two individuals i and j with identical preferences are faced with the same action lottery, then the observer is indifferent between facing ℓ, while being individual i, and facing ℓ, while being individual j. This requirement seems quite natural. Part (2) says that being individual i and getting the most preferred outcome $x_{\max}$ is assumed ethically equivalent to being individual j and getting the (same) most preferred outcome $x_{\max}$. As was convincingly explained by Karni (1998) who, in a different framework, employed a stronger axiom to derive utilitarianism, "This value judgment ... is obtained by default. The methodological framework of revealed preference provides no ground for preferring one individual's most preferred alternative over that of the other. Consequently, strict preference in either direction is either biased or involves considerations other than the rank order of the alternatives". Clearly, the same applies to the worst outcome $x_{\min}$. A similar notion lies behind Segal's (2000) *dictatorship indifference* axiom.

Henceforth we assume that the observer preferences satisfy the impartiality axiom.

To emphasize it, we call her an *impartial* observer.

Axiom 2. *Strong acceptance*: For all $i \in \mathcal{I}$ and $\ell, \ell' \in \Delta(\mathcal{A})$ satisfying $\forall j \neq i \;\; \ell_j = \ell'_j$, if $\gamma_i > 0$ then

$$\ell_i \succcurlyeq_i \ell'_i \quad \Leftrightarrow \quad (\gamma, \ell) \succcurlyeq (\gamma, \ell')$$

This axiom states that the impartial observer sympathizes with individual i and fully adopts his preferences when she imagines herself being this individual with a positive probability, and when all other individuals are unaffected by her choice. This axiom strengthens Harsanyi's *acceptance* principle, according to which this sympathy holds for $\gamma_i = 1$. Axiom 2 also is analogous to an axiom called *strong Pareto*, a version of Harsanyi's *Pareto* principle that was used in his aggregation analysis (see Harsanyi (1955), Weymark (1991) and Epstein and Segal (1992)).[7] To see the connection between our axiom and the strong Pareto principle note that, by sequentially applying our axiom, the following property holds: for *any* $\ell, \ell' \in \Delta(\mathcal{A})$, if $\ell_i \succcurlyeq_i \ell'_i$ for all $i \in Supp(\gamma)$ then $(\gamma, \ell) \succcurlyeq (\gamma, \ell')$.[8] In a sense, strong acceptance unifies two of Harsanyi's main ideas, taken from his two famous analyses of social choice theory. Finally, our axiom is analogous to Karni's (1998) *sympathy* assumption.

The strong acceptance axiom enables us to express the impartial observer's function V as a *social welfare function*. That is, V can be expressed as a function W that, instead of the action lottery ℓ, depends on the individuals' utilities associated with their induced lotteries ℓ_i. More formally, let $V_i(\ell_i) := V(i, \ell)$ be a representing utility the impartial observer attaches to individual i preferences. Note that, by impartiality, $V_i(x_{\min}) = V_j(x_{\min}) := v_{\min}$ and $V_i(x_{\max}) = V_j(x_{\max}) := v_{\max}$, for all $i, j \in \mathcal{I}$, and hence by continuity, the image of V_i, for

[7] *Strong Pareto*: For a given society I, (1) for all lotteries $\ell, \ell' \in \Delta(\mathcal{A})$, if $\ell_i \succcurlyeq_i \ell'_i$ for all i, then $\ell \succcurlyeq \ell'$ and (2) if, furthermore, there exists an individual i' such that $\ell_{i'} \succ_{i'} \ell'_{i'}$, then $\ell \succ \ell'$.

[8] To see it, assume without loss of generality that $Supp(\gamma) = \{1, 2, ..., n\}$ and note that

$$(\gamma, \ell) = (\gamma, (\ell_1, \ell_2, ..., \ell_n)) \succcurlyeq (\gamma, (\ell'_1, \ell_2, ..., \ell_n)) \succcurlyeq (\gamma, (\ell'_1, \ell'_2, ..., \ell_n)) \succcurlyeq \cdots \succcurlyeq (\gamma, (\ell'_1, \ell'_2, ..., \ell'_n)) = (\gamma, \ell')$$

all i, is equal to the closed interval $[v_{\min}, v_{\max}]$. Then, strong acceptance implies that $V(\gamma, \ell)$ can be written as $W\left(\vec{\gamma}, \vec{V}(\ell)\right)$, where W is defined over $\triangle([v_{\min}, v_{\max}])$, the set of lotteries over all attainable utility values in which, for all $i \in Supp(\gamma)$, $\vec{\gamma}_i = \gamma_i$ is the probability of attaining $\left(\vec{V}(\ell)\right)_i = V_i(\ell_i)$. To see how W is constructed assume, for expositional clarity, that $Supp(\gamma) = \{1, ..., n\}$. Then, given V and V_i, for any $\gamma \in \triangle(\{1, ..., n\})$ and $\vec{v} = (v_1, ..., v_n) \in [v_{\min}, v_{\max}]^n$, define W by $W(\vec{\gamma}, \vec{v}) := V(\gamma, \ell)$, for the imaginary lottery γ satisfying $\gamma_i = \vec{\gamma}_i$ and for any ℓ satisfying $v_i = V_i(\ell_i)$, for all $i \in \{1, ..., n\}$. By strong acceptance, W is well defined. Furthermore for a given $\vec{\gamma}$, W is monotonic increasing with respect to v_i whenever $\vec{\gamma}_i > 0$. Note that, by construction, W satisfies $W(1, v) = v$, for all $v \in [v_{\min}, v_{\max}]$.

The following properties will be used in the sequel.

Lemma 1. Assume the observer satisfies *impartiality* and *strong acceptance*. Then
(a) for all $\ell, \ell' \in \triangle(\mathcal{A})$,

$$\succcurlyeq_i = \succcurlyeq_j \quad \text{and} \quad \ell_i = \ell'_j \quad \Rightarrow \quad (i, \ell) \sim (j, \ell')$$

(b) for all $(\gamma, \ell) \in \triangle(\mathcal{I}) \times \triangle(\mathcal{A})$,

$$(i, \ell) \sim (j, \ell) \quad \text{for all } i, j \in Supp(\gamma) \quad \Rightarrow \quad (\gamma, \ell) \sim (k, \ell), \quad \text{for all } k \in Supp(\gamma)$$

(c) for all $(\gamma^e, \ell), (\gamma^e, \ell') \in \triangle(\mathcal{I}) \times \triangle(\mathcal{A})$, where $Supp(\gamma^e) = \{1, ..., n\}$, if there exists a permutation π on $\{1, ..., n\}$ such that $(i, \ell_i) \sim \left(\pi(i), \ell'_{\pi(i)}\right)$ for all i, then

$$(\gamma^e, \ell) \sim (\gamma^e, \ell')$$

The proof appears in Section 3.6.1.

Axiom 3. *Convexity*: Consider an equiprobable lottery $\gamma^e \in \triangle(I)$ and two lotteries $\ell, \ell' \in \triangle(\mathcal{A})$ for which there exist two individuals $i, j \in I$ satisfying $\ell_i \succ_i \ell'_i$ and $\ell_j \prec_j \ell'_j$. Then, for all $\beta \in (0,1)$,

$$(\gamma^e, \ell) \sim (\gamma^e, \ell') \quad \Rightarrow \quad (\gamma^e, \beta\ell + (1-\beta)\,\ell') \succcurlyeq (\gamma^e, \ell)$$

As was explained in the introduction, this axiom is an expression of procedural fairness and is in agreement with Diamond's critique.

We include the requirement of having two individuals with opposing preferences since procedural fairness has greater appeal when real conflict exists. However, it is straightforward to verify that, with continuity, this requirement can be omitted. Hence, in situations where only one individual faces distinct lotteries under the action lotteries ℓ and ℓ', convexity implies that his preferences must also be convex.

Convexity is also related to social stability. Consider a society $I \subset \mathcal{I}$, whose set of available actions is given by a finite $A \subset \mathcal{A}$. For a given identity lottery $\gamma \in \triangle(I)$, the impartial observer's aim is to find the optimal action lottery that maximizes her utility. That is, the impartial observer seeks to solve the problem

$$\max_{\ell \in \triangle(A)} V(\gamma, \ell)$$

For societal stability, it is desirable that the set of optimal action lotteries does not change drastically when only minor changes occur. That is, we want this set to be upper hemi-continuous and convex valued with respect to the set of available actions A. Clearly, the continuity of $\succcurlyeq$ implies upper hemi-continuity, while convexity is equivalent to the optimal set being a convex valued correspondence.

Axiom 4. *Source indifference*: For all societies $\{i_1, ..., i_n\}$ and for all sets of available actions

$\{a^1, ..., a^n\}$, if there exists $k \in \{1, ..., n\}$ such that $\left(i_j, a^k\right) \sim (i_k, a^j)$ for all j, then

$$\left(\gamma^e, a^k\right) \sim (i_k, \ell^e)$$

where $\gamma^e = \sum_{j=1}^n \frac{1}{n} i_j$ and $\ell^e = \sum_{j=1}^n \frac{1}{n} a^j$.

To illustrate, consider the following matrix

	a^1	a^2	$\cdots$	a^k	$\cdots$	a^n
i_1				x_1		
i_2				x_2		
$\vdots$				$\vdots$		
i_k	y^1	y^2	$\cdots$	z	$\cdots$	y^n
$\vdots$				$\vdots$		
i_n				x_n		

and suppose that the impartial observer is indifferent between the following two options, for all j: (1) receiving an outcome x_j while facing the deterministic action a^k and imagining being individual i_j, and (2) receiving an outcome y^j while facing the deterministic action a^j and imagining being individual i_k. There are two ways to randomize, with equal probabilities, over these degenerate pairs of equivalent product lotteries. The product lottery $\left(\gamma^e, a^k\right)$ randomizes over identity lotteries (for the given action a^k), while product lottery (i_k, ℓ^e) randomizes over action lotteries (for the given individual i_k). Then, as was argued by Harsanyi in his response to Diamond and was implicitly assumed by him, the impartial observer should be indifferent between the two randomizations. We want to emphasize that, a priori, there is no clear reason to prefer either of these lotteries. Moreover, and as is explained in the next section (Comment 4), our utilitarian result holds even if Axiom 4 is relaxed and only requires (weak) preference of identity lotteries.

The following lemma shows that, given impartiality and strong acceptance, source indifference for equiprobability lotteries γ^e and ℓ^e implies that this property holds for all lotteries γ and ℓ^γ. This property will be used later on.

Lemma 2. Assume the observer satisfies *impartiality*, *strong acceptance* and *source indifference.* For all societies $\{i_1, ..., i_n\}$ and for all sets of available actions $\{a^1, ..., a^n\}$, if there exists $k \in \{1, ..., n\}$ such that $\left(i_j, a^k\right) \sim (i_k, a^j)$ for all j, then, for all $\gamma = \sum_{j=1}^n \gamma_{i_j} i_j$ and $\ell^\gamma = \sum_{j=1}^n \gamma_{i_j} a^j$,

$$\left(\gamma, a^k\right) \sim (i_k, \ell^\gamma)$$

The proof is relegated to Section 3.6.1.

3.4 Utilitarianism

Our main result shows that the preceding axioms force all individuals to be of the EU type and, in addition, the impartial observer must be a utilitarian. That is, the behavioral assumptions on the impartial observer preferences induce her, as well as all individuals, to satisfy the independence axiom. This is achieved without imposing independence explicitly (neither on individuals nor on the observer).

Theorem. Assume the observer satisfies *impartiality.* Then her preferences satisfy *strong acceptance*, *convexity* and *source indifference* if, and only if, all individuals in $\mathcal{I}$ satisfy *independence* and the observer is a *utilitarian.*

The proof, which is relegated to Section 3.6.2, consists of two parts**.** First, we prove that all individuals in $\mathcal{I}$ must satisfy the independence axiom. Then, we demonstrate that the impartial observer's preferences can be represented by a weighted average of the individual utilities.

Comment 1. Consider the Diamond example, represented by the table

	a^1	a^2
1	1	0
2	0	1

Having the identity lottery $\gamma^e = \left(\frac{1}{2}, \frac{1}{2}\right)$, choosing action a^i corresponds to the pair (γ^e, a^i), while tossing a fair coin corresponds to the pair $(\gamma^e, \ell^e) = \left(\gamma^e, \frac{1}{2}a^1 + \frac{1}{2}a^2\right)$. By source indifference, $(\gamma^e, a^1) \sim (1, \ell^e)$ and $(\gamma^e, a^1) \sim (2, \ell^e)$. Hence, $(1, \ell^e) \sim (2, \ell^e)$ and therefore, by Lemma 1(b), $(1, \ell^e) \sim (\gamma^e, \ell^e)$. But then, by transitivity, $(\gamma^e, a^1) \sim (\gamma^e, \ell^e)$ and the impartial observer is indifferent between the first action (Policy (1)) and the mixture (Policy (2)). Put differently, she does not strictly prefer tossing a fair coin over the pure action a^1. Moreover, it can now be seen (proof omitted) that, by convexity, *any* mixture of the two actions a^1 and a^2 must be indifferent to a^1. This may seem like a significant step towards proving utilitarianism. However, the derivation of these 'straight line indifference segments' from the above extremely symmetric situation does not extend to the general case and cannot be utilized to derive a utilitarian representation.

Comment 2. As noted in the introduction, Blackorby, Donaldson and Mongin (2004) and Mongin and Pivato (2015) also derived utilitarianism without imposing independence. Although these authors work within Harsanyi's aggregation theorem framework, a comparison to our theorem seems natural and is carried out by focusing on the analysis of Mongin and Pivato (2015). Consider a given society I, with a set of actions A, and identify every product lottery (γ, ℓ) with a matrix whose rows correspond to individuals and columns correspond to actions. Mongin and Pivato's *ex ante* analysis is manifested by their *row preference* assumption, an assumption that is analogous to our strong acceptance axiom. Similarly, their *ex post* analysis is manifested by a *column preference* assumption that, in our model, would require an improvement in the impartial observer's situation whenever an action a is replaced by a better action $\bar{a}$. Together with a *coordinate monotonicity* assumption, these two assumptions enable Mongin and Pivato to employ Gorman's (1968) separability theorem and derive a fully separable representation of the observer preferences. As can be seen in Section 3.6.2, our proof uses different arguments. Nevertheless, one might conjecture that, since source indifference implies similar treatment of columns and rows then, together with strong acceptance, Gorman's separability theorem could be applied to yield our result. However, this is not true. As can be seen in Examples 1 and 2 below, strong acceptance and source indifference are not sufficient to imply utilitarianism.

Comment 3. Another result that is close to ours appears in GKPS (2010). Their Theorem 3 roughly states that an observer is a utilitarian if and only if she satisfies acceptance, independence over identity lotteries and their notion of source indifference. We depart from their work in at least three aspects. First, since GKPS (2010) maintain the independence over identity lotteries (for the observer) while we do not assume any form of independence (neither for individuals nor the observer), the current result is stronger than theirs. It should also be noted that the notion of source indifference used by GKPS (2010) (they termed it 'indifference between identity and action lotteries') is stronger than ours. This is formally stated as Lemma 4, which can be seen in Section 3.6.4. Last but not least, our characterisation is based on the societal norm procedural fairness. This is novel and different to GKPS (2010) (and many other previous works), where the utilitarian characterisation is derived from individual behaviours.

As usual, we present the necessity of each axiom. The following first two examples demonstrate the necessity of convexity. The third demonstrates the necessity of source indifference. The fourth demonstrates the necessity of strong acceptance. It is worth mentioning that, since the generalized utilitarian impartial observer of GPKS (2010) in Example 3 satisfies independence but not source indifference, this example justifies the difference between these two axioms.

Example 1. Here we present a non utilitarian impartial observer who satisfies all axioms except for convexity. Assume that all preferences $\succcurlyeq_i$ of individuals $i \in \mathcal{I}$ belong to the rank-dependent utility class (RDU; see Weymark (1981) and Quiggin (1982)). Let $g : [0,1] \to [0,1]$ be an increasing and onto function. For a given simple lottery r and $z \in Supp\,(r)$ define $F_r\,(z) := \sum_{y \leqslant z} r\,(y)$, $F_r\,(z_-) := \sum_{y<z} r\,(y)$ and $\nabla g\,(z; r) := g\,(F_r\,(z)) - g\,(F_r\,(z_-))$. On simple lotteries, RDU preferences are represented by a function of the form $V\,(p) = \sum_x u\,(x)\,\nabla g\,(x; p)$. When g is the identity function, $\nabla g\,(x; p) = p\,(x)$ and RDU preferences are reduced to EU preferences. We assume that, in the eyes of the impartial observer, individual i's preferences are represented by $V_i\,(p) = \sum_x u_i\,(x)\,\nabla g\,(x; p)$, where g is common to all individuals and, for all $i, j \in \mathcal{I}$, $u_i\,(x_{\min}) = u_j\,(x_{\min})$ and $u_i\,(x_{\max}) = u_j\,(x_{\max})$. The observer preferences are also of the RDU type and are represented by

$$V^r\,(\gamma, \ell) = \sum_{i \in I} V_i\,(\ell_i)\,\nabla g\,(V_i\,(\ell_i)\,; \gamma)$$

Impartiality and strong acceptance are satisfied by construction. To verify that source indifference is satisfied consider, without loss of generality, a society $I = \{1, ..., n\}$, a set of available actions $\{a^1, ..., a^n\}$ and assume that there exists k for which $V^r\left(j, a^k\right) = V^r\,(k, a^j)$ for all j. Then, for all j,

$$u_j\left(a^k\,(j)\right) = V_j\left(a^k\,(j)\right) = V^r\left(j, a^k\right) = V^r\left(k, a^j\right) = V_k\left(a^j\,(k)\right) = u_k\left(a^j\,(k)\right)$$

Hence,

$$\begin{aligned} V^r\left(\gamma^e, a^k\right) &= \sum_{j\in I} u_j\left(a^k(j)\right)\nabla g\left(u_j\left(a^k(j)\right);\gamma^e\right) \\ &= \sum_{j\in I} u_k\left(a^j(k)\right)\nabla g\left(u_j\left(a^k(j)\right);\gamma^e\right) \\ &= \sum_{j\in I} u_k\left(a^j(k)\right)\nabla g\left(a^j(k);\ell^e_k\right) = V^r\left(k,\ell^e\right) \end{aligned}$$

as required.

To see that convexity does not hold assume that g is strictly concave and fix $j \in I$. Let $\ell, \ell' \in \Delta(\mathcal{A})$ be two distinct action lotteries satisfying $\ell_i = \ell'_i$ for all $i \neq j$, $\ell_j \neq \ell'_j$ and $V_j(\ell_j) = V_j\left(\ell'_j\right)$ (clearly, such lotteries exist). The strict concavity of g implies $V_j\left(\frac{1}{2}\ell_j + \frac{1}{2}\ell'_j\right) < V_j(\ell_j)$ and hence, for any γ with $\gamma_j > 0$, $V^r\left(\gamma, \frac{1}{2}\ell + \frac{1}{2}\ell'\right) < V^r(\gamma, \ell)$.[9]

Note that, as the following case shows, non-convexity of $\succcurlyeq_i$ (which is manifested by the concavity of g), is not necessary for the non-convexity of $\succcurlyeq$. For this, let $I = \{1, ..., 5\}$,

[9] Perhaps the simplest way to see it is to observe that, for continuous lotteries, $V_j(\ell_j) = \int_z u_j(z)\,\mathrm{d}g\left(F_{\ell_j}(z)\right) = u_j(x_{\max}) - \int_z g\left(F_{\ell_j}(z)\right)u'_j(z)\,\mathrm{d}z$. Therefore

$$\begin{aligned} V_j\left(\frac{1}{2}\ell_j + \frac{1}{2}\ell'_j\right) &= u_j(x_{\max}) - \int_z g\left(F_{\frac{1}{2}\ell_j+\frac{1}{2}\ell'_j}(z)\right)u'_j(z)\,\mathrm{d}z \\ &= u_j(x_{\max}) - \int_z g\left(\frac{1}{2}F_{\ell_j}(z) + \frac{1}{2}F_{\ell'_j}(z)\right)u'_j(z)\,\mathrm{d}z \\ &< u_j(x_{\max}) - \int_z \left[\frac{1}{2}g\left(F_{\ell_j}(z)\right) + \frac{1}{2}g\left(F_{\ell'_j}(z)\right)\right]u'_j(z)\,\mathrm{d}z \\ &= \frac{1}{2}\left[u_j(x_{\max}) - \int_z g\left(F_{\ell_j}(z)\right)u'_j(z)\,\mathrm{d}z\right] \\ &\quad + \frac{1}{2}\left[u_j(x_{\max}) - \int_z g\left(F_{\ell'_j}(z)\right)u'_j(z)\,\mathrm{d}z\right] \\ &= \frac{1}{2}V_j(\ell_j) + \frac{1}{2}V_j\left(\ell'_j\right) = V_j(\ell_j) \end{aligned}$$

Similarly, if g is convex then so is $\succcurlyeq_j$.

consider the two actions described by the following matrix (the entries are the utility values)

	a^1	a^2
1	1	0
2	0	1
3	1	1
4	1	1
5	1	1

let g be given by the convex piecewise linear function

$$g(t) = \begin{cases} 0 & t \leqslant 0.2 \\ -\frac{1}{4} + \frac{5}{4}t & \text{otherwise} \end{cases}$$

and note that, by the convexity of g, each $\succcurlyeq_i$ is convex.

Clearly, for both $j = 1, 2$,

$$V^r\left(\gamma^e, a^j\right) = g(0.2) \times 0 + (1 - g(0.2)) \times 1 = 1$$

Next, consider the lottery $\frac{1}{2}a^1 + \frac{1}{2}a^2$. For $i \in \{1, 2\}$,

$$V_i\left(\frac{1}{2}a^1(i) + \frac{1}{2}a^2(i)\right) = g(0.5) \times 0 + (1 - g(0.5)) \times 1 = \frac{5}{8}$$

while, for $i \in \{3, 4, 5\}$, $V_i\left(\frac{1}{2}a^1(i) + \frac{1}{2}a^2(i)\right) = 1$. Hence, for the impartial observer,

$$\begin{aligned} V^r\left(\gamma^e, \frac{1}{2}a^1 + \frac{1}{2}a^2\right) &= g(0.4) \times \frac{5}{8} + (1 - g(0.4)) \times 1 \\ &= \frac{1}{4} \times \frac{5}{8} + \frac{3}{4} \times 1 = \frac{29}{32} < 1 \end{aligned}$$

and convexity is not satisfied.

Example 2. In the two cases described in Example 1, either individual preferences are non-convex with respect to outcome lotteries (when g is concave) or the impartial observer preferences are non-convex with respect to identity lotteries (when g is convex). This might suggest that convexity would be satisfied if all preferences involved were convex. As we now show, this conjecture is false.

Assume that individual preferences are weighted utility (WU; see Chew (1983)). That is, for all i and $p \in \triangle(\mathcal{X})$,

$$V_i(p) = V(p) = \sum_k p_k \frac{w(x_k)}{\sum_j p_j w(x_j)} u(x_k)$$

where u is a strictly increasing utility function and w is a non constant and positive weighting function. These preferences belong to the betweenness class (see Chew (1989) and Dekel (1986)), a class that is characterized by the property: for all lotteries p and q, $p \succcurlyeq q$ if and only if $p \succcurlyeq \lambda p + (1-\lambda) q \succcurlyeq q$, for all $\lambda \in (0,1)$. Clearly, betweenness implies that WU preferences are convex.

The impartial observer preferences are of the same type and are given by

$$V^w(\gamma, \ell) = \sum_i \gamma_i \frac{w(u^{-1}(V(\ell_i)))}{\sum_j \gamma_j w(u^{-1}(V(\ell_j)))} V(\ell_i)$$

As in Example 1, source indifference is satisfied. To see it, assume (for $k=1$) $V^w(j, a^1(j)) = V^w(1, a^j(1))$, for all j. That is, $u(a^1(j)) = u(a^j(1))$ or, equivalently,

$a^1(j) = a^j(1)$, for all j. Then

$$\begin{aligned} V^w\left(\gamma^e, a^1\right) &= \sum_i \frac{1}{n} \frac{w\left(\left(u^{-1} \circ u\right)\left(a^1(i)\right)\right)}{\sum_j \frac{1}{n} w\left(\left(u^{-1} \circ u\right)\left(a^1(j)\right)\right)} u\left(a^1(i)\right) \\ &= \sum_i \frac{1}{n} \frac{w\left(a^1(i)\right)}{\sum_j \frac{1}{n} w\left(a^1(j)\right)} u\left(a^1(i)\right) \\ &= \sum_i \frac{1}{n} \frac{w\left(a^i(1)\right)}{\sum_j \frac{1}{n} w\left(a^j(1)\right)} u\left(a^i(1)\right) \\ &= V^w\left(1, \ell^e\right) \end{aligned}$$

Next we show that convexity is not satisfied. Consider again the Diamond's example. Assume that $u(x) = x$ and $w(x) = 2 + x$.[10] Then,

$$V^w\left(\gamma^e, a^1\right) = \frac{\frac{1}{2} w(1)}{\frac{1}{2} w(1) + \frac{1}{2} w(0)} = \frac{0.5 \times 3}{0.5 \times 3 + 0.5 \times 2} = \frac{3}{5}$$

and

$$V^w\left(\gamma^e, a^2\right) = \frac{\frac{1}{2} w(1)}{\frac{1}{2} w(0) + \frac{1}{2} w(1)} = \frac{0.5 \times 3}{0.5 \times 2 + 0.5 \times 3} = \frac{3}{5}$$

Let $\ell = 0.8a^1 + 0.2a^2$ be a mixture of a^1 and a^2. Then,

$$V(\ell_1) = \frac{0.8 w(1)}{0.8 w(1) + 0.2 w(0)} = \frac{0.8 \times 3}{0.8 \times 3 + 0.2 \times 2} = \frac{6}{7}$$

$$V(\ell_2) = \frac{0.2 w(1)}{0.8 w(0) + 0.2 w(1)} = \frac{0.2 \times 3}{0.8 \times 2 + 0.2 \times 3} = \frac{3}{11}$$

[10] WU preferences increase with respect to first-order stochastic-dominance when w and $w \cdot u$ are bounded on the outcome interval ($[0, 1]$ in this example) and when $w(x)(u(x) - u(s))$ is monotonic increasing in x for all $s \in [0, 1]$. Clearly, these conditions are satisfied.

and, for the impartial observer,

$$\begin{aligned}
V^w(\gamma^e,\ell) &= \frac{1}{2}\frac{w\left(u^{-1}\left(V\left(\ell_1\right)\right)\right)}{\frac{1}{2}w\left(u^{-1}\left(V\left(\ell_1\right)\right)\right)+\frac{1}{2}w\left(u^{-1}\left(V\left(\ell_2\right)\right)\right)}V\left(\ell_1\right)\\
&\quad+\frac{1}{2}\frac{w\left(u^{-1}\left(V\left(\ell_2\right)\right)\right)}{\frac{1}{2}w\left(u^{-1}\left(V\left(\ell_2\right)\right)\right)+\frac{1}{2}w\left(u^{-1}\left(V\left(\ell_2\right)\right)\right)}V\left(\ell_2\right)\\
&= \frac{\frac{20}{7}}{\frac{20}{7}+\frac{25}{11}}\times\frac{6}{7}+\frac{\frac{25}{11}}{\frac{20}{7}+\frac{25}{11}}\times\frac{3}{11}\\
&= \frac{44}{79}\times\frac{6}{7}+\frac{35}{79}\times\frac{3}{11}\approx 0.598<\frac{3}{5}
\end{aligned}$$

Hence, convexity is violated.

Example 3. A non utilitarian impartial observer who satisfies all axioms except for source indifference is the *generalized utilitarian* impartial observer of GKPS (2010). Consider

$$V^g(\gamma,\ell)=\sum_{i\in I}\gamma_i\phi_i\left[U_i\left(\ell_i\right)\right]$$

where $\phi_i:[v_{\min},v_{\max}]\to\mathbb{R}$ are strictly concave, for all i. It is easy to verify that strong acceptance and convexity are satisfied while, as was shown in GKPS, this observer deems identity lotteries inferior to action lotteries.

Example 4. Here we present a non utilitarian impartial observer who satisfies all axioms except strong acceptance. Consider an impartial observer whose preferences are represented by

$$V^d(\gamma,\ell)=\sum_{i\in I}\sum_{x}\gamma_i\ell_i(x)x$$

That is, the impartial observer evaluates any pair of product lottery (γ,ℓ) by its expected values. Convexity and source indifference are clearly satisfied. While, since individual preferences are always ignored by the observer, strong acceptance is not satisfied.

Comment 4. Consider the following assumption, which is weaker than source indifference.

Preference for identity lotteries: For all societies $\{i_1, ..., i_n\}$ and for all sets of available actions $\{a^1, ..., a^n\}$, if there exists $k \in \{1, ..., n\}$ such that $(i_j, a^k) \sim (i_k, a^j)$ for all j, then

$$(\gamma^e, a^k) \succcurlyeq (i_k, \ell^e)$$

In Section 3.6.3 (Lemma 3) we show that this assumption, in conjunction with strong acceptance and convexity, implies source indifference. Therefore, our theorem could be stated in a slightly stronger form. The current form is chosen because, having no reason to prefer either type of these lotteries, source indifference seems the more natural choice. Moreover, it is more in line with Harsanyi's own arguments.

3.5 Conclusion

As stated in the introduction we argue that, when societal decisions are at stake, postulates must be drawn from society centered behavior. We have chosen to focus on the notion of procedural fairness (exhibited by convexity) and added to it the requirement that the impartial observer is indifferent between identity and action lotteries. In our main result we have shown that these two assumptions (together with strong acceptance) were sufficient to force the impartial observer to be a utilitarian. Unlike most utilitarian results, no form of the independence axiom was required here.

In addition to offering a society centered basis for utilitarianism, our result sheds more light on what is needed in order to always have a strict preference for procedural fairness. Since preference for identity lotteries implies source indifference (Lemma 3, Section 3.6.3) then, in order to have a strict preference for procedural fairness, the impartial observer must display a preference for action lotteries. Two such non-utilitarian models exist in the literature. The first follows from Karni and Safra (2000).[11] In their model, which leads to the representation $V(\gamma, \ell) = \sum_{i \in I} \gamma_i V_i(\ell_i)$, individuals possess a sense of justice and preference for procedural fairness is solely manifested by their behavior (their utilities V_i are assumed to be concave). It can easily be verified that this impartial observer displays a preference for action lotteries. The second model is the generalized utilitarian impartial observer of GKPS (2010). As mentioned above, GKPS show that a preference for action lotteries holds if and only if each ϕ_i is concave, a condition that implies procedural fairness. For a third model, consider a rank dependent, or a Gini, impartial observer, whose preferences are represented by

$$V^{rd}(\gamma, \ell) = \sum_{i \in I} \phi(U_i(\ell_i)) \nabla g(U_i(\ell_i); \gamma)$$

(where each U_i is of the EU type and both ϕ and g are concave). As can easily be verified, a preference for action lotteries follows from Chew, Karni and Safra (1987) while procedural fairness follows from Quiggin (1993, Section 9.1).

[11]See also Grant, Kajii, Polak and Safra (2012b).

3.6 Appendix: Proofs

3.6.1 Proofs of Lemmata 1 and 2

Proof of Lemma 1

(a) Assume $\succcurlyeq_i = \succcurlyeq_j$ and consider $\ell, \ell' \in \triangle(\mathcal{A})$ satisfying $\ell_i = \ell'_j$. Construct an action lottery $\bar{\ell}$ that satisfies $\bar{\ell}_i = \bar{\ell}_j = \ell_i = \ell'_j$. Then

$$(i, \ell) \sim \left(i, \bar{\ell}\right) \sim \left(j, \bar{\ell}\right) \sim (j, \ell')$$

as required (the first and the last indifferences follow from strong acceptance while the second follows from impartiality).

(b) Let $v = V(i, \ell) = V_i(\ell_i)$ and note that, by the arguments that precede the statement of the lemma, $V(\gamma, \ell) = W(\vec{\gamma}, (v, ..., v))$ while $V(k, \ell) = W(1, v)$. That is, the product lottery (γ, ℓ) is equivalent to a utility lottery with n identical outcomes (where n is the number of elements in $Supp(\gamma)$), all equal to v, while (k, ℓ) is equivalent to the degenerate lottery that yields v for sure. The two utility lotteries seem identical but, in order to show that the impartial observer is indeed indifferent between them, the monotonicity property must be employed.

For this, let $c_i(\ell_i) \in \mathcal{X}$ be individual i's certainty equivalent of the lottery ℓ_i (that is, $c_i(\ell_i) \sim_i \ell_i$) and consider the action $\hat{a}$ satisfying $\hat{a}(i) = c_i(\ell_i)$. By strong acceptance, $(\gamma, \ell) \sim (\gamma, \hat{a})$ and $(k, \ell) \sim (k, \hat{a})$. Then, as the unique utility value attained by both $(\gamma, \hat{a})$ and $(k, \hat{a})$ is v, monotonicity implies that $(\gamma, \hat{a}) \sim (k, \hat{a})$. By transitivity, $(\gamma, \ell) \sim (k, \ell)$.

(c) Let (γ^e, ℓ), (γ^e, ℓ') and π satisfy the conditions of the lemma. Construct two actions $\hat{a}$ and $\hat{a}'$ satisfying $\hat{a}(i) = c_i(\ell_i)$ and $\hat{a}'(i) = c_i(\ell'_i)$ where, as above, c_i is the certainty equivalent function of individual i. By strong acceptance, $(\gamma^e, \ell) \sim (\gamma^e, \hat{a})$ and $(\gamma^e, \ell') \sim (\gamma^e, \hat{a}')$. The conditions $(i,\ell_i) \sim \left(\pi(i), \ell'_{\pi(i)}\right)$ imply $V(i,\ell_i) = V\left(\pi(i), \ell'_{\pi(i)}\right)$ for all i, and hence,

$$V(i,\hat{a}) = V(i, c_i(\ell_i)) = V(i,\ell_i) = V\left(\pi(i), \ell'_{\pi(i)}\right) = V\left(\pi(i), c_{\pi(i)}\left(\ell'_{\pi(i)}\right)\right) = V(\pi(i), \hat{a}')$$

By monotonicity, $(\gamma^e, \hat{a}) \sim (\gamma^e, \hat{a}')$ and, by transitivity, $(\gamma^e, \ell) \sim (\gamma^e, \ell')$. ■

Proof of Lemma 2 Consider, without loss of generality, a society $I = \{1, ..., n\}$, a set of available actions $A = \{a^1, ..., a^n\}$ and assume that (again, without loss of generality) $(i, a^1) \sim (1, a^i)$, for all i. Let $\gamma = (\gamma_1, ..., \gamma_n)$.

First assume that γ is rational. That is, $\gamma_i = \frac{n_i}{m_i}$, for all i. Consider a new society $\bar{I} = \{\bar{1}, \bar{2}, ...\}$ with $m_1 \cdots m_n$ individuals, in which the first $n_1 m_2 \cdots m_n$ individuals are identical to individual 1 of I, the next $m_1 n_2 m_3 \cdots m_n$ individuals are identical to individual 2 of I, and so on. Similarly, let the set of actions $\bar{A} = \{\bar{a}^1, \bar{a}^2, ...\}$ consists of $m_1 \cdots m_n$ actions, in which the first $n_1 m_2 \cdots m_n$ actions are identical to action a^1 of A, the next $m_1 n_2 m_3 \cdots m_n$ actions are identical to action a^2 of A, and so on. Finally, let $\bar{\gamma}^e$ and $\bar{\ell}^e$ be the equiprobability lotteries over $\bar{I}$ and $\bar{A}$, respectively. By construction, $(\bar{\imath}, \bar{a}^1) \sim (\bar{1}, \bar{a}^{\bar{\imath}})$, for all $\bar{\imath}$. By source indifference, $(\bar{\gamma}^e, \bar{a}^1) \sim \left(\bar{1}, \bar{\ell}^e\right)$. To conclude note that, by monotonicity, $(\gamma, a^1) \sim (\bar{\gamma}^e, \bar{a}^1)$ and, by Lemma 1(a), $\left(\bar{1}, \bar{\ell}^e\right) \sim (1, \ell^\gamma)$. Transitivity then implies $(\gamma, a^1) \sim (1, \ell^\gamma)$.

Next consider any γ and let $\beta_k \to_{k\to\infty} \gamma$ be a sequence of rational lotteries that converge to γ. By construction, $(\beta_k, a^1) \to_{k\to\infty} (\gamma, a^1)$ and $\left(1, \ell^{\beta_k}\right) \to_{k\to\infty} (1, \ell^\gamma)$. By the argument above, $(\beta_k, a^1) \sim \left(1, \ell^{\beta_k}\right)$ for all k and hence, by continuity, $(\gamma, a^1) \sim (1, \ell^\gamma)$. ■

3.6.2 Proof of the Theorem

The 'if' part is immediate. The proof of the converse is divided into two parts.

Part I[12] In this part we show that all individuals satisfy the independence axiom. Consider

[12]The proof of this part is similar to that of Dekel, Safra and Segal (1991, Theorem 2). However, dealing with social multi-person framework, our proof is more general than (and improves upon) theirs.

an individual $i^* \in \mathcal{I}$ and denote his preferences by $\succsim^*$. We want to demonstrate that for all $p, q, r \in \triangle(\mathcal{X})$, $p \sim^* q \Rightarrow \frac{1}{2}p + \frac{1}{2}r \sim^* \frac{1}{2}q + \frac{1}{2}r$. This, using Herstein and Milnor (1953), would imply that $\succsim^*$ satisfies the independence axiom. Using the continuity of $\succsim^*$, we can restrict attention to equiprobability lotteries with the same number of outcomes: $p = \left(\left(\frac{1}{k}, ..., \frac{1}{k}\right), x\right)$, $q = \left(\left(\frac{1}{k}, ..., \frac{1}{k}\right), y\right)$, and $r = \left(\left(\frac{1}{k}, ..., \frac{1}{k}\right), z\right)$ (to see it, note that (1) any lottery with rational probabilities can be replicated by an equiprobability lottery with not necessarily distinct outcomes and (2) the set of lotteries with rational probabilities is dense in the space of all lotteries).

Consider a society I consisting of $n = 2k$ individuals, all with preferences $\succsim_i = \succsim^*$. Let $\pi_1 = (1, 2, ..., n)$, $\pi_2 = (2, 3, ..., 1)$,..., $\pi_n = (n, 1, 2, ..., n-1)$ be permutations on $\{1, ..., n\}$ (where $\pi_j(i)$ stands for the ith element of the permutation π_j). We concentrate on a set of actions $\dot{A} = \{\dot{a}^1, ..., \dot{a}^n\}$ available to the society that are defined as follows: for $j = 1, ..., k$

$$\dot{a}^j(i) = \begin{cases} x_{\pi_j(i)} & \text{if } 1 \leqslant i \leqslant k \\ z_{\pi_j(i-k)} & \text{if } k < i \leqslant n \end{cases}$$

and, for $j = k+1, ..., n$

$$\dot{a}^j(i) = \begin{cases} z_{\pi_{j-k}(i)} & \text{if } 1 \leqslant i \leqslant k \\ x_{\pi_{j-k}(i-k)} & \text{if } k < i \leqslant n \end{cases}$$

To illustrate, look at the following table

	$\dot{a}^1$	$\dot{a}^2$	$\cdots$	$\dot{a}^k$	$\dot{a}^{k+1}$	$\dot{a}^{k+2}$	$\cdots$	$\dot{a}^n$
1	x_1	x_2	$\cdots$	x_k	z_1	z_2	$\cdots$	z_k
2	x_2	x_3	$\cdots$	x_1	z_2	z_3	$\cdots$	z_1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
k	x_k	x_1	$\cdots$	x_{k-1}	z_k	z_1	$\cdots$	z_{k-1}
$k+1$	z_1	z_2	$\cdots$	z_k	x_1	x_2	$\cdots$	x_k
$k+2$	z_2	z_3	$\cdots$	z_1	x_2	x_3	$\cdots$	x_1
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	z_k	z_1	$\cdots$	z_{k-1}	x_k	x_1	$\cdots$	x_{k-1}

Fact 1 $(\gamma^e, \ell^e) \sim (\gamma^e, \dot{a}^1)$.

Since for all i, j $\ell_i^e = \ell_j^e$, impartiality implies $(i, \ell^e) \sim (j, \ell^e)$ and hence, by Lemma 1(b), $(\gamma^e, \ell^e) \sim (1, \ell^e)$. Next, since $\dot{a}^j(1) = \dot{a}^1(j)$ (x_j if $j \leqslant k$ and z_{j-k} otherwise) then, in both $(1, \dot{a}^j)$ and $(j, \dot{a}^1)$, the impartial observer faces the same deterministic outcome. By Lemma 1(a), $(1, \dot{a}^j) \sim (j, \dot{a}^1)$ for all $j \in I$ and, by source indifference, $(1, \ell^e) \sim (\gamma^e, \dot{a}^1)$. Transitivity then implies $(\gamma^e, \ell^e) \sim (\gamma^e, \dot{a}^1)$.

Fact 2 Let $\ell^k = \frac{1}{k}\sum_{j=1}^{k} \dot{a}^j$. Then $\left(\gamma^e, \ell^k\right) \sim (\gamma^e, \ell^e)$.

Since all actions $\dot{a}^i$ yield the same outcomes then, using impartiality and monotonicity, $(\gamma^e, \dot{a}^i) \sim (\gamma^e, \dot{a}^1)$ for all i. By repeated application of convexity, $\left(\gamma^e, \ell^k\right) = \left(\gamma^e, \frac{1}{k}\sum_{j=1}^{k} \dot{a}^j\right) \succcurlyeq (\gamma^e, \dot{a}^1)$.[13] Hence, by Fact 1 and transitivity, $\left(\gamma^e, \ell^k\right) \succcurlyeq (\gamma^e, \ell^e)$.

For the converse, consider the action lottery $\hat{\ell}^k = \frac{1}{k}\sum_{j=k+1}^{n} \dot{a}^j$. For all $i = 1, ..., k$, $\hat{\ell}_i^k$, the lottery individual i faces under $\hat{\ell}^k$, is identical to ℓ_{k+i}^k, the lottery that individual $k+i$ faces under ℓ^k. By Lemma 1(a), $\left(i, \hat{\ell}_i^k\right) \sim \left(k+i, \ell_{k+i}^k\right)$. Similarly, $\hat{\ell}_{k+i}^k$, the lottery individual

[13] Note that by continuity, the convexity axiom holds even when there are no opposing individuals (see Section 3.3, right after the statement of the convexity axiom).

$k+i$ faces under $\hat{\ell}^k$, is identical to ℓ_i^k, the lottery that individual i faces under ℓ^k and hence, by Lemma 1(a), $\left(k+i, \hat{\ell}_{k+i}^k\right) \sim \left(i, \ell_i^k\right)$. Therefore, by Lemma 1(c), $\left(\gamma^e, \hat{\ell}^k\right) \sim \left(\gamma^e, \ell^k\right)$. Since $\ell^e = \frac{1}{2}\hat{\ell}^k + \frac{1}{2}\ell^k$, convexity implies $(\gamma^e, \ell^e) \succsim \left(\gamma^e, \ell^k\right)$.

Hence, $\left(\gamma^e, \ell^k\right) \sim (\gamma^e, \ell^e)$.

Fact 3 $\frac{1}{2}p + \frac{1}{2}r \sim^* \frac{1}{2}q + \frac{1}{2}r$.

By the first part of the proof of Fact 1, $(\gamma^e, \ell^e) \sim (1, \ell^e)$. Therefore, using transitivity and Fact 2, $\left(\gamma^e, \ell^k\right) \sim (1, \ell^e)$. Note that in the first lottery, the first k individuals face the lottery p and the rest face the lottery r while, in the second, individual 1 is faced with the lottery $\frac{1}{2}p + \frac{1}{2}r$.

Next consider the same set of individuals I with another set of actions $\tilde{A} = \left\{\tilde{a}^1, ..., \tilde{a}^{2k}\right\}$, that is derived from $\dot{A}$ by replacing every x_j by y_j. Clearly, a similar conclusion holds: the impartial observer is indifferent between the product lottery $\left(\gamma^e, \tilde{\ell}^k\right)$, in which the first k individuals face the lottery q and the rest face the lottery r, and the product lottery $\left(1, \tilde{\ell}^e\right)$, in which individual 1 is faced with the lottery $\frac{1}{2}q + \frac{1}{2}r$. But as $p \sim^* q$, all individuals in I are indifferent between p and q and hence, by strong acceptance, $\left(\gamma^e, \ell^k\right) \sim \left(\gamma^e, \tilde{\ell}^k\right)$. By transitivity, $(1, \ell^e) \sim \left(1, \tilde{\ell}^e\right)$. Hence the impartial observer, while imagining herself being individual 1, is indifferent between the lotteries $\frac{1}{2}p + \frac{1}{2}r$ and $\frac{1}{2}q + \frac{1}{2}r$. By strong acceptance, $\frac{1}{2}p + \frac{1}{2}r \sim^* \frac{1}{2}q + \frac{1}{2}r$.

To conclude Part I, note that allowing k to go to infinity implies that $\succsim^*$ satisfies independence over the entire set of lotteries $\triangle(\mathcal{X})$.[14]

Part II In the second part we show that the impartial observer is a utilitarian. Consider a society I (without loss of generality, $I = \{1, ..., n\}$) and let $V(\gamma, \ell)$ be a representation of the impartial observer preferences where $V(i, \ell) = V_i(\ell_i) = \varphi_i(U_i(\ell_i))$, φ_i is monotonic increasing and, by Part I, $U_i(\ell_i) = \sum_{x \in \mathcal{X}} u_i(x)\, \ell_i(x)$ is an EU representation of individual i's

[14]This is where we make use of the infinity of the set $\mathcal{I}$.

preferences. Since u_i is determined up to (positive) affine transformations, we can assume it satisfies $u_i(x_{\min}) = v_{\min}$ and $u_i(x_{\max}) = v_{\max}$ (hence, $\varphi_i(v_{\min}) = v_{\min}$ and $\varphi_i(v_{\max}) = v_{\max}$, for all i).

Fact 4 $\succcurlyeq$ can be represented by a separable function $\bar{V}(\gamma, \ell) = \sum_{i=1}^{n} \gamma_i \phi_i [U_i(\ell_i)]$.

Choose $(\gamma, \ell) \in \Delta(I) \times \Delta(\mathcal{A})$, denote $v_i = \varphi_i(U_i(\ell_i))$ and let $c_i(\ell_i) \in \mathcal{X}$ be individual i's certainty equivalent of the lottery ℓ_i (that is, $u_i(c_i(\ell_i)) = U_i(\ell_i)$). Consider a set of actions $\hat{A} = \{\hat{a}^j \mid j \in \{1, ..., n\}\}$ satisfying $\hat{a}^1(i) = c_i(\ell_i)$ and $\hat{a}^j(1) = (\varphi_1 \circ u_1)^{-1}(v_j)$ for $i, j = 1, ..., n$. By construction, $V(i, \hat{a}^1) = (\varphi_i \circ u_i)(c_i(\ell_i)) = v_i$ and $V(1, \hat{a}^i) = (\varphi_1 \circ u_1) \circ (\varphi_1 \circ u_1)^{-1}(v_i) = v_i$. Hence $(i, \hat{a}^1) \sim (1, \hat{a}^i)$ and, by source indifference and Lemma 2, $(\gamma, \hat{a}^1) \sim (1, \ell^\gamma)$ (ℓ^γ is the action lottery on $\hat{A}$ associated with γ). Put differently, $V(\gamma, \hat{a}^1) = V(1, \ell^\gamma)$. Note that by strong acceptance, $V(\gamma, \ell) = V(\gamma, \hat{a}^1)$. Therefore,

$$
\begin{aligned}
V(\gamma, \ell) &= V\left(\gamma, \hat{a}^1\right) = V(1, \ell^\gamma) = \varphi_1(U_1(\ell_1^\gamma)) \\
&= \varphi_1\left(\sum_{i=1}^{n} \gamma_i u_1\left((\varphi_1 \circ u_1)^{-1}(v_i)\right)\right) \\
&= \varphi_1\left(\sum_{i=1}^{n} \gamma_i \varphi_1^{-1}(v_i)\right) \\
&= \varphi_1\left(\sum_{i=1}^{n} \gamma_i \left(\varphi_1^{-1} \circ \varphi_i\right)(U_i(\ell_i))\right)
\end{aligned}
$$

Denote $\bar{V} = \varphi_1^{-1} \circ V$ and $\phi_i = \varphi_1^{-1} \circ \varphi_i$ (note that $\bar{V}$ also represents the impartial observer preferences and its image is $[v_{\min}, v_{\max}]$). By the above,

$$
\bar{V}(\gamma, \ell) = \sum_{i=1}^{n} \gamma_i \phi_i [U_i(\ell_i)]
$$

Fact 5 $\succcurlyeq$ can be represented by the affine function $\bar{V}(\gamma, \ell) = \sum_{i=1}^{n} \gamma_i U_i(\ell_i)$.

To conclude, we show that for all i, $\bar{V}_i = \phi_i \circ U_i$ is affine which, given $\varphi_i(v_{\min}) = v_{\min}$ and

$\varphi_i(v_{\max}) = v_{\max}$, implies $\bar{V}_i = U_i$. Take $\ell, \ell' \in \Delta(\mathcal{A})$. Since U_i is of the EU type, we have for all $\lambda \in [0,1]$,

$$\begin{aligned}
\bar{V}_i\left(\lambda \ell_i + (1-\lambda)\ell_i'\right) &= \phi_i\left[U_i\left(\lambda \ell_i + (1-\lambda)\ell_i'\right)\right] = \phi_i\left[\lambda U_i(\ell_i) + (1-\lambda)U_i(\ell_i')\right] \quad (3.1)\\
&= \phi_i\left[\lambda u_i(c_i(\ell_i)) + (1-\lambda)u_i(c_i(\ell_i'))\right]\\
&= \phi_i\left[U_i\left(\lambda c_i(\ell_i) + (1-\lambda)c_i(\ell_i')\right)\right]\\
&= \bar{V}_i\left(\lambda \breve{a}^i(i) + (1-\lambda)\breve{a}^j(i)\right) = \bar{V}\left(i, \lambda \breve{a}^i + (1-\lambda)\breve{a}^j\right)
\end{aligned}$$

for actions $\breve{a}^i$ and $\breve{a}^j$ satisfying $\breve{a}^i(i) = c_i(\ell_i)$, $\breve{a}^j(i) = c_i(\ell_i')$ (note that the element $\lambda c_i(\ell_i) + (1-\lambda)c_i(\ell_i')$ that appears in the third line is a lottery, not an outcome). Defining $\breve{a}^i(j) = (\phi_j \circ u_j)^{-1} \circ (\phi_i \circ u_i)(c_i(\ell_i'))$ we get

$$\bar{V}\left(j, \breve{a}^i\right) = (\phi_j \circ u_j) \circ (\phi_j \circ u_j)^{-1} \circ (\phi_i \circ u_i)(c_i(\ell_i')) = (\phi_i \circ u_i)(c_i(\ell_i')) = \bar{V}\left(i, \breve{a}^j\right)$$

and hence, by source indifference and for γ satisfying $\gamma_i = \lambda$, $\gamma_j = 1-\lambda$ and $\gamma_k = 0$ otherwise,

$$\bar{V}\left(i, \lambda \breve{a}^i + (1-\lambda)\breve{a}^j\right) = \bar{V}\left(\lambda i + (1-\lambda)j, \breve{a}^i\right)$$

(note that actions $\breve{a}^k$ for $k \neq i,j$ are irrelevant but can easily be defined so as to fit with the requirements of the axiom). Now, by the structure of $\bar{V}$ and by using the equation $\bar{V}(j, \breve{a}^i) = \bar{V}(i, \breve{a}^j)$,

$$\begin{aligned}
\bar{V}\left(\lambda i + (1-\lambda)j, \breve{a}^i\right) &= \lambda \bar{V}\left(i, \breve{a}^i\right) + (1-\lambda)\bar{V}\left(j, \breve{a}^i\right) = \lambda \bar{V}\left(i, \breve{a}^i\right) + (1-\lambda)\bar{V}\left(i, \breve{a}^j\right)\\
&= \lambda \bar{V}_i\left(\breve{a}^i(i)\right) + (1-\lambda)\bar{V}_i\left(\breve{a}^j(i)\right) = \lambda \bar{V}_i(c_i(\ell_i)) + (1-\lambda)\bar{V}_i(c_i(\ell_i'))\\
&= \lambda \bar{V}_i(\ell_i) + (1-\lambda)\bar{V}_i(\ell_i')
\end{aligned}$$

Summarizing (using equation 3.1),

$$\bar{V}_i\left(\lambda \ell_i+(1-\lambda) \ell_i^{\prime}\right)=\lambda \bar{V}_i\left(\ell_i\right)+(1-\lambda) \bar{V}_i\left(\ell_i^{\prime}\right)$$

and the affinity of $\bar{V}_i$ is established.

Hence,

$$\bar{V}(\gamma, \ell)=\sum_{i=1}^n \gamma_i U_i\left(\ell_i\right)$$

as required. ■

3.6.3 Preference for Identity Lotteries vs Source Indifference

Lemma 3 If the impartial observer preferences satisfy *strong acceptance*, *convexity* and *preference for identity lotteries* then they satisfy *source indifference.*

Proof Consider, without loss of generality, a society $I=\{1, \ldots, n\}$, a set of available actions $A=\{a^1, \ldots, a^n\}$ and assume that (again, without loss of generality) $V\left(i, a^1\right)=V\left(1, a^i\right):=v_i$, for all i. Without loss of generality we can assume that all v_i are pairwise different and that $v_i>v_{i+1}$ for all $i<n$. For $i, j \in\{1, \ldots, n\}$, let $x_{ij} \in \mathcal{X}$ be defined by $V_i\left(x_{ij}\right)=v_{\pi_j(i)}$, where π_j is a permutation on $\{1, \ldots, n\}$ (as defined in the proof of the theorem), and note that, by the monotonicity of each V_i with respect to the outcomes of $\mathcal{X}$, $V_1\left(x_{11}\right)>V_1\left(x_{12}\right)>\cdots>V_1\left(x_{1n}\right)$, $V_2\left(x_{2n}\right)>V_2\left(x_{21}\right)>V_2\left(x_{22}\right)>\cdots>V_2\left(x_{2(n-1)}\right)$,...,$V_n\left(x_{n2}\right)>V_n\left(x_{n3}\right)>\cdots>V_n\left(x_{nn}\right)>V_n\left(x_{n1}\right)$. Consider a new set of actions $\bar{A}=\{\bar{a}^1, \ldots, \bar{a}^n\}$ satisfying $\bar{a}^j(i)=x_{ij}$. By construction,

$$V\left(i, \bar{a}^1\right)=V_i\left(x_{i1}\right)=v_{\pi_1(i)}=v_i=V\left(i, a^1\right)$$

and

$$V\left(1, \bar{a}^i\right)=V_1\left(x_{1i}\right)=v_{\pi_i(1)}=v_i=V\left(1, a^i\right)$$

which implies that, by strong acceptance, $V(\gamma^e, a^1) = W(\gamma^e, (v_1, ..., v_n)) = V(\gamma^e, \bar{a}^1)$ and $V(1, \ell^e)$, given A, is equal to $V(1, \ell^e)$, given $\bar{A}$. Hence it is sufficient to restrict attention to $\bar{A}$ and to show that $V(\gamma^e, \bar{a}^1) = V(1, \ell^e)$ (given $\bar{A}$). For this note that: (i) since $V(\gamma^e, \bar{a}^i) = W(\gamma^e, (v_1, ..., v_n))$ for all i, we have $V(\gamma^e, \bar{a}^i) = V(\gamma^e, \bar{a}^j)$, for all i, j; (ii) by construction, for every $k \in \{1, ..., n\}$, $V\left(i, \bar{a}^k\right) = V(k, \bar{a}^i)$, for all i; (iii) $V(\gamma^e, \ell^e) \in [\min_i V(i, \ell^e), \max_i V(i, \ell^e)]$ and hence, if $V(\gamma^e, \ell^e) = \max_i V(i, \ell^e)$ then $V(\gamma^e, \ell^e) = V(i, \ell^e) = V(j, \ell^e)$, for all i, j; and (iv) individual i strictly prefers action $\bar{a}^{n+2-i}$ (where $\bar{a}^{n+2-1} = \bar{a}^{n+1} := \bar{a}^1$) over all other actions and, by the monotonicity of V_i with respect to first-order stochastic-dominance, he strictly prefers action $\bar{a}^i$ over all mixtures of the other actions. Therefore,

$$V\left(\gamma^e, \bar{a}^1\right) = \max_k V\left(\gamma^e, \bar{a}^k\right) \geqslant \max_k V(k, \ell^e) \geqslant V(\gamma^e, \ell^e) \geqslant V\left(\gamma^e, \bar{a}^1\right)$$

where the equality follows from (i), the first inequality follows from (ii) and from preference for identity lotteries, the second inequality follows from the first part of (iii) and the last inequality follows from (iv) by repeated application of convexity (note that $\ell^e = \frac{1}{n}\sum_j \bar{a}^j$).

Since the first and the last elements are identical, $\max_k V(k, \ell^e) = V(\gamma^e, \ell^e)$ which, by the second part of (iii), implies that $V(1, \ell^e) = \max_k V(k, \ell^e)$ and, therefore, $V(1, \ell^e) = V(\gamma^e, \bar{a}^1)$. Hence the impartial observer is indifferent between identity and action lotteries.
■

3.6.4 GKPS's (2010) Source Indifference Implies Ours

Lemma 4 Assume (as in GKPS 2010) that the impartial observer satisfies the following property:

$\forall \gamma, \gamma' \in \triangle(\mathcal{I})$, $\forall \ell, \ell' \in \triangle(\mathcal{A})$ and $\forall \beta \in (0,1)$,

$$(\gamma, \ell') \sim (\gamma', \ell) \Rightarrow (\beta\gamma + (1-\beta)\gamma', \ell) \sim (\gamma, \beta\ell + (1-\beta)\ell')$$

Then the impartial observer exhibits *source indifference.*

Proof. The proof is by induction. Without loss of generality, consider a society $I = \{1, ..., n\}$, the set of available actions $A = \{a^1, ..., a^n\}$ and assume that $(1, a^i) \sim (i, a^1)$, for all i.

First let $n = 2$. By the GKPS. condition, $(1, a^2) \sim (2, a^1)$ implies

$$(\frac{1}{2}1 + \frac{1}{2}2, a^1) \sim (1, \frac{1}{2}a^1 + \frac{1}{2}a^2)$$

as required.

Next assume it holds for $n-1$ and consider n. Assume, without loss of generality, that the acts of A satisfy $(i, a^j) \sim (i+1, a^{j-1})$ for all $i \in \{1, ..., n-1\}$, $j \in \{2, ..., n\}$. Consider the society $I^{\setminus 1} = \{2, ..., n\}$ and the set of actions $A^{\setminus n} = \{a^1, ..., a^{n-1}\}$. By construction, $(2, a^i) \sim (i+1, a^1)$ for all $i = 1, ..., n-1$ and hence, by the induction hypothesis, $(\frac{1}{n-1}\sum_{i=2}^{n} i, a^1) \sim (2, \frac{1}{n-1}\sum_{i=1}^{n-1} a^i)$. Next apply the same argument to $I^{\setminus n} = \{1, ..., n-1\}$ and $A^{\setminus n} = \{a^1, ..., a^{n-1}\}$, where $(2, a^i) \sim (i, a^2)$ for all i, to get $(2, \frac{1}{n-1}\sum_{i=1}^{n-1} a^i) \sim (\frac{1}{n-1}\sum_{i=1}^{n-1} i, a^2)$. Finally, apply it to $I^{\setminus n} = \{1, ..., n-1\}$ and $A^{\setminus 1} = \{a^2, ..., a^n\}$, where $(1, a^{i+1}) \sim (i, a^2)$ for all i, to get $(\frac{1}{n-1}\sum_{i=1}^{n-1} i, a^2) \sim (1, \frac{1}{n-1}\sum_{i=2}^{n} a^i)$. By transitivity,

$$(\frac{1}{n-1}\sum_{i=2}^{n} i, a^1) \sim (1, \frac{1}{n-1}\sum_{i=2}^{n} a^i)$$

To conclude, mix both sides of the last indifference with $(1, a^1)$ and, by the GKPS. condition, obtain $(\alpha^e, a^1) \sim (1, \ell^e)$ for $I = \{1, ..., n\}$ and $A = \{a^1, ..., a^n\}$, as required. ■

Chapter 4

Social Justice and Risk Perceptions: An Analysis of Contributions to Cancer Research

Abstract: The last few decades have witnessed an increasing proportion of all cancers are lifestyle-related, a trend tends to continue. To assess policies used to address this challenge and to construct a more equitable healthcare system, it is important to gain a better understanding of how individual responsibility is viewed by individual society members. We have constructed a procedure to assess these preferences with respect to lifestyle-related and hereditary cancers by eliciting donations to these two types of cancers. Lifestyle-related cancers involve greater individual responsibility than hereditary cancers, as individuals can control more elements of lifestyle than heredity. The results of implementing our procedure via an online survey demonstrate that subjects take individual responsibility into account by donating about twice as much to cancers with less individual responsibility. Their choices are also affected by group identity, perceptions of cancer likelihood and social demographics.

"Do heavy smokers who contract lung cancer have the same claim, on equity grounds, to resources them to full health (so far as that might be possible) as nonsmokers who contract the diseases?" – Le Grand 1987, page 263-264

4.1 Introduction

As a result of decades of scientific research, it is increasingly clear that personal choices - unhealthy lifestyles - contribute significantly to the chances of contracting diseases (World Health Organization, 2002). In a liberal society, where individual lifestyles are freely chosen, the ethics of healthcare policy[1] consist in part in determining the fair distribution for each: how differences in individual choices should be reflected? While the moral principle that people should be held responsible for their personal choice to remain healthy is widely accepted in Western societies, interpretations and applications of this principle are largely at issue (Wikler, 2002; Brownell et al., 2010; Greenfield, 2011). For instance, the classic issue of whether alcoholics should be excluded from liver transplantation has attracted both controversial debates (Cohen and Benjamin, 1991; Moss and Siegler, 1991; Sherman and Williams, 1995) and policies (The Guardian, 2014). To compromise controversy, a democratic solution is to reflect and understand individual society members' attitudes to individual responsibility.

This essay contributes to this direction, theoretically and empirically, by scrutinizing public preferences in distributing donations between research on lifestyle-related and hereditary cancers. We consider, for the first time, how such preferences may be affected by individual risk perceptions.

Variety of people's perceptions, although often assumed away in economics to avoid complexity in analysis, is itself an inseparable part of decision making (Simon, 1955).[2] In

[1]To clarify the scope of this paper, note that the term "healthcare" is different from "health". The former is one of causal factors for the latter.

[2]The problems studied in the risk analysis literature on how people's risk perceptions vary across groups

this study, we restrict our attention to two aspects of risk perceptions that matter to distribution choices. First, people's perceptions of cancer risks may differ. For instance, physical inactivity is perceived to be a crucial cancer risk for some, but not for others. Similarly, some may perceive cancers as more likely to be linked with lifestyle than others. To capture these differences, we have developed a novel procedure that incorporates people's estimates of various likelihoods linked with cancers.

Second, people's perceptions of cancer risks may be ambiguous. One information source that largely shapes such perceptions is scientific evidence and its associated media coverage. However, this evidence is inconclusive. A recent example that attracted public attention was around prevention rates (i.e., how many cancers are preventable and involve individual responsibility). In 2015, Tomasetti and Vogelstein (2015) estimated that the majority of cancers are due to "bad luck" and are completely out of people's control. However, only a few months later, Wu et al. (2016) provided contradictory evidence, suggesting 70% to 90% of cancers are due to avoidable, extrinsic risk factors. Such inconclusiveness may prevent people from forming unambiguous risk perceptions. Therefore, to capture the potential ambiguity of cancer risks, we use interval probabilities (Flage et al., 2014) allowing subjects to report both intervals of estimates and single estimates.

In constructing our theoretical models and designing our surveys, we made two key decisions that merit discussion. First, while the literature on fairness preferences normally ignore ambiguity, we bring the ambiguous risk perceptions to the forefront of analysing fairness preferences. The significance of ambiguity in decision making is no longer a secret in economics, and there seems no reason that our decisions involving fairness considerations can be an exception. In a related but different study, Dana, Weber and Kuang (2007) has showed that ambiguity in the distributional situations (i.e., what is the likelihood that the

are different to the problem we consider here. However, the finding that risk perceptions differ, for example by status (e.g., Slovic, 1987) and gender (e.g., Flynn, Slovic and Mertz, 1994; Gustafsod, 1998), provides important insights into how to think about the variety of risk perceptions.

unfairness is caused by your decisions) may indeed influence fairness preferences. We study the effect of ambiguity in risk perceptions here. The distributive decisions in our settings can be influenced by both amount of ambiguity and fairness preferences, but this is not a confound in the design. Rather, by explicitly controlling and separating the effect of ambiguity in risk perceptions that implicitly lies in the distribution problems, our design allows a clear pattern of fairness preferences.

Second, while our focus in this chapter and next chapter is the fairness principles on individual responsibility, another interesting dimension worth attention is the ex ante and ex post nature of these problems. As will become clear, our naturally-occurring data in the next chapter does not allow to differentiate between the ex ante and ex post principles. For this reason, rather than modelling the trade-off between ex ante and ex post principle, our purpose of considering those two principles in the theoretical parts is to show that the testable predictions of the models are robust to these two thoughts of the fairness principle. Interesting questions like how individuals invest between prevention and treatment are left for future research.

In the empirical section of this paper, we also record a couple of perceived group identities, including highly publicized cancer cases. We are particularly interested in how Angelina Jolie's proclaimed medical decision to prevent hereditary cancers (The New York Times, 2013) may influence people's donating choices. In the theoretical section, we develop a novel and testable model of to explain how distributions of donations between research on hereditary and lifestyle-related cancers may be affected by potential beneficiaries' individual responsibility. Two types of donors, choice egalitarian and non-choice egalitarian, are modelled, and the former are distinguished from the latter by focusing on individual responsibility in contracting diseases. We have two main theoretical predictions. First, a society consisting of both types of donor will allocate more funds to hereditary cancers. Second, an increase in donors' ambiguity about individual responsibility will result in a decrease in donations to lifestyle-related cancers.

Our empirical data, collected through an online survey based on a representative sample, support both theoretical predictions. The subjects allocated twice as much to alternatives with less individual responsibility than to those with more individual responsibility, and their ambiguity in individual responsibility was negatively related to donations to lifestyle-related cancers. Apart from these two results, we also find a self-interest incentive exhibited in various dimensions.

The previous literature regarding justice in health care distributions can be divided into two main categories. One makes specific proposal directly (see, e.g., Le Grand, 1987; Culyer and Wagstaff, 1993; Cappelen and Norheim, 2005; Brownell et al., 2010). The other, to which this study contributes, reveals public's preferences on distribution fairness, and leaves the final decision on how to reflect these preferences in health care to the political level (Yaari and Bar-Hiller, 1984; Kahneman, Knetsch and Thaler, 1986; Konow, 2000; Cappelen et al., 2007; Cappelen, Sørensen and Tungodden, 2010; Cappelen et al., 2013; Cettolin and Riedl, 2016; Melkonyan, Safra and Ma, 2016). These papers have greatly improved our knowledge of distributive justice. However, the present study differs from all of them by considering differences in risk perceptions and ambiguity.

Our paper also contributes to the literature on conditional altruism (see, e.g., Rabin, 1993; Brekke, Kverndokk and Nyborg, 2003; Konow, 2010). This branch of literature argues that people's altruism is conditional on compliance with moral norms. Since being responsible for one's own choices may be interpreted as one such moral norm, our results resonate well with this branch of literature.

The remainder of the paper is organized as follows. The theoretical model and its predictions are described in Section 2. The empirical model and results are presented in Section 3. The last section draws conclusions.

4.2 Theoretical Model

Consider a donor who is endowed with a sum of £100 to donate to cancer research. The sum can be allocated either to a (typical) hereditary cancer or to a (typical) lifestyle-related cancer. The donor's utility depends on the utilities of potential beneficiaries, who are cancer patients benefitting from improved treatment and a better prevention policy. In what follows, we first consider the case where the donor perceives risk about the adverse outcomes associated with having a cancer. For the case of risk, we begin with characterizing the optimal donations to cancer treatment and then turn to the optimal donations to cancer prevention. After the analysis of choices under risk, we examine the case of ambiguity and, similarly to the risk scenario, we first analyze the cancer treatment and then the cancer prevention.

4.2.1 Donations under Risk

4.2.1.1 Donations to Cancer Treatment

Suppose the donor allocates £100 between two projects with each affecting a distinct cancer patient, one suffering from a hereditary cancer, denoted by hr, and the other suffering from a lifestyle-related cancer, denoted by ls. We will call the patients suffering from cancers hr and ls as patients hr and ls, respectively. Let $d \in [0, 100]$ denote the amount allocated to cancer hr (therefore, $100 - d$ is allocated to cancer ls). The donor assumes that her donation positively affects the effectiveness of cancer treatment. That is, the donation goes toward cancer research which increases the corresponding patients' probability of survival, also called survival rate. The survival rates for the two cancers are assumed, for simplicity, to be equal to each other and given by

$$p_s(d) = p_s^0 + \left(1 - p_s^0\right)\beta(d), \tag{4.1}$$

where $p_s^0 \in (0,1)$ denotes the current survival rate, which is also assumed to be identical across the two cancers, and $\beta : \mathbb{R}_+ \to [0,1)$ is an onto, strictly increasing and strictly concave function with $\beta(0) = 0$.[3] The survival rate function represents the donor's beliefs that (1) her donation has a positive effect on the survival rate and (2) the size of the effect is negatively related to the current survival rate. Note that all values of $p_s(d)$ fall between the current survival rate and 1.

It is assumed that surviving a cancer, an outcome denoted by H, gives both patients, in the eyes of the donor, a utility of one unit; $u^{hr}(H) = u^{ls}(H) = 1$. Dying from a cancer, an outcome denoted by M, yields a lower utility. The utility of patient hr under contingency M, as perceived by the donor, is assumed to be zero; $u^{hr}(M) = 0$. The utility attached by the donor to patient ls, on the other hand, depends on whether or not the donor "penalizes" patient ls for his cancer-inducing behavior. To model this possibility, let $p_a \in (0,1)$ denote the probability that cancer ls is induced by a specific, avoidable, lifestyle (by definition, the corresponding probability for cancer hr is equal to zero). A *choice egalitarian* donor realizes that unhealthy lifestyle is a decision made by supposedly informed individuals and presumes that patients following such lifestyle care less about their possible death. Thus, this donor assumes that the utility difference for patient ls between outcomes H and M is smaller than that of patient hr. More specifically, the choice egalitarian donor assigns $u^{ls}(M) = p_a$. Hence, the utility difference between the two health outcomes is $u^{ls}(H) - u^{ls}(M) = 1 - p_a$, which is decreasing with the probability that cancer ls is lifestyle-related. A donor who is not choice egalitarian uses $u^{ls}(M) = 0$ in her evaluation. Thus, by assigning the same utility

[3]It is assumed that $\beta(\cdot)$ and function $\gamma(\cdot)$ of the next section satisfy standard Inada (1963) conditions that warrant interior solutions.

to patients hr and ls under both contingencies this donor doesn't differentiate between the two types of patients.

The non-choice egalitarian donor assigns the (ex ante) expected utilities EU_n^{hr} and EU_n^{ls} to patients hr and ls, respectively. It follows from our assumptions that

$$\begin{aligned} EU_n^{hr} &= p_s(d)\, u^{hr}(H) + (1 - p_s(d))\, u^{hr}(M) = p_s(d), \\ EU_n^{ls} &= p_s(100-d)\, u^{ls}(H) + (1 - p_s(100-d))\, u^{ls}(M) = p_s(100-d). \end{aligned} \tag{4.2}$$

The non-choice egalitarian donor's objective function $V_n(d)$ is given by the sum of the (ex ante) expected utilities EU_n^{hr} and EU_n^{ls}, which using (4.2) can be written as:

$$V_n(d) = EU_n^{hr} + EU_n^{ls} = p_s(d) + p_s(100-d).$$

The non-choice egalitarian donor's optimal choice of donation is given by the first-order condition[4] of her optimization problem:

$$\begin{aligned} \frac{\partial}{\partial d} V_n(d) &= p_s'(d) - p_s'(100-d) \\ &= \left(1 - p_s^0\right)\left[\beta'(d) - \beta'(100-d)\right] = 0. \end{aligned} \tag{4.3}$$

The choice egalitarian donor assigns the (ex ante) expected utilities EU_{ch}^{hr} and EU_{ch}^{ls}

[4]The first-order condition is sufficient by strict concavity of the objective function:

$$\frac{\partial^2}{\partial d^2} V_n(d) = \left(1 - p_s^0\right)\left[\beta''(d) + \beta''(100-d)\right] < 0.$$

Similarly, all of the other objective functions considered in the paper are strictly concave.

to patients hr and ls, respectively. It follows from our assumptions that

$$\begin{aligned} EU_{ch}^{hr} &= p_s(d)\,u^{hr}(H) + (1 - p_s(d))\,u^{hr}(M) = p_s(d) = EU_n^{hr}, \qquad (4.4) \\ EU_{ch}^{ls} &= p_s(100-d)\,u^{ls}(H) + (1 - p_s(100-d))\,u^{ls}(M) \\ &= p_s(100-d) + p_a(1 - p_s(100-d)) > EU_n^{ls}. \end{aligned}$$

The choice egalitarian donor's objective function $V_{ch}(d)$ is given by the sum of the (ex ante) expected utilities EU_{ch}^{hr} and EU_{ch}^{ls}, which using (4.4) can be written as:

$$V_{ch}(d) = EU_{ch}^{hr} + EU_{ch}^{ls} = p_s(d) + p_s(100-d) + p_a(1 - p_s(100-d)).$$

It then follows that her optimal choice of donation is given by

$$\begin{aligned} \frac{\partial}{\partial d} V_{ch}(d) &= p_s'(d) - (1 - p_a)\,p_s'(100-d) \qquad (4.5) \\ &= \left(1 - p_s^0\right)\left[\beta'(d) - (1 - p_a)\,\beta'(100-d)\right] = 0. \end{aligned}$$

A comparison of the donation choices of the two donor types in (4.3) and (4.5) leads to (the proofs of all propositions are in Appendix A):

Proposition 1 *Consider a society that consists of both choice egalitarian and non-choice egalitarian donors. Suppose that each donor is given £100 to allocate between the hereditary and lifestyle-related cancers, where both allocations are for research that improves treatment of the disease. Then the average allocation to the hereditary cancer is larger than that for the lifestyle-related cancer.*

Thus, unless all of the donors are non-choice egalitarian, the hereditary cancer will receive a larger donation. It also follows immediately from the derivations leading to Proposition 1 that the larger the share of choice egalitarian donors the larger the total contribution to the hereditary cancer.

4.2.1.2 Donations to Cancer Prevention

Similarly to the analysis for the cancer treatment case, we assume that the donor considers two cancer patients, hr and ls, and that $d \in [0, 100]$ denotes the amount allocated to cancer hr. Here, the donor focuses on contributing to research to prevent cancer and assumes that her donation positively affects cancer prevention by decreasing the potential patients' probability of getting the cancer. Let $p_c^{hr}(d)$ and $p_c^{ls}(d)$ denote the resulting cancer incidence rates, which for brevity are called cancer rates, for cancers hr and ls, respectively. The cancer rates for the two cancer types are assumed to be equal to each other and given by

$$p_c(d) = p_c^0 \gamma(d),$$

where $p_c^0 \in (0, 1)$ denotes the current cancer rate, which is also assumed to be identical across the two cancers, and $\gamma : \mathbb{R}_+ \to (0, 1]$ is an onto, strictly decreasing and strictly convex function (note that $\gamma(0) = 1$). Similarly to the preceding subsection, the cancer rate function represents the donor's beliefs that (1) her donation can affect cancer rates and (2) the size of the effect is negatively related to the current cancer rate. All values of $p_c(d)$ are between the current cancer rate and 0. The survival rates for both cancers are fixed and equal to $p_s \in (0, 1)$.

It is assumed that either not getting a cancer, or surviving it and being healthy, is represented by the same outcome H for both potential patients. Hence, in the eyes of the donor, $u^{hr}(H) = u^{ls}(H) = 1$.[5] Dying from a cancer, an outcome denoted by M, yields the same utilities as in the previous section: $u^{hr}(M) = 0$ and $u^{ls}(M) = p_a$ for the choice

[5]For simplicity, it is assumed that the outcome when the individual doesn't get the cancer yields her the same utility as the outcome when she gets the cancer and then gets cured. All of our results hold under the assumption when the latter outcome yields a lower utility.

egalitarian donor while $u^{ls}(M) = 0$ for the non-choice egalitarian donor.

The non-choice egalitarian donor assigns the (ex ante) expected utilities EU_n^{hr} and EU_n^{ls} to patients hr and ls, respectively.[6] We have that

$$\begin{aligned} EU_n^{hr} &= (1 - p_c(d)) u^{hr}(H) + p_c(d) \left(p_s u^{hr}(H) + (1 - p_s) u^{hr}(M)\right) \\ &= 1 - (1 - p_s) p_c(d), \\ EU_n^{ls} &= (1 - p_c(100 - d)) u^{ls}(H) + p_c(100 - d) \left(p_s u^{ls}(H) + (1 - p_s) u^{ls}(M)\right) \\ &= 1 - (1 - p_s) p_c(100 - d). \end{aligned} \tag{4.6}$$

The non-choice egalitarian donor's objective function $V_n(d)$ is given by the sum of the (ex ante) expected utilities EU_n^{hr} and EU_n^{ls}, which using (4.6) can be written as:

$$V_n(d) = EU_n^{hr} + EU_n^{ls} = 2 - (1 - p_s) [p_c(d) + p_c(100 - d)].$$

The non-choice egalitarian donor's optimal choice of donation is given by

$$\begin{aligned} \frac{\partial}{\partial d} V_n(d) &= -(1 - p_s) [p_c'(d) - p_c'(100 - d)] \\ &= -(1 - p_s) p_c^0 [\gamma'(d) - \gamma'(100 - d)] = 0. \end{aligned} \tag{4.7}$$

The choice egalitarian donor assigns the (ex ante) expected utilities EU_{ch}^{hr} and EU_{ch}^{ls}

[6] Note that we slightly abuse the notation and, similarly to Section 4.2.1.1, use EU_n^{hr} and EU_n^{ls} to denote the (ex ante) expected utilities. The same comment applies to the notation for the choice and non-choice egalitarian donors' objective functions, as well as the corresponding constructs in the analysis of donations under ambiguity.

to patients hr and ls, respectively. It follows from our assumptions that

$$\begin{aligned} EU_{ch}^{hr} &= (1 - p_c(d))\, u^{hr}(H) + p_c(d)\left(p_s u^{hr}(H) + (1 - p_s)\, u^{hr}(M)\right) \\ &= 1 - (1 - p_s)\, p_c(d) = EU_n^{hr}, \\ EU_{ch}^{ls} &= (1 - p_c(100 - d))\, u^{ls}(H) + p_c(100 - d)\left(p_s u^{ls}(H) + (1 - p_s)\, u^{ls}(M)\right) \\ &= 1 - (1 - p_s)(1 - p_a)\, p_c(100 - d) > EU_n^{ls}. \end{aligned} \tag{4.8}$$

The choice egalitarian donor's objective function $V_{ch}(d)$ is given by the sum of the (ex ante) expected utilities EU_{ch}^{hr} and EU_{ch}^{ls}, which using (4.8) can be written as:

$$V_{ch}(d) = EU_{ch}^{hr} + EU_{ch}^{ls} = 2 - (1 - p_s)\left[p_c(d) + (1 - p_a)\, p_c(100 - d)\right].$$

Her optimal donation choice is given by

$$\begin{aligned} \frac{\partial}{\partial d} V_{ch}(d) &= -(1 - p_s)\left[p_c'(d) - (1 - p_a)\, p_c'(100 - d)\right] \\ &= -(1 - p_s)\, p_c^0\left[\gamma'(d) - (1 - p_a)\, \gamma'(100 - d)\right] = 0 \end{aligned} \tag{4.9}$$

Using (4.7) and (4.9) we obtain:

Proposition 2 *Consider a society that consists of both choice egalitarian and non-choice egalitarian donors. Suppose that each donor is given £100 to allocate between the hereditary and lifestyle-related cancers, where both allocations are for research that improves prevention of the disease. Then the average allocation to the hereditary cancer is larger than that for the lifestyle-related cancer.*

Propositions 1 and 2 also apply to the aggregate (across all potential donors) allocations. That is, the aggregate allocation to the hereditary cancer under both scenarios (treatment and prevention) exceeds the aggregate allocation to the lifestyle-related cancer.

4.2.2 Donations under Ambiguity

4.2.2.1 Donations to Cancer Treatment

Similarly to Section 4.2.1.1, the survival rate for both cancer types is given by $p_s(d) = p_s^0 + (1 - p_s^0)\beta(d)$, where p_s^0 has some unique value in the interval $(0,1)$. In contrast to Section 4.2.1.1 where p_a also had a unique value, the decision-maker's beliefs about the likelihood that cancer ls is induced by a specific, avoidable, lifestyle are given by the *probability interval* $[\underline{p}_a, \bar{p}_a]$, where $0 \leq \underline{p}_a \leq \bar{p}_a \leq 1$. $\underline{p}_a$ and $\bar{p}_a$ are called minimal and maximal probabilities, respectively. As before, the corresponding probability for cancer hr is assumed to be equal to zero.

The elements of $[\underline{p}_a, \bar{p}_a]$ can be thought of as representing at least two factors: the decision-maker's information on the possible probabilities and his or her degree of confidence in the existing theories surrounding these probabilities.[7] So, for example, if there are several competing hypotheses about the likelihood that cancer ls is induced by a specific lifestyle, but the donor is convinced that only one is truly valid, then $[\underline{p}_a, \bar{p}_a]$ would be a singleton. In this case, the decision-maker faces pure risk. In contrast, complete ambiguity or pure uncertainty is characterized by $[\underline{p}_a, \bar{p}_a] = [0,1]$. This represents situation where the donor has no information on the current likelihood that cancer ls is induced by a specific lifestyle other than that it falls somewhere in $[0,1]$. We will call the difference between the maximal and minimal probabilities a *degree of ambiguity* and denote it by $\omega \equiv \bar{p}_a - \underline{p}_a$.

[7] Gajdos et al. (2004) provide a complementary interpretation of the set of probabilities. The decision-maker in their model maximizes the minimum expected utility computed with respect to a subset of the set of initially given priors. The extent to which the set of initially given priors is reduced is a measure of aversion to information imprecision.

The donor has $\alpha-$MMEU preferences

$$W_i(d) = \alpha \min_{p_a \in \left[\underline{p_a}, \bar{p}_a\right]} \left\{EU_i^{hr} + EU_i^{ls}\left(p_a\right)\right\} \tag{4.10}$$
$$+ (1-\alpha) \max_{p_a \in \left[\underline{p_a}, \bar{p}_a\right]} \left\{EU_i^{hr} + EU_i^{ls}\left(p_a\right)\right\},$$

where $\alpha \in [0,1]$ is the parameter characterizing the donor's attitude to ambiguity, $i \in \{n, ch\}$ reflects whether the donor is a choice egalitarian or a non-choice egalitarian, and EU_i^{hr} and $EU_i^{ls}\left(p_a\right)$ are given by (4.2) for $i = n$ and (4.4) for $i = ch$. $\alpha-$MMEU preference structure is a generalization of Arrow-Hurwicz criterion (Hurwicz, 1952, Arrow and Hurwicz, 1972) and it reduces to an expected utility preference functional when $\left[\underline{p_a}, \bar{p}_a\right]$ is a singleton.

$\alpha-$MMEU preference structure is also a natural generalization of the maximin and maximax decision rules (Gilboa and Schmeidler, 1989). When $\alpha = 1$, $\alpha-$MMEU preferences have *maximin expected utility* form (Gilboa and Schmeidler, 1989) which corresponds to *complete ambiguity aversion* (also called *complete pessimism*) on the donor's part. An MMEU donor is completely pessimistic in the sense that, when evaluating stochastic outcomes, he or she always uses the probability distribution that yields the lowest possible expected utility over $\left[\underline{p_a}, \bar{p}_a\right]$. In our case, an MMEU donor will use the smallest possible probability $\underline{p_a}$ that cancer ls is induced by lifestyle.

In contrast, $\alpha = 0$ corresponds to a donor with *maximax expected utility* preferences and reflects a situation where the donor is *completely ambiguity tolerant*. A donor with $\alpha = 0$ focuses all attention on the *most optimistic* probability distribution, which in our case is equal to $\bar{p}_a$.

Given that $\alpha-$MMEU utility is a weighted linear functional of the most pessimistic and most optimistic scenarios, it is natural to call α a *measure of ambiguity aversion* or a *coefficient of pessimism*. A donor with a larger α is said to be more ambiguity averse (or more pessimistic).[8] A donor with $\alpha > 0.5$ is said to be ambiguity averse while a donor with $\alpha < 0.5$ is said to be ambiguity loving. Note also that a donor with $\alpha = 0.5$ is not ambiguity neutral. In the present model, a donor is ambiguity neutral if and only if the set of probabilities $\left[\underline{p_a}, \bar{p}_a\right]$ is a singleton.

The utilities assigned by the choice egalitarian and non-choice egalitarian donors to the two cancer patients are identical to the corresponding utilities in Section 4.2.1.1. Under this assumption and using (4.1), (4.2) and (4.10), we can write the non-choice egalitarian donor's objective function as

$$W_n(d) = p_s(d) + p_s(100 - d),$$

while the choice egalitarian donor's objective as

$$W_{ch}(d) = p_s(d) + p_s(100 - d) + \left(\alpha \underline{p_a} + (1 - \alpha)\bar{p}_a\right)(1 - p_s(100 - d)).$$

The latter expression reveals that the choice-egalitarian donor with $\alpha-$MMEU preference and beliefs $\left[\underline{p_a}, \bar{p}_a\right]$ behaves similarly to the choice-egalitarian donor with expected utility preferences and beliefs given by the singleton probability $\left(\alpha \underline{p_a} + (1 - \alpha)\bar{p}_a\right)$. It also follows immediately from the objective functions above that the non-choice egalitarian donor's optimal donation is unchanged from Section 4.2.1.1 and is given by (4.3). In contrast, the condition characterizing the choice egalitarian donor's optimal donation changes from (4.5) to:

$$\left(1 - p_s^0\right)\left[\beta'(d) - \left(1 - \left(\alpha \underline{p_a} + (1 - \alpha)\bar{p}_a\right)\right)\beta'(100 - d)\right] = 0. \tag{4.11}$$

Using the expression for $W_n(d)$ and (4.11), we obtain

[8] Note that "more ambiguity averse" is a comparative rather than absolute notion. Thus, a more ambiguity averse donor may very well be ambiguity loving.

Proposition 3 *If the donor is non-choice egalitarian then her donation to the treatment of the lifestyle-related cancer will be unaffected by her ambiguity aversion and her beliefs* $\left[\underline{p_a}, \bar{p}_a\right]$ *that cancer ls is lifestyle-related. Conversely, if the donor is choice egalitarian then*

(a) *a more ambiguity averse donor will donate more to the treatment of the lifestyle-related cancer;*

(b) *an increase in the minimal likelihood of the lifestyle-related cancer, holding its degree of ambiguity constant,* $d\underline{p_a} = d\bar{p}_a > 0$ *will result in a decrease in the donation to the treatment of the lifestyle-related cancer.*

(c) *an increase in the ambiguity of the lifestyle-related cancer, holding its minimal likelihood constant,* $d\bar{p}_a > 0 = d\underline{p_a}$ *will result in a decrease in the donation to the treatment of the lifestyle-related cancer.*

The intuition behind Proposition 3 can be grasped immediately by examining the expressions for the two donor types' objective functions. While the attitude and perception of ambiguity don't affect the non-choice egalitarian donor, these behavioral traits enter the choice egalitarian donor's objective $W_{ch}(d)$ as the sum $\left(\alpha\underline{p_a} + (1-\alpha)\,\bar{p}_a\right)$ of the minimal and maximal probabilities weighted by the respective ambiguity attitudes. This term is decreasing in the degree of ambiguity aversion and increasing in $\underline{p_a}$ and $\bar{p}_a$. Recall also that the larger the magnitude of $\left(\alpha\underline{p_a} + (1-\alpha)\,\bar{p}_a\right)$ the larger the utility assigned by the choice egalitarian to the adverse outcome $u^{ls}(M)$ associated with the lifestyle-related cancer. As a result, donations to the lifestyle related cancer become relatively unattractive as $\left(\alpha\underline{p_a} + (1-\alpha)\,\bar{p}_a\right)$ increases.

4.2.2.2 Donations to Cancer Treatment

Similarly to Section 4.2.1.2, the cancer rate for both cancer types is given by $p_c(d) = p_c^0\gamma(d)$. The utilities assigned by the choice egalitarian and non-choice egalitarian donors to the two cancer patients are also identical to the corresponding utilities in Section 4.2.1.2. The beliefs about p_a and attitudes to ambiguity are the same as in the preceding section. Also similarly to Section 4.2.2.1, the donor has $\alpha-$maximin expected utility preferences.

Under these assumptions, we can write the non-choice egalitarian donor's objective function as

$$W_n(d) = 2 - (1 - p_s)[p_c(d) + p_c(100 - d)].$$

while the choice egalitarian donor's objective as

$$W_{ch}(d) = 2 - (1 - p_s)\left[p_c(d) + \left(1 - \left(\alpha \underline{p_a} + (1 - \alpha)\bar{p}_a\right)\right)p_c(100 - d)\right].$$

The choice egalitarian donor's optimal choice is given by

$$-(1 - p_s)p_c^0\left[\gamma'(d) - \left(1 - \left(\alpha \underline{p_a} + (1 - \alpha)\bar{p}_a\right)\right)\gamma'(100 - d)\right] = 0 \qquad (4.12)$$

Using the expression for $W_n(d)$ and (4.12) we obtain:

Proposition 4 *If the donor is non-choice egalitarian then her donation to the prevention of the lifestyle-related cancer will be unaffected by her ambiguity aversion and her beliefs* $\left[\underline{p_a}, \bar{p}_a\right]$ *that cancer ls is lifestyle-related. Conversely, if the donor is choice egalitarian then*

(a) *a more ambiguity averse donor will donate more to the prevention of the lifestyle-related cancer;*

(b) *an increase in the minimal likelihood of the lifestyle-related cancer, holding its degree of ambiguity constant,* $d\underline{p_a} = d\bar{p}_a > 0$ *will result in a decrease in the donation to the prevention of the lifestyle-related cancer.*

(c) *an increase in the ambiguity of the lifestyle-related cancer, holding its minimal likelihood constant,* $d\bar{p}_a > 0 = d\underline{p_a}$ *will result in a decrease in the donation to the prevention of the lifestyle-related cancer.*

4.3 Survey Data

The study was conducted from November 24-28, 2014 using the online survey platform Maximiles. A total of 166 valid responses were obtained out of a group of 203 respondents recruited from the UK general population.[9]

To elicit how charitable giving is affected by the individual responsibility of the beneficiaries of hypothetical donations, the respondents faced a series of hypothetical choice questions (see Appendix B). For two of the choice questions, the *prevention* and *treatment* questions, the respondents were asked to allocate £100 between two alternatives. For the first of these questions, the choice pertained to the allocation between research on the prevention of hereditary versus lifestyle-related cancers, while for the second question, it was between research on the treatment of hereditary versus lifestyle-related cancers. The third choice question, coined the *hazard* question, asked the participants how they would allocate some fixed amount of money between a smoker and non-smoker who both have been exposed to a lung cancer hazard.

We chose the equal division between the alternatives as the default allocation for all choice questions. The deviations of the respondents' selections from this default are denoted by y_p, y_t and y_h for the prevention, treatment and hazard questions, respectively. In what follows, we will call these variables as *allocations relative to the default* or, simply, *relative allocations*. Positive values of these variables reflect preference for giving to hereditary cancers in the case of variables y_p and y_t and to non-smokers in the case of variable y_h. The sample statistics for the three relative allocations are reported in Table 1. It reveals that 66% of the respondents allocated a strictly larger share to the prevention of hereditary

[9]Catch trial questions preceded the main body of the survey to ensure that only data for the respondents who paid sufficient attention was retained. In some of the survey questions, the respondents were asked to choose minimum and maximum values for a certain variable. If they ignored the instructions that "the minimum should be less than the maximum", they were removed from the sample (31 respondents). There were also 6 respondents who failed to answer at least one of the choice questions.

cancers while 12% strictly preferred the prevention of lifestyle-related cancers. The pattern for the treatment question is very similar. For the hazard question, 67% of the respondents strictly favored the non-smoker and only 3% strictly favored the smoker. The percentages of the respondents who chose the equal split were 22% for the prevention question, 22% for the treatment question, and 30% for the hazard question.

Table 1. Descriptive Statistics of Choice Questions

(Based on 166 observations from Maximiles platform – November 24-28, 2014)

Task	Mean	Standard deviation	Median	Share of subjects with $-50 \leqslant y_i < 0$	Share of subjects with $y_i = 0$	Share of subjects with $0 < y_i \leqslant 50$
Overall						
Prevention	16.26	22.92	17.00	12.05	21.69	66.27
Treatment	15.47	23.00	12.00	13.86	22.29	63.86
Hazard	17.02	18.42	16.00	3.01	30.12	66.87
Non-smokers						
Prevention	17.35	23.51	20.00	12.50	16.91	70.59
Treatment	18.04	22.50	20.00	11.76	16.91	71.32
Hazard	19.68	18.76	20.00	2.21	23.53	74.26
Smokers						
Prevention	11.33	19.59	0.00	10.00	43.33	46.67
Treatment	3.83	21.97	0.00	23.33	46.67	30.00
Hazard	4.97	10.40	0.00	6.67	60.00	33.33

We have also elicited information on several factors which may affect the respondents' answers to the choice questions and some additional questions of interest. First, we categorized the respondents according to their awareness of two highly publicized cancer cases. We asked the participants whether they have knowledge of Angelina Jolie's decision to take preventive actions against a hereditary cancer (variable AJ in Table 2) and Sean Connery's treatment for ailment which was likely to be a lifestyle-related cancer (variable SC in Table 2).

In 2013, Angelina Jolie, an actress, director, and humanitarian, revealed in a New York Times article that she took a preventive double mastectomy[10] after learning she had an estimated 87% risk of developing breast cancer due to an inherited defective gene, BRCA1. In her case, the chance of developing breast cancer dropped to under 5% after the surgery. The share of the respondents who knew about this or similar cases was substantial (84%).

In 1993, it was revealed that Sean Connery, who is probably best known for his role as James Bond, had received radiation therapy for undisclosed throat ailment. This news sparked media reports that the actor was suffering from throat cancer following years of heavy smoking (he started smoking when he was nine years old). The respondents who knew about this or similar cases were not as many as in Angelina Jolie's case (only 34%).[11]

The proportions of the respondents in our sample who had suffered from cancer or had relatives who had cancer were 6.6% and 63.9%, respectively. The respondents' self-descriptions of lifestyles were recorded on a five-point scale (1 = very healthy to 5 = very unhealthy). The average score for this variable was 2.6. We also categorized the participants as smokers or non-smokers because we hypothesized that the respondents in these groups might exhibit favoritism toward individuals from the same group, especially for the *hazard* question. The proportion of smokers in our sample is 18.1%, which is very close to the proportion of the wider UK population (HSCIC, 2015).

[10] Preventive double mastectomy is a prophylactic surgery of removing both breasts to reduce the risk of breast cancer in women.

[11] Our choice of these two individuals was based on a pilot study that attempted to identify the well-known cancer cases that involved celebrities.

Table 2. Means and Standard Deviations of the Independent Variables

(Based on 166 observations from Maximiles platform – November 24-28, 2014)

Variable	Mean	Std. dev.	Variable	Mean	Std. dev.
Group identity			*Likelihood perceptions*		
AJ	0.843	0.365	Personal_ambiguity	0.241	0.221
SC	0.343	0.476	Personal_hereditary_minimum	0.214	0.197
Personal history	0.066	0.249	Personal_hereditary_ambiguity	0.195	0.189
Family history	0.639	0.482	Personal_lifestyle-related_minimum	0.231	0.184
Personal lifestyle	2.578	0.749	Personal_lifestyle-related_ambiguity	0.204	0.177
Smoke	0.181	0.386	*Risk factor perceptions*		
Presenting order_tasks first	0.560	0.498	Avoid RFs, benefit minimum	0.292	0.150
Likelihood perceptions			Avoid RFs, benefit ambiguity	0.238	0.152
UK_minimum	0.286	0.142	*Socio-demographic data*		
UK_ambiguity	0.278	0.179	Gender_male	0.500	0.502
UK_hereditary_minimum	0.288	0.164	Age bands	3.620	1.504
UK_hereditary_ambiguity	0.218	0.142	Education level	4.723	1.831
UK_lifestyle-related_minimum	0.375	0.171	Household size (more than 1)	0.747	0.436
UK_lifestyle-related_ambiguity	0.247	0.166	Income level	2.018	0.937
Personal_minimum	0.292	0.192	Have children	0.620	0.487

The survey also elicited the respondents' perceived likelihoods of getting various cancers; the overall likelihood of getting cancer, the likelihood of getting hereditary cancer, and the likelihood of getting lifestyle-related cancer. The respondents were asked to provide estimates of these likelihoods separately for the UK population as a whole and for them individually. For each of these questions, the respondents could choose between intervals of estimates and single estimates. The difference between the maximal and minimal likelihoods for each question, called as *the degree of ambiguity*, was used to measure a respondent's perceived ambiguity about the corresponding hazard. Table 2 reports the descriptive statistics for the *minimum,* which is equal to the minimal likelihood estimate, and degree of ambiguity variables.[12]

We have also obtained estimates of the respondents' cancer risk factor perceptions (CRUK, 2011) which included the minimum and maximum benefits from avoiding risk factors. Finally, we collected information on the socio-demographic characteristics of the respondents including their gender, age bands, education level, household size, income level and number of children.

4.4 Results

We first examine the differences across the respondents' attitudes toward the role of beneficiaries' in affecting the likelihood of a cancer. The results of both Wilcoxon sign-rank test and t-test indicate that the relative allocations y_p, y_t and y_h are all significantly different from zero ($p = 0.00$ for both tests). The statistical tests also indicate that the respondents prefer to contribute more to hereditary cancers (in the prevention and treatment questions) and to non-smokers (in the hazard question).

Result 1. *The decision-makers give more to causes with less individual responsibility.*

[12]The subjects' perceptions about the likelihoods of getting hereditary and lifestyle-related cancers for the UK population are in line with the estimates of the corresponding likelihoods based on incidence data. According to the CRUK (2017), the lifetime cancer risk is estimated 50% while the likelihood of a lifestyles-related cancer is estimated at 42%. Both of these numbers fall in the corresponding intervals between the minimum and maximum estimates provided by our subjects.

This finding, combined with Propositions 1 and 2, suggests that a considerable share of individuals favor choice egalitarianism (see, e.g., Konow 1996, 2000, Roemer, 1998, Cappelen et al., 2013), which stipulates that only inequalities caused by differences in personal choices should remain following a redistribution policy. This result is consistent with a series of laboratory experiments which address the issue of social preferences by eliciting experimental subjects' willingness to make transfers to subjects who have taken risks versus to those who have not (see, e.g., Cappelen et al., 2007; Cappelen et al., 2010; Cappelen et al., 2013; and Melkonyan et al., 2016).

The relative allocations in our survey don't differ, however, across the prevention, treatment, and hazard questions. All of the differences are found to be statistically insignificant (Wilcoxon signed-rank tests yield $p = 0.27$ for the comparison of the prevention and treatment questions, $p = 0.66$ for the prevention and hazard questions, and $p = 0.98$ for the treatment and hazard questions). This finding suggests that our respondents hold similar fairness views for all three situations.

We also examined how the smoking status affects the relative allocations. As one may have expected, the non-smokers tend to give more to the alternatives with less individual responsibility ($p = 0.00$ for Wilcoxon sign-rank test and t-test in all three choice situations). Somewhat unexpectedly, the smokers tend to give more to hereditary cancers in the prevention question and more to the nonsmokers in the hazard question ($p = 0.01$ and 0.02, respectively, for the Wilcoxon sign-rank test). The null hypothesis that the smokers are indifferent between hereditary and lifestyle-related cancers, however, cannot be rejected in the treatment question. Furthermore, we find that the smokers tend to give more to lifestyle-related cancers in the treatment question than in the prevention question ($p = 0.01$ for the Wilcoxon sign-rank test). Finally, the Mann-Whitney test reveals that the smokers on average allocate less to causes with less individual responsibility ($p = 0.07$ for the prevention question, 0.00 for the treatment question and 0.00 for the hazard question, respectively), compared to the non-smokers. Such behavior is likely to be based on a belief that smokers will benefit themselves or people in similar circumstances by allocating more to causes with less individual responsibility. We summarize these findings in the following:

Result 2. *The non-smokers allocated significantly more to causes with less individual responsibility than the smokers. However, the smokers still allocated significantly more to hereditary cancers than to lifestyle-related cancers in the prevention question and allocated significantly more to the nonsmokers in the hazard question.*

We now turn to exploring the effects of other individual characteristics on the respondents' relative allocations. Note that the relative allocations belong to the interval $[-50\%, 50\%]$ for each of the three choice questions. Since the relative allocations are censored at 50%, we used the Tobit model (Tobin, 1958) in our estimation. We let y_i^* $(i = p, t, h)$ denote the unobserved latent variable which corresponds to the observed variable y_i (Greene, 1997):

$$y_i = \begin{cases} -50 & \text{if } y_i^* \leqslant -50 \\ y_i^* & \text{if } -50 \leqslant y_i^* \leqslant 50 \\ 50 & \text{if } y_i^* \geqslant -50 \end{cases} .$$

We separately estimated the following version of the Tobit model for each of the three questions:[13]

$$y_i^* = \beta_0 + \mathbf{C}_i'\delta + \mathbf{D}_i'\gamma + \varepsilon_i, \tag{4.13}$$

[13] Since the relative allocations were chosen contemporaneously, the disturbances in the estimated equations for the three choice questions may be correlated. To investigate this possibility, the equations in (4.13) were estimated simultaneously as a system, using the seemingly unrelated regression technique (Zellner, 1962). The results were similar to the case of a separate estimation.

where $\mathbf{C}_i$ is the vector of all continuous variables (*likelihood perceptions* and *cancer awareness* variables, Table 2), and $\mathbf{D}_i$ is the vector of all dummy variables and dummy coded categorical variables (*group identity* and *socio-demographic data*, Table 2). In our estimation, we also accounted for heteroskedasticity by using White's (1980) estimator (see also Cameron and Trivedi, 2010). The main regression results are reported in Table 3.[14]

The estimation of equation (4.13) yields a number of interesting findings. The respondents' awareness of highly publicized cancer cases had a statistically significant impact on the corresponding relative allocations. The respondents who were aware of Angelina Jolie's or similar cases are estimated to invest 16.50% more in the treatment of a hereditary cancer. Similarly, the respondents who were aware of Sean Connery's or similar cases are estimated to invest 6.98% more in the prevention of a lifestyle-related cancer. Thus, possessing information of this nature positively affects the corresponding donation irrespective of whether the cancer in question is hereditary or lifestyle-related.

Result 3. *Awareness of highly publicized or similar cases affects the relative allocations in favor of the publicized cancer type.*

This finding is consistent with previous studies, where media coverage has been found to be positively related to prosocial behaviour toward the reported issue (see, e.g., Simon, 1997; Brown and Minty, 2008). There are a number of potential drivers of this relationship. Preferences over donations to different causes may depend on the information about these causes. In our context, this information may pertain, for example, to the group affected by the disease and the urgency of the need (see, e.g., Bekkers and Wiepking, 2011).

Media coverage may also decrease the "social distance"[15] between beneficiaries and donors. The closer the social distance, the stronger the empathy. When a disease has been

[14]The multicollinearity is not a cause for concern since the variance inflation factors for all of our variables are below 4.

[15]According to Loewenstein and Small (2007), social distance refers to "feelings of connection (or lack thereof) between two individuals".

covered by the media, subjects may feel connected with its (potential) victims. As a result, these victims may become "identifiable" and no longer be perceived as simply unknown or "statistical" lives (Schelling, 1968).

Our estimation results suggest that the respondents' personal history does not affect the relative allocations. A possible explanation for this phenomenon is the small proportion of the respondents (6.6%) who reported relevant personal history.

In contrast to personal history, a much larger share of the participants in our survey reported that one or more of their family members suffered from cancer (about 64%). It was found that having a family history of cancer results in a statistically significant decrease of 9.28% in the allocation to hereditary cancers for the treatment question. The effects of a family history for the other two questions are statistically insignificant.

Our estimation results show that the relative allocations were affected by the respondents' perceptions of own lifestyles. The less healthy the respondents' self-described lifestyle, the more she/he allocated to the smokers in the hazard question. Specifically, a 1 unit decrease in the self-reported healthiness of lifestyle, represented by a 1 unit increase in the corresponding index, is estimated to lead to a statistically significant increase of £4.27 in the amount allocated to the smokers.

As one could have also expected, the smoking status impacted the respondents' relative allocations. Compared to the non-smokers in our sample, the smokers are estimated to allocate £12.48 more to a lifestyle-related cancer in the treatment question and £11.97 more to the smoker in the hazard question. Thus, the respondents exhibit "in-group favoritism" (Sumner, 1906, Akerlof and Kranton, 2000; Rudman and Goodwin, 2004; Chen and Li, 2009) conditional on both the smoking status and lifestyle.

Table 3. Regression Results

Variable	Prevention	Treatment	Hazard
Angelina Jolie's case	10.62	16.50^{***}	−0.51
	(7.17)	(6.14)	(4.32)
Sean Connery's case	-6.98^{*}	−5.26	−1.49
	(4.21)	(4.05)	(2.97)
Personal history	2.58	−8.95	−6.86
	(9.16)	(7.04)	(5.07)
Family history	−6.11	-9.28^{**}	−4.16
	(4.17)	(3.80)	(3.05)
Personal lifestyle	3.67	1.11	-4.27^{**}
	(3.16)	(2.51)	(1.97)
Smoke	−5.14	-12.48^{***}	-11.97^{***}
	(4.43)	(4.67)	(3.76)
Presenting order_tasks first	3.24	0.60	-9.21^{***}
	(4.08)	(3.80)	(2.78)
UK_minimum	−0.23	−0.16	−0.01
	(0.21)	(0.19)	(0.16)
UK_ambiguity	0.22	0.38^{***}	0.07
	(0.15)	(0.13)	(0.13)
UK_hereditary_minimum	−0.05	−0.24	−0.03
	(0.16)	(0.15)	(0.12)
UK_hereditary_ambiguity	−0.24	-0.57^{***}	−0.15
	(0.19)	(0.17)	(0.17)
UK_lifestyle-related_minimum	0.28^{**}	0.35^{***}	−0.02
	(0.13)	(0.12)	(0.10)
UK_lifestyle-related_ambiguity	0.36^{**}	0.30^{**}	0.12
	(0.15)	(0.15)	(0.10)
Personal_minimum	0.17	0.08	−0.18
	(0.21)	(0.20)	(0.18)
Personal_ambiguity	−0.04	−0.01	−0.05
	(0.11)	(0.10)	(0.09)
Personal_hereditary_minimum	0.08	0.17	0.31^{***}
	(0.18)	(0.17)	(0.11)
Personal_hereditary_ambiguity	0.05	−0.02	0.07
	(0.12)	(0.09)	(0.08)
Personal_lifestyle-related_minimum	-0.47^{***}	−0.26	−0.12
	(0.18)	(0.17)	(0.13)
Personal_lifestyle-related_ambiguity	−0.05	−0.04	0.03
	(0.15)	(0.13)	(0.11)
Avoid RFs, benefit minimum	0.18	0.03	0.17^{*}
	(0.15)	(0.14)	(0.10)
Avoid RFs, benefit ambiguity	0.03	0.19	−0.16
	(0.17)	(0.16)	(0.14)
Gender_male	-8.06^{*}	−5.13	0.57
	(4.23)	(3.75)	(2.87)
Age bands	0.37	1.71	−0.17
	(1.39)	(1.31)	(1.09)
Education level	−0.86	−1.29	−0.11
	(1.15)	(1.07)	(0.74)
Household size (more than 1)	−0.47	2.31	0.67
	(5.67)	(4.95)	(3.83)
Income level	3.17^{*}	2.25	−0.29
	(1.85)	(1.78)	(1.48)
Have children	−6.04	−5.92	−1.53
	(4.65)	(4.53)	(3.64)

*,**,*** indicate significance at the 10, 5, and 1 percent levels, respectively.
Robust standard errors are reported in parentheses.

Result 4. *The amount allocated to the smokers is larger for the decision-makers who reported smoking compared to the non-smokers as well as for the decision-makers who reported a relatively unhealthy lifestyle compared to those with a healthy lifestyle.*

The relative allocations chosen by the respondents also vary with their perceptions of various cancer risks at both the UK and individual levels. According to our estimation results, a 1 unit increase in the perceived minimal likelihood (equivalent to 1% chance of getting the disease) of a lifestyle-related cancer in the UK leads to £0.28 decrease in the allocation to lifestyle-related cancers for the prevention question and £0.35 decrease to lifestyle-related cancers for the treatment question. This result confirms the hypothesis stemming from part (b) of Propositions 3 and 4 that an increase in the minimal probability $\underline{p_a}$ of the lifestyle-related cancer will lead to a decrease in donations to the cancer.

Result 5. *The amount allocated to lifestyle-related cancers is decreasing in its perceived minimal likelihood of a lifestyle-related cancer in the UK for both the prevention and treatment questions.*

Behavior is also sensitive to the respondents' perceptions of ambiguity. To the best of our knowledge, this channel has been overlooked in the existing literature and ours is the first study to connect donation behavior with perceived ambiguity of cancer risk. A 1 percent increase in the degree of ambiguity of hereditary cancer in the UK leads to £0.57 decrease in the allocation to hereditary cancers in the prevention question; while a 1 percent increase in the degree of ambiguity of lifestyle-related cancers in the UK leads to £0.36 decrease in the allocation to lifestyle-related cancers in the prevention question and £0.30 decrease in the treatment question. The result regarding the effect of the degree of ambiguity of lifestyle-related cancers is precisely what part (c) of Propositions 3 and 4 predicts.

Result 6. *The decision-makers allocate more to cancers for which they perceive relatively* *small degree of ambiguity.*

The relationship between the respondents' perceptions of contracting cancer themselves and their allocations reveals that self-interest plays an important role in their choices. A 1 percent increase in a respondent's perceived minimal likelihood of contracting a lifestyle-related cancer herself/himself leads to an estimated £0.47 increase in the allocation to lifestyle-related cancers in the prevention question. In addition, a 1 percent increase in a respondent's perceived minimal likelihood of contracting a hereditary cancers herself/himself themselves leads to an estimated £0.31 increase in the allocation to the non-smoker in the hazard question.

Result 7. *The decision-makers exhibit a certain degree of self-interest in their choices by allocating more to cancers for which they perceive a larger minimal individual likelihood.*

The subjects' risk factor perceptions also affect their choices. For a 1% increase in the perceived minimum benefits from avoiding risk factors, the share allocated to hypothetical smokers decreased by 0.17% in the hazard question.

None of the socio-demographic characteristics are statistically significant in affecting the relative allocations for the treatment and hazard questions while gender and income are the only statistically significant effects for the prevention question. It was estimated that, compared to men, women donate £8.06 more to hereditary cancers in the prevention question. This might be due to the fact that most known hereditary cancers (such as breast and ovarian cancers caused by defective genes) occur only in women. As a result, women may possess more information on hereditary cancers and perceive prevention in this category to be more important. Alternatively, women may be more likely to exhibit choice egalitarianism. We also found that decision-makers with more financial resources donate significantly more to hereditary cancers in the prevention question; an estimated increase of £3.17 as a result of an increase of £20,000 in the income category.

Result 8. *Women and decision-makers in relatively high-income categories donate more to hereditary cancers in the prevention question.*

4.5 Conclusion

Views on individual responsibility are of great importance to numerous public debates on healthcare policies. This study elicits such views from the general public by analysing what they perceive to be just allocations between lifestyle-related and hereditary cancers. The findings of this research are multidimensional. First, people take individual responsibility into account. On average, our participants donate about twice as much to cancer with less individual responsibility (hereditary cancers) than to cancer with more individual responsibility (lifestyle-related cancers). This result is consistent with our results 1 and 2 in the next chapter, which are based on a different data set. For clarity, Result 1 in the next chapter indicates that, starting from the scenario where no information on individual responsibility is provided in the project description, adding information on hereditary/lifestyle-related causes of a cancer to the description has a positive/negative effect on donation, while Result 2 indicates that donors contribute more to cancers with smaller prevention rates. Taken together, these results from the empirical parts of this thesis provide us consistent supporting evidence to choice egalitarianism. The evidence is useful, not only because it resonates well with the previous literature on fairness principles, but also because it extracts precious views from the general public that may help construct a more equitable healthcare system or society.

Second, people exhibit "in-group favoritism" with regard to smoking status and self-reported lifestyle. The amount allocated to the smokers is larger for the decision-makers who reported smoking compared to the non-smokers as well as for the decision-makers who reported a relatively unhealthy lifestyle compared to those with a healthy lifestyle. Since in-group favoritism is revealed also in the next chapter, where women/men are found to favor to donate to women/men only cancer, our results from these two chapters are consistent on this dimension.

Third, risk perceptions of various cancer likelihoods affect peoples decisions. People allocate more to cancers which they perceive to have a relatively low level of ambiguity and a larger minimal individual likelihood of contracting. These results justify the importance of studying risk perceptions and ambiguity. They may be an important dimension for future research on fairness preferences. Lastly, apart from gender and income, social demographics seem to have no effect.

The intended contribution of this chapter (and next chapter) is to understand the publics views of justice with regard to individual responsibility, which is an essential preliminary step in constructing a more equitable healthcare system. A natural next step is to reflect these views in healthcare policies, in which multiple concerns need to be balanced. For instance, concerns for individual responsibility need to be balanced against other justice concerns such as justice according to need; and justice concerns need to be balanced against non-justice concerns such as efficiency. These issues are left for future research.

4.6 Appendix

4.6.1 Appendix A: Proofs of Propositions

Proof of Proposition 1: First, consider the non-choice egalitarian donor. It follows from her first-order condition (4.3) that

$$\beta'(d) = \beta'(100 - d).$$

By strict concavity of β, β' is strictly decreasing and hence

$$d = 100 - d \Longrightarrow d = 50.$$

Thus, the optimal allocation for the non-choice egalitarian donor entails an equal split between the two cancer types.

Next consider the choice egalitarian donor. From her first-order condition (4.5), we have

$$\beta'(d) = (1 - p_a)\,\beta'(100 - d)$$

which, together with $\beta' > 0$, implies that

$$\beta'(d) < \beta'(100 - d).$$

Since β' is strictly decreasing, it then follows that

$$d > 100 - d \Longrightarrow d > 50.$$

Thus, thve optimal allocation to the hereditary cancer is greater than 50.

Summarizing, the average allocation to the hereditary cancer is larger than that to the lifestyle-related cancer. ■

Proof of Proposition 2: First, consider the non-choice egalitarian donor. From her first-order condition (4.7), we have

$$\gamma'(d) = \gamma'(100 - d)$$

By strict convexity of γ, γ' is strictly increasing and hence

$$d = 100 - d \Longrightarrow d = 50.$$

Next, consider the choice egalitarian donor. From her first-order condition (4.9), we have

$$\gamma'(d) = (1 - p_a)\,\gamma'(100 - d)$$

which implies (note that $\gamma' < 0$) that

$$\gamma'(d) > \gamma'(100 - d)\,.$$

Since γ' is strictly increasing, the last inequality imlies

$$d > 100 - d \Longrightarrow d > 50.$$

Summarizing, the average allocation to the hereditary cancer is larger than that for the lifestyle-related cancer. ■

Proof of Proposition 3: Differentiating the left-hand-side of (4.11) with respect to α, we obtain $(1 - p_s^0)\left(\underline{p_a} - \bar{p}_a\right)\beta'(100 - d) < 0$. Part (a) then follows by the implicit function theorem.

The marginal effect of change $d\underline{p_a} = d\bar{p}_a > 0$ on the left-hand-side of (4.11) is given by $(1-p_s^0)\,\beta'(100-d)\,d\underline{p}_a > 0$. Part (b) then follows by the implicit function theorem.

The marginal effect of change $d\bar{p}_a > 0 = d\underline{p}_a$ on the left-hand-side of (4.11) is given by $(1-p_s^0)\,(1-\alpha)\,\beta'(100-d)\,d\bar{p}_a > 0$. Part (c) then follows by the implicit function theorem. ■

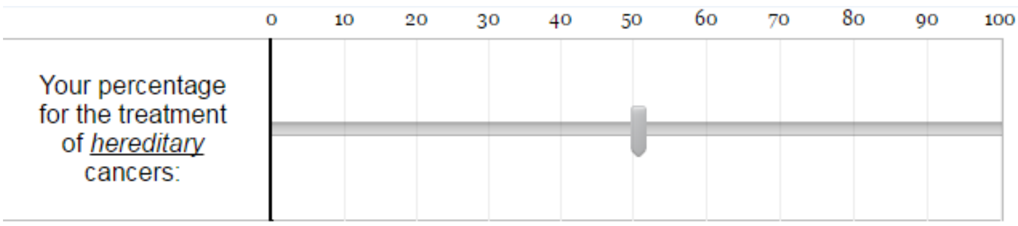
Proof of Proposition 4: Differentiating the left-hand-side of (4.12) with respect to α, we obtain $(1-p_s)\,p_c^0\left(\bar{p}_a - \underline{p}_a\right)\gamma'(100-d) < 0$. Part (a) then follows by the implicit function theorem.

The marginal effect of change $d\underline{p}_a = d\bar{p}_a > 0$ on the left-hand-side of (4.12) is given by $-(1-p_s)\,p_c^0\gamma'(100-d)\,d\underline{p}_a > 0$. Part (b) then follows by the implicit function theorem.

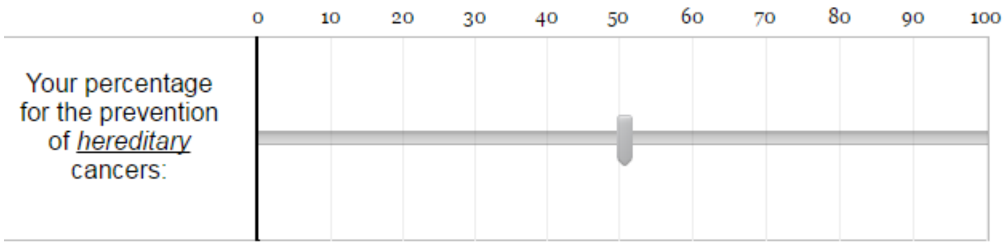
The marginal effect of change $d\bar{p}_a > 0 = d\underline{p}_a$ on the left-hand-side of (4.12) is given by $-(1-p_s)\,p_c^0\,(1-\alpha)\,\gamma'(100-d)\,d\bar{p}_a > 0$. Part (c) then follows by the implicit function theorem. ■

4.6.2 Appendix B: Choice Questions

Question 1. Suppose that £100 would be donated on your behalf to research on cancer treatment (such as chemotherapy and radiotherapy). Please indicate the percentage of this amount that you would allocate to the treatment of hereditary cancers (caused by an inherited genetic defect). The rest of the funds will go to the treatment of lifestyle-related cancers (such as smoking, poor diet, and physical inactivity).

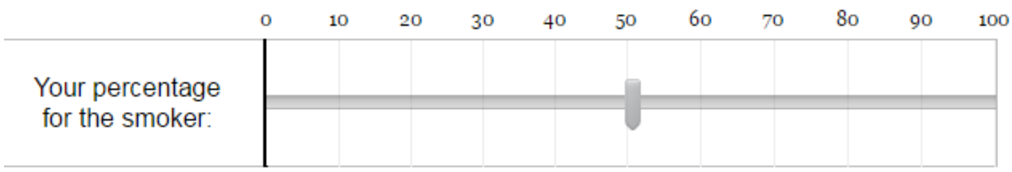


Question 2. Suppose that £100 would be donated on your behalf to research on cancer prevention (such as genetic testing). Please indicate the percentage of this amount that you would allocate to the prevention of hereditary cancers (caused by an inherited genetic defect). The rest of the funds will go to the prevention of lifestyle-related cancers (such as smoking, poor diet, and physical inactivity).



Question 3. Consider the following scenario. Suppose that a segment of general population has been exposed to a cancer hazard (for example, due to negligence by some third party). After this incident, some proportion of the exposed population contracted lung cancer. In response to this adverse outcome, the government allocated a fixed amount of funds to compensate the individuals that were exposed to the hazard.

Consider the compensation scheme of the following two individuals; both were exposed to the hazard and contracted the disease, while the first was a heavy smoker and the second not. A fixed amount of money has been allocated to compensate these two individuals. Please indicate the percentage of this amount that you would allocate to compensating the smoker (the rest of the funds will go to the non-smoker):



Chapter 5

Justice in an Uncertain World: Evidence on Donations to Cancer Research

Abstract: The paper uses information on charitable contributions to cancer research in the United Kingdom to elicit information on fairness principles endorsed by donors. The latter face a choice between contributing to several hereditary and lifestyle-related cancers and their choices of how much to donate to different cancers reveal how they view luck vis-a-vis individual choices. We find that provision of information on lifestyle-related causes of cancer adversely affects contributions. In contrast, information on hereditary causes has a positive effect on donations. Thus, a non-negligible share of the donors lean toward choice egalitarianism, which conditions outcomes on the potential beneficiaries' choices, and this is mainly due to preferences of women who tend to strongly favor choice egalitarianism.

5.1 Introduction

What a society perceives as fair is at the center stage of public debates surrounding financial bailouts of companies and countries, healthcare policies, and welfare programs. Under a democratic system of governance, these perceptions and policies reflect attitudes of individual society members to different fairness principles. Hence, assessment of these individual attitudes is of paramount importance for understanding the drivers of public policies. The present study makes an important step in this direction by using data on charitable contributions to cancer research in the United Kingdom.

In a non-deterministic world, fairness principles can be differentiated according to the degree to which individuals are held responsible for their choices vis-a-vis their luck. At one end of this spectrum of fairness principles is *strict egalitarianism* (Nielsen, 1985), which does not hold individuals responsible for any causes of inequality. According to the principle, social redistribution should be based solely on outcomes. At the opposite end of the spectrum is *libertarianism* (Nozick, 1974), which postulates that individuals should bear full responsibility for their circumstance even if they are caused by bad luck. Some theories of distributive justice combine egalitarian principles with concerns for individual responsibility[1]. One of the most notable among these is *choice egalitarianism* (Dworkin, 1981a, 1981b; Arneson 1989, Cappelen et al., 2013) that holds people responsible for their choices but not for their luck.

To understand which of these fairness principles people endorse, we use charitable contributions to research on cancer treatment and prevention made via the online platform recently developed by the Cancer Research United Kingdom (CRUK). The data collected via the platform offered a unique opportunity to distinguish the incentives to contribute to

[1] See, e.g., Rawls, 1971, Dworkin, 1981a, 1981b, Arneson 1989, Cohen, 1989, Sen, 1993, Roemer, 1998, Fleurbaey, 2008.

cancers where chance plays a central role, namely hereditary cancers, and cancers where individual choices are more important, lifestyle-related cancers. The platform enables potential donors to choose a cancer research project for sponsorship as well as provides them with full flexibility over the donation amount. The list of available projects includes both hereditary and lifestyle-related cancers and this characteristic of research projects is explicitly stated for several of them. Moreover, donors can easily access information regarding each cancer's prevention rate (this term stands for the probability that the cancer victim can avoid the cancer in question by some choice).

To disentangle different incentives of donors in our field experiment we develop a theoretical model of donation behavior for decision-makers who embrace one of different fairness principles. Our theory yields a number of testable hypotheses which are examined in the empirical part of the paper. The main theoretical result is that choice egalitarians' donations decrease with prevention rates while non-choice egalitarians are insensitive to these changes. Our empirical analysis, which is the first study to use naturally-occurring data to elicit attitudes to different fairness principles, reveals that a non-negligible part of the donors, especially women, in our data set are choice egalitarian. In a companion paper (Safra, Ma, and Melkonyan, 2017), we found a similar pattern of preferences for data obtained from an online survey where respondents faced a series of hypothetical choice questions and reported various socio-demographic characteristics.

The considerations of fair treatment of risk taking play a central role in a number of contexts. A notable illustration is offered by the functioning of a wide range of healthcare systems (Cappelen and Norheim, 2005). Two very recent examples are offered by policy changes in the operation of the National Health System (NHS) in the United Kingdom. In 2014, the NHS Blood and Transplant Service announced that it was changing its current policy by allowing people with severe drink-related liver diseases to be considered for liver transplants (The Guardian, 2014). Many questioned the appropriateness of this decision mentioning that individuals who are likely to have harmed their own health are not as deserving of treatment. The other example involves a policy that is tilted in the reverse direction. In 2016, hospital leaders in North Yorkshire, UK announced that overweight patients and smokers will be prohibited from most standard hip and knee surgeries for up to a year (The Telegraph, 2016).

Redistribution of income and wealth is another important area where individual responsibility for poor choices frequently has a key role. Diverse fairness views of how poor choices should influence redistributive policies reverberated loudly during the recent financial crisis. Many individuals and interest groups vehemently objected to using government resources to help the troubled financial institutions while others defended it on the basis that the alternative was even worse. A group of prominent economists wrote to Congress cautioning against a bailout of "particular investors and institutions whose choices proved unwise", with fairness being their primary concern (Wolfers, 2008).

The experimental evidence suggests that a considerable fraction of laboratory subjects tend to accept inequalities reflecting differences in choice (Konow, 2000; Frohlich et al., 2004; Cappelen et al., 2007; Cappelen et al., 2010; Krawczyk, 2010; Cappelen et al., 2013). The laboratory experiments provide important insights into possible behavior in real-life situations and guidance for theoretical developments. The potential for generalizability of these predictions to a wide range of domains hinges upon a combination of theory and empirical evidence from naturally occurring environments (Winkler and Murphy, 1973; Harrison and List, 2004). We make one of the first steps in this direction in the context of understanding individuals' attitudes to different distributive justice principles. In addition to providing a positive reconciliation of the findings regarding these attitudes from the laboratory and naturally-occurring environments, we offer a number of additional insights about factors that affect charitable contributions. Thus, the paper also contributes to a burgeoning literature on charitable giving (see, e.g., Auten et al., 2002; Landry et al., 2010) and provides guidance for structuring fund-raising activities.

In Section 5.2, we develop and analyze a novel model of charitable contributions that yields a number of novel insights and testable hypotheses. In addition to the result mentioned above, the model predicts that survival rates adversely affect donations. In Section 5.3, we develop the empirical model and report its results. In addition to the result mentioned above, there are three main findings. First, a non-negligible share of the donors tends to favor choice egalitarianism. Provision of information on hereditary causes of a cancer in a project description has a positive effect on donations. In contrast, information on lifestyle-related causes negatively affects contributions. We also find that at least a significant part of the latter effect is driven by women. Second, in line with our theoretical predictions, contributions are larger for cancers with smaller survival rates. Third, there is "in-gender favoritism" with each gender donating significantly more to cancers that affect only their own gender. The final section of the paper concludes and outlines avenues for future research.

5.2 Theoretical Model

We model a donor who decides how much to contribute to a cancer of certain type.[2] The donor possesses other-regarding preferences. She cares about her own utility, denoted by U^D, that is derived from her own income, and the utility of a representative cancer patient, denoted by U^P. Her preferences are represented by the function $W\left(U^D, U^P\right)$, which is assumed to be strictly increasing in both arguments and concave with non-negative cross-partial derivatives.[3] For simplicity, we assume that her own utility U^D displays risk neutrality so that $U^D(x) = x$ for all income levels x. The donor's initial income equals y and she considers donating a non-zero part d of her income to research on a specific type of cancer,

[2]The comparative statics results in this section hold in a wide range of environments where donors choose the cancer type toward which to direct their contributions. For space considerations, we do not model this choice.

[3]Utilitarianism corresponds to the case of linear W.

assuming her donation will have a positive effect on the patient's utility U^P. More specifically, she believes that her donation will increase the patient's probability of survival. Let $p_s(d)$ denote the probability of survival, or survival rate, for a patient diagnosed with this type of cancer. It is defined as

$$p_s(d) = p_s^0 + \left(1 - p_s^0\right) e(d),$$

where $p_s^0 \in (0,1)$ is the cancer's current survival rate and $e : \mathbb{R}_+ \to [0,1]$ is a strictly increasing and strictly concave function with $e(0) = 0$. The survival rate function represents the donor's beliefs that (1) her donation has a positive effect on the survival rate and (2) the size of the effect is negatively related to the current survival rate. Note that for our specification all values of $p_s(d)$ fall between p_s^0 and 1.

It is assumed that surviving the cancer, an outcome denoted by H, yields the patient, in the eyes of the donor, a utility of one unit. That is, $U^P(H) = 1$. Dying from the cancer, an outcome denoted by M, yields a lower utility. The level of utility under contingency M depends on whether the donor takes into account the extent to which this cancer may be induced by the patient's lifestyle and on whether the donor's preferences exhibit dissatisfaction with the patient for his cancer-inducing behavior. To model this possibility, let $p_r \in [0,1)$ denote the probability that the cancer victim can avoid the cancer by modifying his behavior, called the *prevention rate*. When $p_r = 0$, the cancer victim could not have avoided the cancer. In contrast, when p_r is arbitrarily close to 1, the cancer patient could have avoided the cancer with almost certainty.

The donor believes that an unhealthy lifestyle is a choice made by a supposedly informed individual. A *choice egalitarian* donor acts upon this belief in the following sense. She reasons that since the patient seems to care less about his own death by following an unhealthy lifestyle, his utility difference between being healthy and being dead is smaller than that of patients who do not follow this kind of risky behavior. To reflect this trait of the donor's preferences, we assume that $U^P(M) = p_r$ for a choice egalitarian donor. Hence, the utility difference between the two health outcomes under this assumption is given by $U^P(H) - U^P(M) = 1 - p_r$, which is decreasing with the prevention rate. If the donor is not a choice egalitarian (that is, either a strict egalitarian or a libertarian), then she uses $U^P(M) = 0$.[4]

In addition to the egalitarian dimension, the fairness principles may be differentiated depending on whether a principle of fairness is applied ex ante or ex post. At one end of this spectrum is the idea that a principle of fairness should be applied ex ante – to the expected utilities of the individuals involved (Diamond, 1967; Weymark, 1991; Epstein and Segal, 1992; Mongin, 2001; Karni and Safra, 2002; and Grant et al., 2010). At the other end is the idea that a principle should be applied ex post – to the final outcome allocations (Harel et al., 2005; Adler and Sanchirico, 2006; Fleurbaey, 2010; and Grant et al., 2012a). These two principles coincide when society adopts utilitarianism (Harsanyi, 1977b). However, as most fairness notions involve some degree of non-linearity, the two principles usually lead to different social distributions.

For the problem considered in our paper, the donor follows an *ex ante* principle if the expected utility of the patient is integrated as an argument into the function W. She follows an *ex post* principle if the values of W for all possible final outcomes are computed before taking expectations. When W is not linear, the two principles and the corresponding donation behavior differ. As will become clear below, our data does not allow to differentiate between the ex ante and ex post principles. The reason for considering those two extremes is to show that the testable predictions of our theoretical model are robust to these two formulations of the fairness principle.

Depending on their fairness views, there are two possible types of ex ante donors;

[4]Note that although libertarians are opposed to compulsory donations, they are not advocating against voluntary ones.

choice egalitarian and non-choice egalitarian donors. A *choice egalitarian ex ante* donor's objective function is given by

$$\begin{aligned} V_{ch}^{xa}(d) &= W\left(U^{D}(d), EU^{P}(d)\right) \\ &= W\left(y-d, p_s(d)\,U^{P}(H) + (1-p_s(d))\,U^{P}(M)\right) \\ &= W\left(y-d, p_s(d) + (1-p_s(d))\,p_r\right), \end{aligned}$$

where $EU^{P}(d)$ is the patient's expected utility. A *non-choice egalitarian ex ante* donor's objective is given by

$$\begin{aligned} V_{nc}^{xa}(d) &= W\left(U^{D}(d), EU^{P}(d)\right) \\ &= W\left(y-d, p_s(d)\,U^{P}(H) + (1-p_s(d))\,U^{P}(M)\right) \\ &= W\left(y-d, p_s(d)\right), \end{aligned}$$

which was derived from $V_{ch}^{xa}(d)$ by setting p_r equal to 0.

Similarly, there are two possible types of ex post donors. A *choice egalitarian ex post* donor's objective is

$$\begin{aligned} V_{ch}^{xp}(d) &= p_s(d)\,W\left(y-d, U^{P}(H)\right) + (1-p_s(d))\,W\left(y-d, U^{P}(M)\right) \\ &= p_s(d)\,W(y-d, 1) + (1-p_s(d))\,W(y-d, p_r), \end{aligned}$$

while for a *non choice egalitarian ex post* donor it is given by

$$\begin{aligned} V_{nc}^{xp}(d) &= p_s(d)\,W\left(y-d, U^{P}(H)\right) + (1-p_s(d))\,W\left(y-d, U^{P}(M)\right) \\ &= p_s(d)\,W(y-d, 1) + (1-p_s(d))\,W(y-d, 0), \end{aligned}$$

which was derived from $V_{ch}^{xp}(d)$ by setting p_r equal to 0.

In Appendix, we prove the following two propositions about how different variables affect behavior of these four types of donors.

Proposition 1

(i) *Both ex ante and ex post choice egalitarian donors donate smaller amounts to cancers with higher prevention rates.*

(ii) *Donations of all non-choice egalitarian donors are unaffected by prevention rates.*

The higher the prevention rate the smaller the difference between the recipient's utility of the positive outcome H and his utility of the adverse outcome M. Because a choice egalitarian donor's optimal donation is increasing in this difference of utilities of different outcomes, she donates less to cancers for which lifestyle has a relatively large role. Non-choice egalitarian donors don't differentiate between the prevention rates in their assessment of the adverse outcome, which immediately implies part (ii) of Proposition 1. Note also that Proposition 1 applies both to ex ante and to ex post preferences. Thus, our findings are robust against variations in the donor's fairness views along this dimension.

Proposition 2 *All donors donate smaller amounts to cancer types with higher current survival rates.*

For all four types of donors, donations are more effective in changing the survival rate when the current survival rate is relatively small. As a result, all donor types contribute relatively more to cancers with relatively dire prognoses. In the following section, we empirically test these and other theoretical predictions.

5.3 Empirical Analysis

5.3.1 Data

Our analysis utilizes a data set that is based on a newly developed online donation platform "My Project" (CRUK, 2015a). The platform is operated by CRUK, which is the world's largest independent cancer research charity. "My Project" presents potential donors with a multi-layer system to choose a project for a donation and its amount. As showed in Figure 1, this system contains four stages. In the first stage, when a potential donor enters the website, she is presented with 24 different categories, including almost all cancer types (20 categories) and some general cancer-related activities and services (4 types).[5] After one of these 24 categories is selected, in the second stage the potential donor can review all of the available projects under the chosen category.

Each project is accompanied by a description, which is presented in the third stage after the selection of a specific project. CRUK (2015a) contains full details of all projects that are currently available for sponsorship or were recently closed. To give a taste of such projects, here is an excerpt from the description of an actual project: "this project aims to understand how breast cancer establishes its nutrient supply and how these supply lines could be shut down to control breast cancer growth." (CRUK, 2015b).

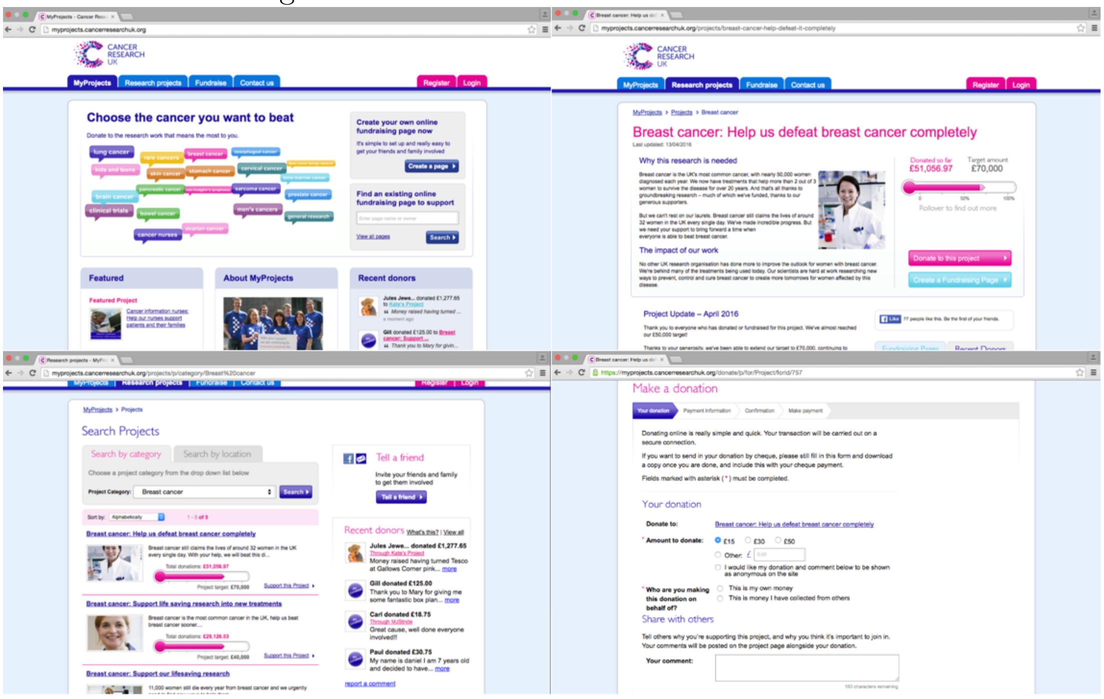
If the potential donor decides to donate to a project she then chooses the amount of donation. She can also leave comments which are then made publicly available. This is the last stage. The platform records the donation time, name of donor, chosen project, amount of donation and comments (if any).

One of the main novelties of the donation platform is the donors' full control over the destinations of their contributions. This is in stark contrast to the standard way of

[5] These four are "basic cancer biology research", "cancer nurses", "clinical trials" and "general research".

allocating donations within charities, where "directed service" is only available to very large donors[6] and donations from the rest of donors are allocated by charities on a need basis. The platform provides a unique opportunity to test the responsiveness of contributions to various aspects of sponsored projects.

Figure 1. Donation Process on the Platform



Our data covers the period from April 1, 2014 to May 11, 2015.[7] Given our research objectives, we cannot use some of the data. Since the donors' preferences over beneficiaries' individual responsibility in causing a cancer can be estimated only for specific cancer types, data for donations to four cancer categories on the platform (women's, men's, child and rare cancers) and four cancer-related services are removed from the sample. We have also combined the categories for the cervical and womb cancers and sarcomas into "cervical/womb

[6]Charities may have different interpretations for what a "very large" donation is. Irrespective of the interpretation, "directed service" is in general not available to everyone. We are grateful to David Milton, the head of the fundraising team from Worldwide Cancer Research, for clarifying these and many other specifics of charitable giving to cancer.

[7]This is all of the data that was made available to us by the CRUK. The estimation results are very similar for the data set that covers exactly one year.

cancer"[8] category since a number of projects under these categories are cross-listed. The rest of the categories are left intact. The resulting sample contains 4,879 donations to 13 categories and 32 projects.[9] Table 1 contains a summary of different categories, the number of donors for each category, and sample statistics for these donations.

Table 1 reveals that the average donation in our sample was £112 while the total amount of donations was £547,393. Among distinct cancer types, breast cancer attracted the most aggregate donation. This is a consequence of a relatively large number of donors who make contributions to breast cancer (a total of 1,334 donors) and in spite of the fact that the average donation to breast cancer is relatively small (an average of £65, which is only slightly larger than the lowest average donation of £62 to stomach cancer). The top three average donations were to cervical/womb cancer (£239), skin cancer (£235) and bone cancer (£208), respectively. The standard deviations for these donations are also considerably large (at 532, 1,115 and 769, respectively).

[8] There was only one sarcoma project and it fell into both cervical and womb categories. The numbers of the newly diagnosed cases are used as the weights in the aggregation of different categories (see the second footnote following Table 3 for details).

[9] Out of a total of 32 projects on the platform, 23 mentioned cancer treatment only and 9 mentioned both treatment and prevention.

Table 1. Descriptive Statistics for Donations by Cancer Type†

(Based on "My Project" platform – April 1, 2014 to May 11, 2015)

Cancer Type	Number of donors	Total donation	Mean	Standard deviation	Median
Bone	111	23,111	208	769	50
Bowel	282	43,901	156	558	38
Brain	584	62,879	108	354	30
Breast	1,334	86,807	65	367	25
Lung	381	47,489	125	268	30
Lymphoma‡	88	8,033	91	167	38
Oesophageal	587	51,863	88	538	25
Ovarian	197	16,133	82	199	25
Pancreatic	447	70,756	158	829	31
Prostate	315	26,046	83	223	25
Skin	227	53,293	235	1,115	38
Stomach	118	7,349	62	150	25
Cervical/Womb	208	49,732	239	532	50
Total	4,879	547,393	112	514	25

† UK taxpayers can add a 25% gift to their donation at no additional cost, since CRUK claims this additional amount from the UK's tax and customs authority. This amount is included in the reported data.

‡ Lymphoma refers to non-Hodgkin's lymphoma.

The summary statistics for the variables in our data set are reported in Table 2. We classified the projects in the dataset based on the information about hereditary and lifestyle-induced causes provided in the projects' descriptions (*description* in Table 2). Since none of the project descriptions mentions both hereditary and lifestyle-induced causes, three categories were created: (i) only hereditary causes are mentioned, (ii) only lifestyle-induced causes are mentioned, and (iii) neither hereditary nor lifestyle-induced causes are mentioned. Two projects, one for lung cancer and one for skin cancer, mention lifestyle-induced causes in their descriptions. There are also two projects that mention hereditary causes. One of these is for breast cancer while the other is for ovarian cancer.

At first glance, the statistics in Table 2 may suggest that including information about lifestyle-induced causes attracts higher donations than mentioning hereditary causes or not providing any information. However, such conclusions might be erroneous since they don't reflect all of the information about donation incentives. As the estimation results based on a model that controls for cancer type and other available information demonstrate, the opposite holds - mentioning hereditary causes has a positive effect on donations while providing information about lifestyle-related causes has the reverse effect.

The comments left by the donors provide insights into the rationale behind donations.[10] We have found four main events/factors associated with donations: attendance of a fund-raising event (variable *attend* in Table 2), loss of a family member or a friend (*loss*), donation as a gift to another person (*gift*)[11], and fighting with a cancer (*fight*). The last refers to the scenarios where the donor herself, a relative, or a friend are fighting with a cancer. About 66% of the donors in our sample left comments. Among those who provided comments, most (55%) mention attendance of a fund-raising event. The shares of donors who mention loss of a family member or a friend, current fight with the disease and donation as a gift are given by 23%, 10% and 3%, respectively.

[10]The coding of the comments was independently performed by two researchers. The results were then compared to each other.

[11]An example of a donation in this category is a Christmas gift from person A to person B that specifies that a specific contribution to the charity will be made in the name of person B.

Table 2. Descriptive Statistics

(Based on "My Project" platform – April 1, 2014 to May 11, 2015)

Variable	Number of donors	Total donation	Mean	Standard deviation	Median
Description					
Hereditary	235	18,507	79	184	26
Lifestyle-induced	454	75,402	166	792	31
Neither	4,190	453,483	108	488	25
Comment†					
Attend	1,763	192,194	109	421	25
Loss	735	155,693	212	817	50
Fighting	312	47,020	151	748	38
Gift	95	9,474	100	152	50
Others‡	2,232	201,763	90	474	25
Gender					
Female	2,541	250,545	99	412	25
Male	1,637	193,937	118	591	31
Unisex	701	102,911	147	639	30
Donor Type					
Institution	80	50,836	635	1,738	181
Non-institution	4,799	496,557	103	463	25
Total	4,879	547,393	112	514	25

† The number of donors who left comments is 3,222. Among these donors, 2,647 attributed their donations to one or more of the four categories of comments we have created.

‡ "Others" represents donors who either left no comments (1,657 donors) or ascribed their donations to factors other than those captured by the four categories of the comments (575 donors).

We used the donors' names to identify their gender and whether the donor is an institution such as a company or a non-profit organization.[12] Although men's donations are

[12]There are three categories for gender - male, female, and unisex. The latter encompasses three groups

on average 20% higher than women's, the number of women donors is substantially higher than the number of men[13] and the latter effect is so strong that the total amount of donations by women exceeds that by men. This pattern across the two genders is in line with other studies of charitable giving in the UK (see, e.g., CAF, 2015). Finally, and not surprisingly, the institutional donors contributed considerably more than the non-institutional donors. The share of the institutional donors is only 1.6%, but they account for 9.3% of the overall donations.

To understand donation behavior driven by the degree a cancer could be prevented by taking precautionary actions, we use data (CRUK, 2016a) on the prevention rate for each cancer category in our dataset. This information is easily accessible on the donation platform. According to CRUK (2016a), the prevention rate estimates the fraction of cancers that can be attributed to "modifiable lifestyle" and "other theoretically avoidable factors" (see also Parkin et al., 2011). The prevention rate variable is used as a proxy for the individual responsibility in preventing a cancer.

The ten-year survival rate in the UK for people with different types of cancer is used as a proxy for the probability of survival discussed in the theoretical part of the paper. We have used the estimated ten-year survival rate for patients diagnosed during 2010-2011. This was the most recent year for which data was available from the CRUK's website in May 2016 (CRUK, 2016b). The prevention and ten-year survival rates are reported in Table 3. Given ample evidence that donations vary with the time of the year (see, e.g., Eckel, Grossman and Milano, 2007; Tilcsik and Marquis, 2013), we also control for the month the donation was made.

of donors: donors with names that could belong to a man or a woman, couple-donors, and instutitions. The procedure of coding the names was also independently performed by two researchers. Similarly to the outcome for coding of the comments, the comparison revealed almost identical results.

[13] The ratio of the numbers of male to female donors in our sample is 0.64. This number is considerably smaller than the ratio of the two genders (0.98) in the UK population.

Table 3. Prevention and Ten-Year Survival Rates by Cancer Type[†,‡]

Cancer Type	Prevention Rate	Ten-Year Survival Rate
Bone	0.5	33
Bowel	54	57
Brain	0.5	13
Breast	27	78
Lung	89	5
Lymphoma	6	63
Oesophageal	89	12
Ovarian	21	35
Pancreatic	37	1
Prostate	0	84
Cervical/Womb[‡]	54	73
Skin	86	89
Stomach	75	15

[†] Given that CRUK (2015b) specifies the prevention rates to be "less than 0.5%" for bone and brain cancer, we have set them to 0.5%.

[‡] The prevention rate for cervical/womb category is calculated as (3,064·100%+8,475·37%)/(3,064+8,475)=54% while the ten-year survival rate as (3,064·63%+8,475·77%)/(3,064+8,475)=73%, where 3,064 is the number of cervical cancers diagnosed in 2011 while 8,475 is the number of the newly diagnosed cases of womb cancer for the same year. Note that 2011 was the most recent year for which data was available on CRUK's web site in May 2015 (CRUK, 2015c).

5.3.2 Empirical Model of Donation Behavior and Its Findings

We estimate the following linear lognormal model:[14]

$$\ln Y_i = \beta_0 + \beta_1 R_i + \beta_2 S_i + \mathbf{D}_i'\gamma + \mathbf{C}_i'\delta + \varepsilon_i, \tag{5.1}$$

where Y_i is the amount of a donor's contribution, R_i is the prevention rate, S_i is the ten-year survival rate, $\mathbf{D}_i$ is the vector of all dummy variables (attend, loss, fighting, gift, and institution), and $\mathbf{C}_i$ is the vector of all categorical variables (gender, cancer type, month, and description). Since heteroskedasticity is still present after taking the logarithmic transformation (Breusch-Pagan/CookWeisberg test, $p = 0.00$), White's (1980) heteroskedasticity-consistent estimator is used in all specifications of the empirical model.

The main estimation results are reported in Table 4. We compare four specifications, A through D, which are differentiated by inclusion of dummy variables for the donation month and selected cancer type.[15] Specification A does not control for either factor. Specifications B and C add controls for only one of these factors. Finally, specification D includes both month and cancer type controls.

According to the estimation results, each of these sets of dummy variables is statistically significant (robust Wald test, $p = 0.00$ for both sets). The coefficients of determination R^2 for these specifications indicate that the model with both sets of dummy variables (specification D) provides a better fit of the data than the other specifications. The signs of the coefficient estimates are consistent across all specifications, with the exception of the coefficient for the dummy variable that characterizes mentioning lifestyle-induced causes of a cancer in the project's description. A comparison of specifications $A - D$ also highlights

[14]A logarithmic transformation of the donation amount reduces the skewness of the distribution from 17.00 to 0.87 and the kurtosis from 358.57 to 4.47.

[15]Our choice of these two sets of dummy variables to form specifications $A - D$ is predicated about the large number of elements in each set.

why controlling for cancer type and other relevant information may reverse the effects of information about lifestyle-induced and hereditary causes of a cancer that are based solely on summary statistics in Table 2 (see also the preceding section). Given the estimation results of specifications $A - D$, in what follows we focus on specification D.

Table 4. Regression Results

	Specification			
Variable	*A*	*B*	*C*	*D*
Prevention Rate	-0.001 (0.001)	-0.000 (0.001)	-0.052^{***} (0.008)	-0.052^{***} (0.008)
Ten-Year Survival Rate	-0.001^{**} (0.001)	-0.000 (0.001)	-0.213^{***} (0.039)	-0.216^{***} (0.038)
Attend	0.138^{***} (0.036)	0.147^{***} (0.037)	0.138^{***} (0.036)	0.147^{***} (0.037)
Loss	0.740^{***} (0.054)	0.744^{***} (0.054)	0.724^{***} (0.055)	0.733^{***} (0.054)
Fighting	0.344^{***} (0.076)	0.348^{***} (0.075)	0.347^{***} (0.077)	0.352^{***} (0.077)
Gift	0.332^{***} (0.119)	0.283^{**} (0.120)	0.287^{**} (0.122)	0.240^{*} (0.123)
Institution	1.829^{***} (0.176)	1.796^{***} (0.176)	1.728^{***} (0.175)	1.705^{***} (0.175)
Gender: Unisex	0.025 (0.052)	0.034 (0.052)	0.092^{*} (0.051)	0.097^{*} (0.051)
Gender: Male	0.200^{***} (0.038)	0.196^{***} (0.037)	0.246^{***} (0.037)	0.240^{***} (0.037)
Description: Hereditary	0.025 (0.077)	0.019 (0.076)	0.206^{***} (0.080)	0.184^{**} (0.078)
Description: Lifestyle-induced	0.205^{***} (0.072)	0.151^{**} (0.074)	-0.230^{*} (0.129)	-0.247^{*} (0.129)
Month effects (11)	–	✓	–	✓
Cancer type (12)	–	–	✓	✓
R^2	0.085	0.109	0.130	0.150

Note. All specifications are based on the whole sample. Robust standard errors (White 1980) are reported in parentheses. *,**,*** significant at the 10%, 5%, and 1% levels, respectively.

We follow Halvorsen and Palmquist (1980) and Kennedy (1981) by transforming the estimated coefficients from (5.1) into what Halvorsen and Palmquist (1980) call relative effects and what we call estimated effects. For continuous variables R_i and S_i, the estimated effects are calculated as $100 \cdot \widehat{\beta}_i$, where $\widehat{\beta}_i$ is the estimate of β_i in (5.1). For the dummy variables in D_i, the estimated effects are calculated as $\exp(\widehat{D}_i - \frac{1}{2}\widehat{V}(\widehat{D}_i)) - 1$, where $\widehat{D}_i$ is the estimate of D_i and $\widehat{V}(\widehat{D}_i)$ is the estimate of the variance of $\widehat{D}_i$. The categorical variables in C_i are dummy-coded. For this reason, their estimated effects are calculated similar to those for the dummy variables. In what follows, unless stated otherwise, we report the estimated effects based on specification D in Table 4.

We begin with the results on the relationship between the individual responsibility variables and contributions. Potential donors have access to two types of information about the role of individual responsibility for different cancers. The first type of information is related to the statements about hereditary and lifestyle-induced causes in the project descriptions. The second type pertains to the prevention rates for different cancers.

If donors are choice egalitarians and take this information into account, then, according to Proposition 1, they will express their likes and dislikes via their donations. We find that this indeed is the case for both types of information (see Results 1 and 2).

The variables that characterize the information on individual responsibility in the project descriptions have a jointly significant effect on donations (robust Wald test, $p = 0.01$). Each of these variables is also statistically significant individually.

Result 1. **(a)** *Starting from the scenario where no information on individual responsibility is provided in the project description, adding information on hereditary causes of a cancer to the description has a positive effect on donations;*
(b) *Starting from the scenario where no information on individual responsibility is provided in the project description, adding information on lifestyle-induced causes of a cancer to the description has a negative effect on donations.*

Based on our estimation results, supplying information on the hereditary causes of a cancer in the project description results in a 18.4% increase in the donation; in contrast, mentioning lifestyle-induced causes in the project description leads to a 24.7% decrease in the donation (see Table 4). Result 1 has immediate implications for the design of fund-raising activities by charities. If their objective is to maximize the amount of charitable contributions, they may want to include information on hereditary causes and to exclude information on lifestyle-induced causes from the project descriptions. We also find:

Result 2. *Donors contribute more to cancers with smaller prevention rates.*

According to our estimation results, an increase of 1 unit in the prevention rate (which, in this case, is equal to 1%) leads to a 5.2% decrease in the amount of donation (see Table 4). Thus, both types of information about individual responsibility have a significant impact on donations. This finding is consistent with Proposition 1 of our theoretical model. Proposition 1 and Results 1 and 2 also suggest that a non-negligible share of the donors in our data set are choice egalitarian.

From Proposition 2, donations are positively affected by the severity of a cancer, measured by the survival rate. Our empirical model supports this finding:

Result 3. *Donors contribute more to cancers with a lower ten-year survival rate.*

The estimated effect is considerable. An increase of 1 unit in the ten-year survival rate (which is equal to 1%) leads to a 21.6% decrease in the amount of donation. Note that the effect of a change in the survival rate on donations (21.6%) is around four times the effect of a change in the prevention rate (5.2%).

We now turn to the effects of other variables on donations. Our estimates indicate that, ceteris paribus, men are expected to donate 24.0% more than women (Table 4). Recall, however, that the overall donation of women exceeds that of men since the number of women donors is 55.2% higher than the number of men (Table 2).

It is informative to juxtapose our findings to the existing estimates of the generosity of the two genders.[16] Similar to our results, a number of studies report that women are more likely to donate than men (see, e.g., Andreoni et al., 2003; Piper and Schnepf, 2008).[17] The evidence on the effect of gender on the amount of donations is more mixed. Although many studies find that men give higher amounts (Lyons and Passey, 2005; De Wit and Bekkers, 2015), some find higher contributions by women (Chang, 2005). The difference in the contributions depends significantly on the type of charity under consideration (Andreoni et al., 2003; Eckel et al., 2005). For example, men favor sports and recreation charities while women prefer health and human services and many other charity types (see, e.g., Andreoni et al., 2003 and Piper and Schnepf, 2008).

Wiepking and Bekkers (2012) argue that "the more socioeconomic variables, such as age, income and educational level, that are included in the models examining charitable giving, the smaller the reported gender differences in giving are." Since our data set doesn't contain information on income and wealth levels of the donors, a part of the estimated gender difference in donations may be due to higher levels of these two economic variables for men.

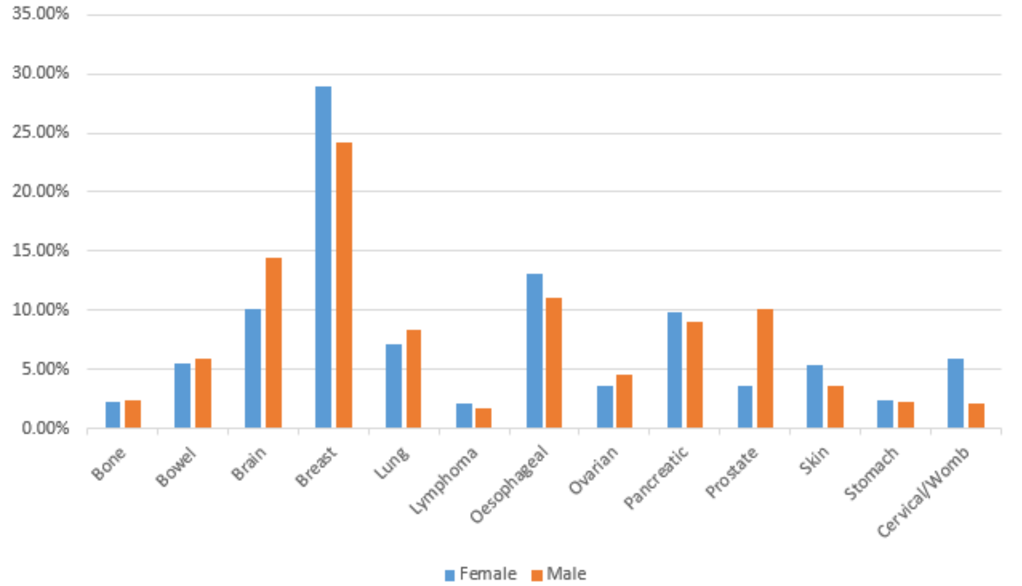
In addition to examining the relationship between gender differences and amounts of donations, we investigated how gender affects the destinations of donations. We found a significant statistical relationship ($\chi^2(12) = 138$, $p = 0.00$) by testing the hypothesis of independence between gender and cancer type (see, e.g., Agresti 2013). Figure 2 depicts the empirical distributions of female and male donors for each cancer category. For two out of three "women only" cancer categories (breast cancer and womb/cervical cancer), the number of women as a percentage of all women donors substantially exceeds the corresponding figure for men. However, for ovarian cancer the percentage for men slightly exceeds the percentage

[16] For surveys of gender differences in behavior and preferences, see, e.g. Croson and Gneezy (2009) and Wiepking and Bekkers (2012).

[17] However, there are exceptions, albeit very few. For a survey that used a variation of the dictator game, Bekkers (2007) finds that men are more likely to give to health charities than women.

for women (see Figure 2). Similar to women, men tend to favor own gender. For prostate cancer, which is the only "men only" cancer category in our data set, the number of men as a percentage of all men donors substantially exceeds the corresponding figure for women.[18]

Figure 2. Destinations of Donations by Gender
(Based on 4879 observations from "My Project" platform – April 1, 2014 to May 11, 2015)



There are a number of explanations for such behavior. First, self-interested individuals may be concerned that in the future they may get cancers affecting own gender. As a result of this purely selfish motive, they will contribute more to research on these cancers. Second, "in-group favoritism" (Sumner, 1906, Rudman and Goodwin, 2004; Chen and Li, 2009) for gender may be a result of preference for individuals of own group. The third potential explanation for "in-gender favoritism" is that donors have superior information about cancers that affect own gender and, as a result, contribute more to those cancers. The

[18] The pattern is similar when one compares "men's cancer" and "women's cancer" categories which were removed from our data set. For men's cancer, there are 32.9% female donors and 48.6% male donors; while for women's cancer, there are 56.7% female donors and 21.7% male donors. When the gender variable is interacted with the cancer types variable, we however have no evidence on males contribute more in amount than females to cancer types which have a high incidence on males or females contribute more in amount than males to cancer types which have a high incidence on females.

existing literature on in-group favoritism in charitable giving links donations to the diversity of communities in terms of ethnicity, religion, income, and other characteristics (see, e.g., Andreoni et al., 2011). According to our knowledge, there are no studies that explore "in-gender favoritism" in charitable giving. Thus, the present paper is the first attempt to tackle this important phenomenon.

We now turn to the donors' comments.[19] First, the estimation reveals that donations by individuals who mention a loss of a family member or a friend are greater by 73.3% than those of others in our sample (Table 4). This is consonant with the existing empirical results that personal experience promotes donations by increasing donors' awareness of the needs of a victim group (Burgoyne, Young and Walker, 2005; Bekkers, 2008) and reducing the social distance[20] between the donor and victims (Small and Simonsohn, 2008).[21] Second, the donors who state fighting a cancer donate 35.2% more (Table 4).

The other two factors also have a significant positive effect on donations. Those who attend a fund-raising event are estimated to donate 14.7% more (Table 4). Those who donate as a gift give 24.0% more than others (Table 4). Brown and Ferris (2007) and Wiepking and Maas (2009) also find a positive relationship between social networking and amount of donations.

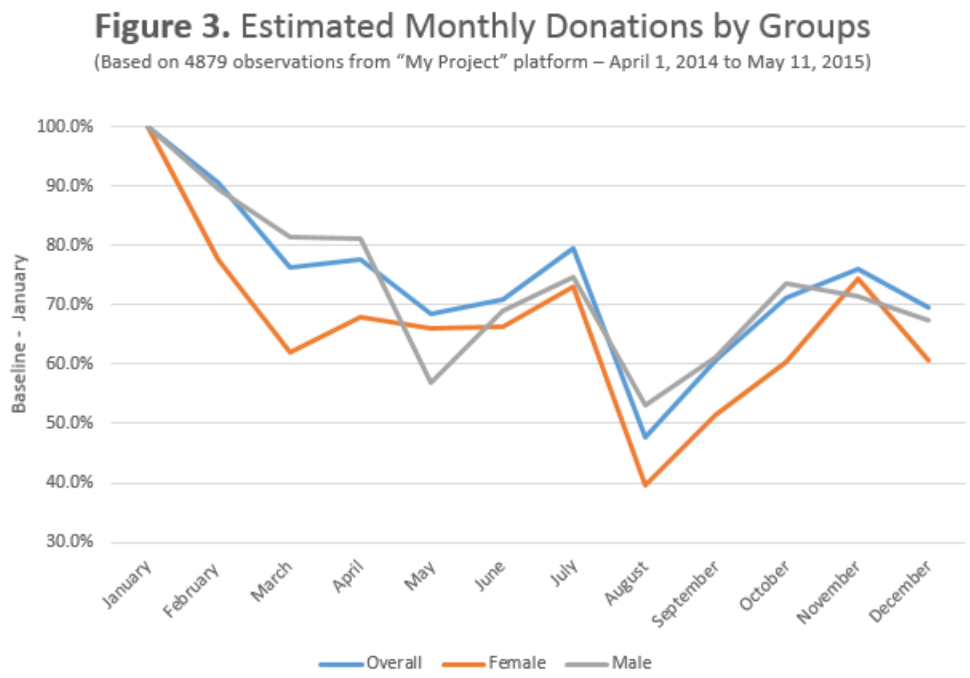
Figure 3 depicts the estimated effects of different months for the whole sample and for male-only and female-only groups. January is used as a baseline. The largest estimated effect is for January which is followed by a steady decrease in the effects until May. The estimated effects fluctuate in the following months with the bottom reached in August. For the whole sample, the estimated effect for August is only 47.8% of that for January. This

[19]Due to lack of statistical significance we have excluded the dummy variable which characterizes whether a comment was provided.

[20]Social distance refers to "feelings of connection (or lack thereof) between two individuals" (Loewenstein and Small, 2007).

[21]However, some studies find no evidence that experience of illness matters for decisions whether or not and how much to give (e.g., Smith, Kehoe and Cremer, 1995).

pattern is most likely attributable to the structure of the UK tax system where the deadlines for filing the tax returns are on January 31 and July 31 (see, e.g., Romney-Alexander, 2002). Donations in the UK are normally made through the Gift Aid system, which allows both matching and rebates from HM Revenue and Customs (Scharf and Smith, 2015). Tax relief can only be claimed at the January deadline. This may explain why the estimated donations are higher in January than in July, even though both are tax payment months.



We now conduct a closer examination of the differences among different groups of donors. Specifications E and F in Table 5 correspond to the estimation for female-only and male-only groups, respectively. A comparison of the estimation results for these two specifications reveals that the variables that proxy the individual responsibility have a statistically significant effect only for women. Mentioning lifestyle-induced causes of a cancer causes women to decrease their donations by 48.0% while a 1 percent increase in the prevention rate leads to a 5.7% decrease in donations by women (Table 5). Thus, the significant effects of these two variables for the whole sample are driven by women. Our companion paper (Safra, Ma and Melkonyan, 2017) supports this finding by demonstrating that after controlling for various socioeconomic factors women are more likely to be choice egalitarian than men.

Table 5. Regression Results by Gender

	Specification	
	Female only	*Male only*
Variable	*E*	*F*
Prevention Rate	-0.057^{***} (0.009)	-0.035 (0.027)
Ten-Year Survival Rate	-0.274^{***} (0.047)	-0.137 (0.104)
Attend	0.144^{***} (0.051)	0.147^{**} (0.065)
Loss	0.735^{***} (0.073)	0.770^{***} (0.097)
Fighting	0.311^{***} (0.095)	0.424^{***} (0.164)
Gift	0.145 (0.141)	0.404 (0.307)
Description: Hereditary	0.136 (0.108)	0.189 (0.132)
Description: Lifestyle-induced	-0.480^{***} (0.177)	-0.153 (0.217)
Number of donors	2541	1637

Note. All specifications control both time effects and cancer types. Robust standard errors (White 1980) are reported in parentheses. *,**,*** significant at the 10%, 5%, and 1% levels, respectively.

As one might have expected, the institutional donors contribute significantly more than others. According to the estimations reported in Table 4, the institutional donors give 170.5% more. Our analysis of the differences between the institutional and non-institutional donors is reported in Table 6. Specifications G and H correspond to the models for the samples of the non-institutional and institutional donors, respectively. Specification I is based on the whole sample. It includes the interaction terms for the prevention rates with the institutional dummy. The estimation results reveal that the institutional donors, similar to their non-institutional counterparts, are significantly sensitive to the individual responsibility. For a 1% increase in the prevention rate, the donation amount decreases by 5.6% for the institutional donors and by 5.2% for the non-institutional donors. Finally, the estimated effect of the interaction term indicates that the institutional donors are more sensitive to individual responsibility than non-institutional donors.

Table 6. Regression Results by Institutional and Non-Institutional Donors

	Specification		
	Non-Institutions	Institutions	Whole Sample
Variable	*G*	*H*	*I*
Prevention Rate	-0.052^{***} (0.008)	-0.056^{**} (0.024)	-0.052^{***} (0.008)
Ten-Year Survival Rate	-0.219^{***} (0.039)	0.005 (0.134)	-0.217^{***} (0.038)
Attend	0.150^{***} (0.037)	-0.295 (0.422)	0.144^{***} (0.037)
Loss	0.742^{***} (0.054)	-1.100 (0.690)	0.731^{***} (0.054)
Fighting	0.356^{***} (0.077)	1.276 (1.441)	0.351^{***} (0.077)
Gift	0.234^{*} (0.124)		0.240^{*} (0.123)
Gender: Unisex	0.098^{*} (0.051)		0.098^{*} (0.051)
Gender: Male	0.241^{***} (0.037)		0.241^{***} (0.037)
Description: Hereditary	0.182^{**} (0.079)	-0.263 (1.171)	0.181^{**} (0.078)
Description: Lifestyle-induced	-0.266^{**} (0.130)	-0.919 (0.923)	-0.251^{*} (0.129)
PreventionRate × Institution			-0.010^{**} (0.004)
Number of donors	4799	80	4879

Note. All specifications control both time effects and cancer types. Robust standard errors (White 1980) are reported in parentheses. *,**,*** significant at the 10%, 5%, and 1% levels, respectively.

5.4 Concluding Remarks

Luck and individual choices play a central role in the distribution of income, health status, and social standing across individuals within a society. We examine how people view inequalities arising from these two factors. An online platform designed by CRUK to solicit donations to cancer research offers the potential donors an ability to choose the destinations of their contributions. For some of these destinations, hereditary causes of the disease are more prevalent while for others lifestyle causes are dominant. Moreover, this information is explicitly stated for some of the cancer types on the online platform. Thus, through their contributions donors are revealing how they view the adverse health outcomes that are more likely to be caused by luck versus those for which individual choices play a relatively large role. By testing the predictions of our theoretical model, we find that donors contribute more to hereditary cancers. Based on our estimations, we conclude that a non-negligible share of the donors embrace choice egalitarianism as a guiding principle in making their contributions. Interestingly, this effect is mainly due to strong preference for choice egalitarianism among women. Among our other results are the findings that donations decrease with survival rates and that there is a significant amount of "in-gender favoritism" in donation behavior.

There are a number of interesting avenues for future research. It would be interesting to see whether preferences toward fairness principles vary across different domains. Healthcare policy, unemployment benefits, education system, and income redistribution all seem to be good candidates for an inclusion in such comparison. Another appealing area for research would be to compare attitudes to different fairness principles between the United States and European countries. As was discussed in Alesina et al. (2001) and Alesina and Angeletos (2005), a relatively high degree of redistributive taxation and welfare programs in the latter is at least in part attributable to perceptions in Europe that bad choices are less important than bad luck and more balanced perceptions in the United States. Estimating both these perceptions and preferences toward fairness principles and then juxtaposing the estimates from the United States and Europe would offer a more complete picture of the differences in social policies across the two regions. Finally, given the empirical focus of the paper, we have left a number of interesting extensions of the theoretical model to future research.

5.5 Appendix

Proof of Proposition 1: (i) We start with the analysis of the choice egalitarian ex ante donor's choice problem. The first-order condition for her optimization problem is given by

$$\begin{aligned}\frac{\partial V_{ch}^{xa}(d)}{\partial d} &= -W_1(y-d, p_s(d) + (1-p_s(d))p_r) \\ &\quad + p_s'(d)(1-p_r)W_2(y-d, p_s(d) + (1-p_s(d))p_r) \\ &= 0.\end{aligned} \tag{5.2}$$

Note that the strict concavity of the functions W and e implies that the second-order condition for the donor's optimization problem is always satisfied. Denote the optimal level of d by d^*. By the implicit function theorem, the effect of p_r on the optimal d^* is given by

$$\begin{aligned}\frac{\partial d^*}{\partial p_r} &= -\frac{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d \partial p_r}}{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d^2}} \\ &= -\frac{-W_{12}(\cdot)(1-p_s(d^*)) - p_s'(d^*)W_2(\cdot) + p_s'(d^*)(1-p_r)(1-p_s(d^*))W_{22}(\cdot)}{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d^2}} \\ &< 0,\end{aligned}$$

where the last inequality follows from the sign assumptions on the second-order derivatives of W. Hence, the optimal donation is negatively affected by the prevention rate.

Next consider a choice egalitarian ex post donor. Differentiating V_{ch}^{xp} with respect to d yields the first-order condition for her optimization problem

$$\begin{aligned}\frac{\partial V_{ch}^{xp}(d)}{\partial d} &= p_s'(d)[W(y-d,1) - W(y-d,p_r)] \\ &\quad - p_s(d)[W_1(y-d,1) - W_1(y-d,p_r)] - W_1(y-d,p_r) \\ &= 0.\end{aligned} \tag{5.3}$$

By the implicit function theorem and the properties of W, we have

$$\begin{aligned}\frac{\partial d^*}{\partial p_r} &= -\frac{\frac{\partial^2 V_{ch}^{xp}(d^*)}{\partial d \partial p_r}}{\frac{\partial^2 V_{ch}^{xp}(d^*)}{\partial d^2}} \\ &= -\frac{-p_s'(d^*)\, W_2(\cdot) - (1 - p_s(d^*))\, W_{12}(\cdot)}{\frac{\partial^2 V_{ch}^{xp}(d^*)}{\partial d^2}} \\ &< 0,\end{aligned}$$

which implies that the optimal donation of the choice egalitarian ex post donor is also negatively affected by the prevention rate.

(ii) Since the prevention rate does not appear in either V_{nc}^{xa} or V_{nc}^{xp}, the optimal donation of either donor is independent of p_r. ■

Proof of Proposition 2: Since the analysis of the choice egalitarian donor's problem subsumes that of the non-choice egalitarian's (by setting $p_r = 0$), we only deal with the former. Rewriting the first-order condition (5.2) for the choice egalitarian ex ante donor so that it explicitly features p_s^0, we obtain

$$\begin{aligned}\frac{\partial V_{ch}^{xa}(d)}{\partial d} &= -W_1\left(y - d, \left(p_s^0 + \left(1 - p_s^0\right) e(d)\right)(1 - p_r) + p_r\right) \\ &\quad + \left(1 - p_s^0\right) e'(d)\,(1 - p_r)\, W_2\left(y - d, \left(p_s^0 + \left(1 - p_s^0\right) e(d)\right)(1 - p_r) + p_r\right) \\ &= 0.\end{aligned}$$

By the implicit function theorem and the properties of W, we have

$$\begin{aligned}\frac{\partial d^*}{\partial p_s^0} &= -\frac{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d \partial p_s^0}}{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d^2}} \\ &= -\frac{-W_{12}(\cdot)(1-e(d^*))(1-p_r) - e'(d^*)(1-p_r)W_2(\cdot)}{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d^2}} \\ &\quad -\frac{(1-p_s^0)\, e'(d^*)(1-p_r)^2(1-e(d^*))W_{22}(\cdot)}{\frac{\partial^2 V_{ch}^{xa}(d^*)}{\partial d^2}} \\ &< 0,\end{aligned}$$

which implies that the optimal donation the choice egalitarian ex ante donor is negatively affected by the survival rate.

Next, rewriting the first-order condition (5.3) to explicitly include p_s^0 yields

$$\begin{aligned}\frac{\partial V_{ch}^{xp}(d)}{\partial d} &= \left(1-p_s^0\right) e'(d)\left[W(y-d,1) - W(y-d,p_r)\right] \\ &\quad -\left(p_s^0 + \left(1-p_s^0\right) e(d)\right)\left[W_1(y-d,1) - W_1(y-d,p_r)\right] - W_1(y-d,p_r) \\ &= 0.\end{aligned}$$

By the implicit function theorem and the properties of W,

$$\begin{aligned}\frac{\partial d^*}{\partial p_s^0} &= -\frac{\frac{\partial^2 V_{ch}^{xp}(d^*)}{\partial d \partial p_s^0}}{\frac{\partial^2 V_{ch}^{xp}(d^*)}{\partial d^2}} \\ &= -\frac{-e'(d^*)\left[W(y-d^*,1) - W(y-d^*,p_r)\right] - (1-e(d^*))\left[W_1(y-d^*,1) - W_1(y-d^*,p_r)\right]}{\frac{\partial^2 V_{ch}^{xp}(d^*)}{\partial d^2}}, \\ &< 0,\end{aligned}$$

which implies that the optimal donation of the choice egalitarian ex post donor is also negatively affected by the survival rate. ■

Chapter 6

Conclusion

Fairness is a central but controversial concern in social and individual decisions. In practice, many debates invoking fairness are endless and unresolvable because, even under the same circumstances, people may perceive fairness in very different ways (e.g. Konow, 2000). Is the moral sentiment of fairness something commonly shared by everyone, as Adam Smith (1759) commented in the eighteenth century? Or is fairness just a vacuous concept that people use to satisfy their self-serving purposes? To answer these questions, this thesis has examined fairness views from three different perspectives.

The first perspective is how theoretical economists model fairness. Chapter 3 revisited Harsanyi's utilitarian impartial observer theorem. Departing from Harsanyi's individual-centered method, a defence of utilitarianism was provided based on the notion of procedural fairness. This axiomatic justification of utilitarianism is probably based on the weakest assumptions made in the literature. The analysis also indicates that a desire for strict procedural fairness is incompatible with the source indifference axiom in Harsanyi's framework. This incompatibility links the current research with various non-utilitarian models in the literature.

In Chapter 4, an empirical method was adopted to examine what the general public perceives as justice in healthcare distribution. This focused on fairness issues around the rapid rise in lifestyle-related diseases, and whether people who are thought to bring bad health to themselves deserve equal treatment. Tested through an online survey, in which subjects were asked to allocate funds to two beneficiaries with different levels of individual responsibility, the results reveal that members of the general public take individual responsibility into consideration. They allocate twice as much to alternatives with less individual responsibility (hereditary cancers) as to competing alternatives with more individual responsibility (lifestyle-induced cancers). The results also show that ambiguity in individual responsibility, social-demographics and other perception factors affect allocation decisions.

Lastly, Chapter 5 investigated whether the fairness concern for individual responsibility holds in reality. This was done by analysing real-life donations to CRUK's newly-developed platform, where donors can choose between various cancers with different levels of individual responsibility. The results indicate that donors do consider individual responsibility. Their estimated donations increased when hereditary causes of cancers were mentioned, and decreased when lifestyle-induced causes of cancer were mentioned. A similar result was also found for disease prevention rates, another proxy of individual responsibility: an increase in the prevention rate leads to a decrease in estimated donations.

During the exploration of this thesis, we sincerely feel that debates about fairness in economics are likely to continue, as they have come through all the way in philosophy since Aristotle's time. This is because fairness is not only a fundamental element in almost all social and individual decisions, but also an inherent part of humanity. Ultimately, a topic like fairness may be impossible to resolve, but it is our hope that this thesis will draw some attention to and shed light on current fairness studies in economics, leading toward a fairer and better society.

Bibliography

[1] Adler, M. D. and C. W. Sanchirico. 2006. "Inequality and uncertainty: theory and legal applications." *University of Pennsylvania Law Review* 155, 279-377.

[2] Agresti, A. 2013. *Categorical Data Analysis.* Hoboken: Wiley.

[3] Akerlof, G. A. and J. L. Yellen. 1990. "The fair wage-effort hypothesis and unemployment." *Quarterly Journal of Economics* 105, 255-283.

[4] Akerlof, G. A. and R. E. Kranton. 2000. "Economics and identity." *Quarterly Journal of Economics* 115, 715-753.

[5] Alesina, A., E. L. Glaeser and B. Sacerdote. 2001. "Why doesn't the United States have a European-style welfare state?" *Brookings Papers on Economic Activity*, 187-254.

[6] Alesina, A. and G.-M. Angeletos. 2005. "Fairness and redistribution." *American Economic Review* 95, 960-980.

[7] Andreoni, J., E. Brown and I. Rischall. 2003. "Charitable giving by married couples who decides and why does it matter?" *Journal of Human Resources* 38, 111-133.

[8] Andreoni, J., A. Payne, J. D. Smith and D. Karp. 2011. "Diversity and donations: The effect of religious and ethnic diversity on charitable giving." *Journal of Economic Behavior & Organization* 128, 47-58.

[9] Anscombe, F. J. and R. J. Aumann. 1963. “A definition of subjective probability.” *Annals of Mathematical Statistics* 34, 199-205.

[10] Arneson, R. 1989. “Equality and equal opportunity for welfare.” *Philosophical Studies* 56, 77-93.

[11] Arrow, K. J. 1951. *Social Choice and Individual Values.* London: Yale University Press, 1963.

[12] Arrow, K. J. and L. Hurwicz. 1972. “An optimality criterion for decision-making under ignorance.” In C. F. Carter and J. L. Ford (eds.), *Uncertainty and expectations in economics: Essays in honour of G.L.S. Shackle.* Oxford: Basil Blackwell, 1-11.

[13] Arrow, K. J. 1977. “Extended sympathy and the possibility of social choice.” *The American Economic Review* 67, 219-225.

[14] Auten, G. E., H. Sieg and C. T. Clotfelter. 2002. “Charitable giving, income, and taxes: An analysis of panel data.” *American Economic Review* 92, 371-382

[15] Bekkers, R. 2008. “Straight from the heart.” In *Patients, Consumers and Civil Society* (Advances in Medical Sociology, Volume 10), S. Chambré, M. Goldner (ed.). Emerald Group Publishing Limited, 197-221.

[16] Bekkers, R. and P. Wiepking. 2010. “A literature review of empirical studies of philanthropy: eight mechanisms that drive charitable giving.” *Nonprofit and Voluntary Sector Quarterly* 40, 924-973.

[17] Brekke, K. A., S. Kverndokk and K. Nyborg. 2003. “An economic model of moral motivation.” *Journal of Public Economics* 87, 1967-1983.

[18] Ben-Porath, E., I. Gilboa and D. Schmeidler. 1997. “On the measurement of inequality under uncertainty.” *Journal of Economic Theory* 75, 194-204.

[19] Bentham, J. 1789. *An Introduction to the Principles of Morals and Legislation*. Oxford: Blackwell, 1948.

[20] Blackorby, C., D. Donaldson and J. A. Weymark. 1999. "Harsanyi's social aggregation theorem for state-contingent alternatives." *Journal of Mathematical Economics* 32, 365-387.

[21] Blackorby, C., D. Donaldson and P. Mongin. 2004. "Social aggregation without the expected utility hypothesis." Discussion paper 2004-020, Ecole Polytechnique.

[22] Brock, J. M., A. Lange, and E. Y. Ozbay. 2013. "Dictating the risk: Experimental evidence on giving in risky environments." *American Economic Review* 103, 415-437.

[23] Broome, J. 1978. "Trying to value a life." *Journal of Public Economics* 9, 91-100.

[24] Broome, J. 1984. "Uncertainty and fairness." *The Economic Journal* 94, 624-632.

[25] Broome, J. 1993. "A cause of preference is not an object of preference." *Social Choice and Welfare* 10, 57-68.

[26] Brown, E. and J. M. Ferris. 2007. "Social capital and philanthropy: An analysis of the impact of social capital on individual giving and volunteering." *Nonprofit and Voluntary Sector Quarterly* 36, 85-99.

[27] Brown, P. H. and J. H. Minty. 2008. "Media coverage and charitable giving after the 2004 Tsunami." *Southern Economic Journal* 75, 9-25.

[28] Brownell, K. D., R. Kersh, D. S. Ludwig, R. C. Post, R. M. Puhl, M. B. Schwartz and W. C. Willett. 2010. "Personal responsibility and obesity: a constructive approach to a controversial issue." *Health Affairs* 29, 379-387.

[29] Burgoyne, C. B., B. Young and C. M. Walker. 2005. "Deciding to give to charity: A focus group study in the context of the household economy." *Journal of Community & Applied Social Psychology* 15, 383-405.

[30] Cameron, A. C. and P. K. Trivedi. 2010. *Microeconometrics using stata*. Revised edition. College Station, Texas: Stata Press.

[31] Chew, S. H. 1983. "A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox." *Econometrica* 51, 1065-1092.

[32] Chew, S. H. 1989. "Axiomatic utility theories with the betweenness property." *Annals of operations Research* 19, 273-298.

[33] Chew, S. H., E. Karni and Z. Safra. 1987. "Risk aversion in the theory of expected utility with rank dependent probabilities." *Journal of Economic Theory* 42, 370-381.

[34] Chew, S. H. and J. S. Sagi. 2012. "An inequality measure for stochastic allocations." *Journal of Economic Theory* 147, 1517-1544.

[35] Cancer Research UK (CRUK). 2011. "Cancer awareness measure toolkit version 2.1." http://www.cancerresearchuk.org/sites/default/files/health_professional_cancer_awareness_measure_toolkit_version_2.1_09.02.11.pdf.

[36] Cancer Research UK. 2015a. http://myprojects.cancerresearchuk.org/.

[37] Cancer Research UK. 2015b. http://myprojects.cancerresearchuk.org/projects/breast-cancer-support- life-saving-research-into-new-treatments.

[38] Cancer Research UK. 2015c. http://www.cancerresearchuk.org/content/cancer-statistics-for-the-uk.

[39] Cancer Research UK. 2016a. http://www.cancerresearchuk.org/health-professional/cancer-statistics/risk/preventable-cancers#heading-One.

[40] Cancer Research UK. 2016b. http://www.cancerresearchuk.org/health-professional/cancer-statistics/survival/common-cancers-compared#heading-Zero

[41] Cancer Research UK. 2017. "Cancer risk statistics." http://www.cancerresearchuk.org/health-professional/cancer-statistics/risk.

[42] Cappelen, A. W. and O. F. Norheim. 2005. "Responsibility in health care: A liberal egalitarian approach." *Journal of Medical Ethics* 31, 476-480.

[43] Cappelen, A. W., A. D. Hole, E. Ø. Sørensen and B. Tungodden. 2007. "The pluralism of fairness ideals: An experimental approach." *American Economic Review* 97, 818-827.

[44] Cappelen, A. W., E. Ø. Sørensen and B. Tungodden. 2010. "Responsibility for what? Fairness and individual responsibility." *European Economic Review* 54, 429-441.

[45] Cappelen, A. W., J. Konow, E. Ø. Sørensen and B. Tungodden. 2013. "Just luck: An experimental study of risk-taking and fairness." *American Economic Review* 103, 1398-1413.

[46] Cettolin, E. and A. Riedl. 2016. "Justice under uncertainty." *Management Science*, forthcoming.

[47] Chang, W. C. 2005. "Religious giving, non-religious giving, and after-life consumption." *The BE Journal of Economic Analysis & Policy* 5, 1-33.

[48] Charities Aid Foundation (CAF). 2015. "UK giving 2014." https://www.cafonline.org/docs/default-source/about-us-publications/caf-ukgiving2014.

[49] Chen, Y. and S. X. Li. 2009. "Group identity and social preferences." *American Economic Review* 99, 431-457.

[50] Cohen, C. and M. Benjamin. 1999. "Alcoholics and liver transplantation." *Jama* 265, 1299-301.

[51] Cohen, G. A. 1989. "On the currency of egalitarian justice." *Ethics* 99, 906-944.

[52] Croson, R. and U. Gneezy. 2009. "Gender differences in preferences." *Journal of Economic Literature* 47, 448-474.

[53] Culyer, A. J. and A. Wagstaff. 1993. "Equity and equality in health and health care." *Journal of Health Economics* 12, 431-57.

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

[55] De Wit, A. and R. Bekkers. 2016. "Exploring gender differences in charitable giving: The Dutch case." *Nonprofit and Voluntary Sector Quarterly* 45, 741-761.

[56] Dekel, E. 1986. "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom." *Journal of Economic Theory* 40, 304-318.

[57] Dekel, E., Z. Safra and U. Segal. 1991. "Existence and dynamic consistency of Nash equilibrium with non-expected utility preferences." *Journal Economic Theory* 55, 229-246.

[58] Dhillon, A. and J. F. Mertens. 1999. "Relative utilitarianism." *Econometrica* 67, 471-498.

[59] Diamond, P. A. 1967. "Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment." *Journal of Political Economy* 75, 765-766.

[60] Dworkin, R. 1981a. "What is equality? Part 1: equality of welfare." *Philosophy & Public Affairs* 10, 185-246.

[61] Dworkin, R. 1981b. "What is equality? Part 2: equality of resources." *Philosophy & Public Affairs* 10, 283-345.

[62] Eckel, C. C., P. J. Grossman and R. M. Johnston. 2005. "An experimental test of the crowding out hypothesis." *Journal of Public Economics* 89, 1543–1560.

[63] Eckel, C. C., P. J. Grossman and A. Milano. 2007. "Is more information always better? An experimental study of charitable giving and Hurricane Katrina." *Southern Economic Journal* 74, 388-411.

[64] Edgeworth, F. Y. 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London: Kegan Paul.

[65] Elster, J. 1989. *Solomonic Judgements: Studies in the Limitation of Rationality*. Cambridge: Cambridge University Press.

[66] Elster, J., 1992. *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*. Cambridge: Cambridge University Press.

[67] Epstein, L. G. and U. Segal. 1992. "Quadratic social welfare functions." *Journal of Political Economy* 100, 691-712.

[68] Fehr, E. and K. M. Schmidt. 1999. "A theory of fairness, competition, and cooperation." *Quarterly Journal of Economics* 114, 817-868.

[69] Flage, R., T. Aven, E. Zio and P. Baraldi. 2014. "Concerns, challenges, and directions of development for the issue of representing uncertainty in risk assessment." *Risk Analysis* 34, 1196-1207.

[70] Fleming, M. 1952. "A cardinal concept of welfare." *Quarterly Journal of Economics* 66, 366-384.

[71] Fleurbaey, M. 2008. *Fairness, Responsibility, and Welfare*. New York: Oxford University Press.

[72] Fleurbaey, M. 2010. "Assessing risky social situations." *Journal of Political Economy* 118, 649-680.

[73] Fleurbaey, M. and P. Mongin. 2016. "The utilitarian relevance of the aggregation theorem." *American Economic Journal: Microeconomics* 8, 289-306.

[74] Flynn, J., P. Slovic and C. K. Mertz. 1994. "Gender, race, and perception of environmental health risks." *Risk Analysis* 14, 1101-1108.

[75] Frohlich, N., J. Oppenheimer and A. Kurki. 2004. "Modeling other-regarding preferences and an experimental test." *Public Choice* 119, 91–117.

[76] Gajdos, T. and E. Maurin. 2004. "Unequal uncertainties and uncertain inequalities: An axiomatic approach." *Journal of Economic Theory* 116, 93-118.

[77] Gajdos, T., J. M. Tallon and J. C. Vergnaud. 2004. "Decision making with imprecise probabilistic information." *Journal of Mathematical Economics* 40, 647-681.

[78] Gilboa, I. and D. Schmeidler. 1989. "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics* 18, 141-153.

[79] Gilboa, I., D. Samet and D. Schmeidler. 2004. "Utilitarian aggregation of beliefs and tastes." *Journal of Political Economy* 112, 932-938.

[80] Gorman, W. M. 1968. "The structure of utility functions." *Review of Economic Studies* 35, 367-390.

[81] Grant, S., A. Kajii, B. Polak and Z. Safra. 2010. "Generalized utilitarianism and Harsanyi's impartial observer theorem." *Econometrica* 78, 1939-1971.

[82] Grant, S., A. Kajii, B. Polak and Z. Safra. 2012a. "Equally-distributed equivalent utility, ex post egalitarianism and utilitarianism." *Journal of Economic Theory* 147, 1545-1571.

[83] Grant, S., A. Kajii, B. Polak and Z. Safra. 2012b. "A generalized representation theorem for Harsanyi's ('impartial') observer." *Social Choice and Welfare* 39, 833–846.

[84] Greene, W. H. 1997. *Econometric analysis.* 3rd ed. New Jersey: Prentice Hall.

[85] Greenfield, K. 2011. *The myth of choice: Personal responsibility in a world of limits.* New Haven: Yale University Press.

[86] Gustafsod, P. E. 1998. "Gender Differences in risk perception: Theoretical and methodological erspectives." *Risk Analysis* 18, 805-811.

[87] Halvorsen, R. and R. Palmquist. 1980. "The interpretation of dummy variables in semilogarithmic equations." *American Economic Review* 70, 474–475.

[88] Hare, R. M. 1952. *The Language of Morals.* London: Oxford University Press.

[89] Harel, A., Z. Safra and U. Segal. 2005. "Ex-post egalitarianism and legal justice." *Journal of Law, Economics, and Organization* 21, 57-75.

[90] Harrison, G. W. and J. A. List. 2004. "Field experiments." *Journal of Economic Literature* 42, 1009-1055.

[91] Harsanyi, J. C. 1953. "Cardinal utility in welfare economics and in the theory of risk-taking." *Journal of Political Economy* 61, 434-435.

[92] Harsanyi, J. C. 1955. "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility." *Journal of Political Economy* 63, 309-321.

[93] Harsanyi, J. C. 1975a. "Nonlinear social welfare functions." *Theory and Decision* 6, 311-332.

[94] Harsanyi, J. C. 1975b. "Can the maximin principle serve as a basis for morality? A critique of John Rawls's theory." *American Political Science Review* 69, 594-606.

[95] Harsanyi, J. C. 1977a. "Morality and the theory of rational behavior." *Social Research* 44, 623-656.

[96] Harsanyi, J. C. 1977b. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

[97] Harsanyi, J. C. 1979. "Bayesian decision theory, rule utilitarianism, and Arrow's impossibility theorem." *Theory and Decision* 11, 289-317.

[98] Health & Social Care Information Centre (HSCIC). 2015. "Statistics on smoking: England 2015." http://www.hscic.gov.uk/catalogue/PUB17526/stat-smok-eng-2015-rep.pdf.

[99] Herstein, I. N. and J. Milnor. 1953. "An axiomatic approach to measurable utility." *Econometrica* 21, 291-297.

[100] Hume, D. 1740. *Treatise of Human Nature*. Oxford: Oxford University Press, 2014.

[101] Hurwicz, L. 1952. "A criterion for decision-making under uncertainty." *Cowles Commission Technical Report* 355.

[102] Inada, K. I. 1963. "On a two-sector model of economic growth: comments and a generalization." *The Review of Economic Studies* 30, 119-127.

[103] Kahneman, D., J. L. Knetsch and R. Thaler. 1986. "Fairness as a constraint on profit seeking: entitlements in the market." *American Economic Review* 76, 728-741.

[104] Kant, I. 1785. *Groundwork of the Metaphysic of Morals*. New York: Harper & Row, 1964.

[105] Karni, E. 1996. "Social welfare functions and fairness." *Social Choice and Welfare* 13, 487-496.

[106] Karni, E. 1998. "Impartiality: Definition and representation." *Econometrica* 66, 1405-1415.

[107] Karni, E. and J. A. Weymark. 1998. "An informationally parsimonious impartial observer theorem." *Social Choice and Welfare* 15, 321-332.

[108] Karni, E. and Z. Safra. 2000. "An extension of a theorem of von Neumann and Morgenstern with an application to social choice theory." *Journal of Mathematical Economics* 34, 315-327.

[109] Karni, E. and Z. Safra. 2002. "Individual sense of justice: A utility representation." *Econometrica* 70, 263–284.

[110] Kennedy, P. E. 1981. "Estimation with correctly interpreted dummy variables in semilogarithmic equations." *American Economic Review* 71, 801.

[111] Konow, J. 1996. "A positive theory of economic fairness." *Journal of Economic Behavior & Organization* 31, 13-35.

[112] Konow, J. 2000. "Fair shares: Accountability and cognitive dissonance in allocation decisions." *American Economic Review* 90, 1072-1091.

[113] Konow, J. 2003. "Which is the fairest one of all? A positive analysis of justice theories." *Journal of Economic literature* 41, 1188-1239.

[114] Konow, J. 2010. "Mixed feelings: Theories of and evidence on giving." *Journal of Public Economics* 94, 279-297.

[115] Krawczyk, M. 2010. "A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution." *Journal of Public Economics* 94, 131-141.

[116] Landry, C. E., A. Lange, J. A. List, M. K. Price and N. G. Rupp. 2010. "Is a donor in hand better than two in the bush? Evidence from a natural field experiment." *American Economic Review* 100, 958-983.

[117] Le Grand, J. 1987. "Equity, health, and health care." *Social Justice Research* 1, 257-274.

[118] Loewenstein, G. and D. A. Small. 2007. "The scarecrow and the tin man: The vicissitudes of human sympathy and caring." *Review of General Psychology* 11, 112-126.

[119] Lyons, M. and A. Passey. 2005. *Giving Australia: Research on Philanthropy in Australia*. Sydney: University of Technology.

[120] Machina, M. J. 1987. "Choice under uncertainty: Problems solved and unsolved." *Journal of Economic Perspectives* 1, 121-154.

[121] Machina, M. J. 1989. "Dynamic consistency and non-expected utility models of choice under uncertainty." *Journal of Economic Literature* 27, 1622-1668.

[122] Marschak, J. 1950. "Rational behavior, uncertain prospects, and measurable utility." *Econometrica* 18, 111-141.

[123] Melkonyan, T., Z. Safra, S. Ma. 2016. "Justice in an uncertain world: evidence on donations to cancer research." *University of Warwick* Working Paper.

[124] Mill, J. S. 1863. *Utilitarianism.* Cambridge: Cambridge University Press, 2014.

[125] Mongin, P. 2001. "The impartial observer theorem of social ethics." *Economics and Philosophy* 17, 147-179.

[126] Mongin, P. and M. Pivato. 2015. "Ranking multidimensional alternatives and uncertain prospects." *Journal of Economic Theory* 157, 146-171.

[127] Moss, A. H. and M. Siegler. 1991. "Should alcoholics compete equally for liver transplantation?" *Jama* 265, 1295-1298.

[128] Nielsen, K. 1985. *Equality and Liberty: A Defense of Radical Egalitarianism.* Totowa, NJ: Rowman and Allanheld.

[129] Nozick, R. 1974. *Anarchy, State, and Utopia.* New York: Basis Books.

[130] Oberholzer-Gee, F., I. Bohnet and B. S. Frey. 1997. "Fairness and competence in democratic decisions." *Public Choice* 91, 89-105.

[131] Parkin, M., L. Boyd, S. Darby, D. Mesher, P. Sasieni, L. Walker and R. Peto. 2011. "The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010." *British Journal of Cancer* 105, Si-S81.

[132] Pattanaik, P. K. 1968. "Risk, impersonality, and the social welfare function." *Journal of Political Economy* 76, 1152-1169.

[133] Piper, G. and S. V. Schnepf. 2008. "Gender differences in charitable giving in Great Britain." *Voluntas* 19, 103-124.

[134] Quiggin J. 1982. "A theory of anticipated utility." *Journal of Economic Behavior and Organization* 3, 323-343.

[135] Quiggin J. 1993. *Generalized Expected Utility Theory: The Rank-Dependent Model.* Kluwer Academic Publishers.

[136] Rabin, M. 1993. "Incorporating fairness into game theory and economics." *American Economic Review* 83, 1281-1302.

[137] Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.

[138] Robbins, L. 1938. "Interpersonal comparisons of utility: A comment." *The Economic Journal* 43, 635-641.

[139] Roemer, J. E. 1998. *Equality of Opportunity.* Cambridge: Harvard University Press.

[140] Romney-Alexander, D. 2002. "Payroll giving in the UK: Donor incentives and influences on giving behaviour." *International Journal of Nonprofit and Voluntary Sector Marketing* 7, 84-92.

[141] Rudman, L. A. and S. A. Goodwin. 2004. "Gender differences in automatic in-group bias: Why do women like women more than men like men?" *Journal of Personality and Social Psychology* 87, 494-509.

[142] Safra, Z. and E. Weissengrin. 2003. "Harsanyi's impartial observer theorem with a restricted domain." *Social Choice and Welfare* 20, 177-187.

[143] Saito, K. 2013. "Social preferences under risk: Equality of opportunity versus equality of outcome. " *American Economic Review* 103, 3084-3101.

[144] Savage, L. J. 1954. *The Foundations of Statistics*. New York: Wiley.

[145] Scharf, K. and S. Smith. 2015. "The price elasticity of charitable giving: Does the form of tax relief matter?" *International Tax and Public Finance* 22, 330-352.

[146] Schelling, T. C. 1968. "The life you save may be your own." In B. C. Samuel (ed.), *Problems in Public Expenditure Analysis*. Washington DC: The Brookings Institute.

[147] Scitovsky, T. 1986. *Human Desire and Economic Satisfaction: Essays on the Frontiers of Economics*. Brighton: Wheatsheaf Books.

[148] Segal, U. 2000. "Let's agree that all dictatorships are equally bad." *Journal of Political Economy* 108, 569-589.

[149] Sen, A. K. 1970. *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

[150] Sen, A. K. 1976. "Welfare inequalities and Rawlsian axiomatics." *Theory and Decision* 7, 243-262.

[151] Sen, A. K. 1977. "Non-linear social welfare functions: A reply to Professor Harsanyi." In R. Butts and J. Hintikka (eds.), *Foundational Problems in the Special Sciences*. Dordrecht: Reidel Publishing Company, 297-302.

[152] Sen, A. 1979. "Equality of what?" In S. M. McMurrin (ed.), *Tanner Lectures on Human Values, Volumn I*. Cambridge: Cambridge University Press, 197-220.

[153] Sen, A. 1993. "Capability and well-being." In M. Nussbaum and A. Sen, *The Quality of Life*. Oxford: Clarendon Press.

[154] Sen, A. 2008. "Social choice." in S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*. New York: Palgrave Macmillan.

[155] Sherman, D. and R. Williams. 1995. "Liver transplantation for alcoholic liver disease." *Journal of Hepatology* 23, 474-479.

[156] Sidgwick, H. 1874. *The Methods of Ethics*. Cambridge: Cambridge University Press, 2012.

[157] Simon, A. F. 1997. "Television news and international earthquake relief." *Journal of Communication* 47, 82-93.

[158] Simon, H. A. 1955. "A behavioral model of rational choice." *Quarterly Journal of Economics* 69, 99-118.

[159] Slovic, P. 1987. "Perception of risk." *Science* 236, 280-285.

[160] Small, D. A. and U. Simonsohn. 2008. "Friends of victims: Personal experience and prosocial behavior." *Journal of Consumer Research* 35, 532-542.

[161] Smith, A. 1759. *The Theory of Moral Sentiments*. Chicago : University of Chicago Press, 1976.

[162] Smith, A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Chicago: University of Chicago Press, 1976.

[163] Smith, V. H., M. R. Kehoe and M. E. Cremer. 1995. "The private provision of public goods: Altruism and voluntary giving." *Journal of Public Economics* 58, 107-126.

[164] Sumner, W.G. 1906. *Folkways: A study of the sociological importance of usages, manners, customs, mores, and morals.* Boston: Ginn and Company.

[165] The Guardian, 2014. "Heavy drinkers to be considered for NHS liver transplants." By J. Meikle, 3 April 2014.

[166] The New York Times. 2013. "My medical choice." By Jolie A., 14 May 2013.

[167] The Telegraph, 2016. "Obese patients and smokers banned from routine surgery in 'most severe ever' rationing in the NHS." By H. Bodkin, 2 September 2016.

[168] Tilcsik, A. and C. Marquis. 2013. "Punctuated generosity: How mega-events and natural disasters affect corporate philanthropy in US communities." *Administrative Science Quarterly* 58, 111-148.

[169] Tobin, J. 1958. "Estimation of relationships for limited dependent variables." *Econometrica* 26, 24-36.

[170] Tomasetti, C. and B. Vogelstein. 2005. "Variation in cancer risk among tissues can be explained by the number of stem cell divisions." *Science* 347, 78-81.

[171] Vickrey, W. 1945. "Measuring marginal utility by reactions to risk." *Econometrica* 13, 319-333.

[172] Von Neumann, J. and O. Morgenstern. 1944. *The Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

[173] Weymark, J. A. 1981. "Generalized Gini inequality indices." *Mathematical Social Sciences* 1, 409-430.

[174] Weymark, J. A. 1991. "A reconsideration of the Harsanyi-Sen debate on utilitarianism." In J. Elster and J. E. Roemer (eds.), *Interpersonal Comparisons of Well-Being*. Cambridge: Cambridge University Press, 255-320.

[175] White, H. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica* 48, 817-838.

[176] Wiepking, P. and R. Bekkers. 2012. "Who gives? A literature review of predictors of charitable giving. Part Two: Gender, family composition and income." *Voluntary Sector Review* 3, 217-245.

[177] Wiepking, P. and I. Maas. 2009. "Resources that make you generous: Effects of social and human resources on charitable giving." *Social Forces* 87, 1973-1995.

[178] Wikler, D. 2002. "Personal and social responsibility for health." *Ethics & International Affairs* 16, 47-55.

[179] Winkler, R. L. and A. H. Murphy. 1973. "Experiments in the laboratory and the real world." *Organizational Behavior and Human Performance* 10, 252-270.

[180] Wolfers, J. 2008. "Economists on the Bailout." Freakonomics/Nytimes.com, September 23, 2008.

[181] World Health Organization. 2002. *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*. Geneva: World Health Organization.

[182] Wu, S., S. Powers, W. Zhu and Y. A. Hannun. 2016. "Substantial contribution of extrinsic risk factors to cancer development." *Nature* 529, 43-47.

[183] Yaari, M. E. and M. Bar-Hillel. 1984. "On dividing justly." *Social Choice and Welfare* 1, 1-24.

[184] Young, H. P. 1994. *Equity in Theory and Practice*. Princeton: Princeton University Press.

[185] Zellner, A. 1962. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American statistical Association* 57, 348-368.

[186] Zhou, L. 1997. "Harsanyi's utilitarianism theorems: General societies." *Journal of Economic Theory* 72, 198-207.