Linking Social and Personal Preferences:
Theory and Experiment[*]

William R. Zame, Bertil Tungodden, Erik Ø. Sørensen,

Shachar Kariv, and Alexander W. Cappelen[†]

April 4, 2020

Abstract

The attitudes of a Decision Maker toward riskless and risky choices—both personal choices and social choices—enter virtually every realm of individual decision-making. This paper asks when it is possible to *link* these attitudes. We provide a simple formalization of this question and necessary and sufficient conditions that such a link exists. We also offer an experimental test of the theory in which subjects were confronted with choices (involving monetary outcomes) in three domains: risky personal choices, riskless social choices and risky social choices. Revealed preference tests show that subject choices are generally consistent with utility maximization *within* each choice domain but frequently involve at least some errors. We test for consistency *across* choice domains using a novel nonparametric revealed preference test that accounts for these errors.

JEL Classification Numbers: C91, D63, D81.

Keywords: social preferences, risk preferences, revealed preference, experiments.

1 Introduction

Many individuals make choices that have consequences only for themselves – choices in the *personal domain*—and choices that have consequences both for themselves and for others—choices in the *social domain*. Many of these choices involve risk, so a full understanding of choice behavior in these domains requires a commensurate understanding of both the individual's preferences over consequences and the individual's attitude toward risk. It is then natural to ask: Is there a connection between an individual's attitude toward risk in the personal domain and the same individual's attitude toward risk in the social domain? This paper offers formalizations of this question, theoretical responses to this question, and an experimental test of the theory.

Our motivation for asking (and answering) this question arises not only from intellectual curiosity but also from pragmatism because we often choose— or at least influence— which individuals will be in a position to make choices that have consequences for us (and for themselves): Chairs, Deans, Mayors, Governors, Congresspersons, Senators—even Presidents. And we certainly care not only whether the President prefers peace to war and seeks tax and civil rights reforms, but what actions the President would be prepared to take to alter the risks of peace or war and the likelihood of achieving those reforms. Blockading Soviet ships bound for Cuba (as John Kennedy did) risks war, and putting forward both tax reform and civil rights legislation simultaneously (as Kennedy also did) risks accomplishing neither.[1]

But can we draw any inferences at all about the President's risky choices in the social domain from the fact that the President chooses to conduct an illicit affair or smoke in secret or invest aggressively or exaggerate his accomplishments (to mention just a few personal choices that have made headlines in recent memory)? What these personal choices have in common is that they involve (personal) risk—to the President's marriage or health, or finances or reputation. Drawing inferences about (present and future) risky social choices from knowledge of (past and present) risky personal choices would seem useful, and also possible—provided that there is a *linkage* between attitude toward risk in the personal domain and attitude toward risk in the social domain.

In this paper, we establish a theoretical linkage between preferences and risk attitudes in the social domain and in the personal domain, and provide an experimental test of this theory.[2] We formulate our problem abstractly by assuming that the individual Decision

---

[1] In fact, Kennedy accomplished neither tax reform nor civil rights legislation; both were pushed through by Lyndon Johnson after Kennedy's assassination.

[2] One important qualification needs to be remembered when interpreting our results. As the examples above illustrate, there is a question that seems both puzzling and important: what is the meaning of an individual's "attitude toward risk" in a domain in which the consequences are not monetary—or more generally, involve consequences other than consumption? We confess that we have no answer to offer to this question; indeed we suspect it has no entirely satisfactory. Keeping this in mind, we formulate an abstract model that avoids this issue entirely (we explain below) but we construct our experiment so that the consequences are *monetary* (for the subject and for one other).

Maker ($\mathcal{DM}$) has preferences over risky choices in the social domain but that that we can observe only preferences over risky choices in the personal domain and preferences over non-risky choices in the social domain. (It is convenient to view the personal domain as a subset of the social domain in which the consequences for others are fixed.) We then ask under what assumptions it will be possible to *deduce*—on the basis of these observations— preferences over risky choices in the social domain. Our theoretical results provide necessary and sufficient conditions that such deduction be possible. The required conditions depend on what is observed and, in one setting, about our assumptions as to the $\mathcal{DM}$'s degree of rationality.

The formal model considers a $\mathcal{DM}$, characterized by a fixed preference relation $\succeq$ over the set $L(\Omega)$ of lotteries on a set Ω of *social states*.[3] A subset $P \subset \omega$ of these social states have consequences only for the $\mathcal{DM}$—these are *personal states*—while the others have consequences both for the $\mathcal{DM}$ and for others. We do not observe the entire preference relation $\succeq$ on $L(\Omega)$ but only some portion of it. In our main theoretical result, and in our experimental work, we assume that we can observe the restrictions $\succeq_0$ of $\succeq$ to Ω and to $L(P)$. That is, we observe comparisons between social states (including personal states) and comparisons between personal lotteries, but we *do not* observe comparisons between social states and personal lotteries.[4]

We ask: in what circumstances it is possible to *deduce* the entire preference relation $\succeq$ from restriction $\succeq_0$? In other words, in what circumstances does $\succeq_0$ admit a *unique* extension to the preference relation $\succeq$ over the full domain of lotteries $L(\Omega)$ on social states? If we assume that the $\mathcal{DM}$'s preferences obey the usual axioms of individual choice under uncertainty – Completeness, Transitivity, Continuity, Reduction of Compound Lotteries and the Sure Thing Principle—together with a (relatively weak) axiom that we call State Monotonicity, then a necessary and sufficient condition that it be possible to deduce the *complete* preference relation $\succeq$ from the *incomplete* relation $\succeq_0$ is that the $\mathcal{DM}$ finds every social state to be indifferent to some personal state. We offer an interpretation of this condition in Section 3.[5]

This theoretical result seems clean and satisfying but it is another question entirely whether is also descriptive of reality. To address this latter question, we designed and executed an experiment in which subjects were confronted with choices in three domains:

- PERSONAL RISK The objects of choice are risky personal choices (equiprobable binary

---

[3]To make the analysis simpler and sharper, we assume that Ω is finite. This avoids subtle issues about the topology of Ω and the continuity of $\succeq$.

[4]For completeness, we do provide theoretical analysis of the setting in which we observe comparisons between social states and personal lotteries—that is, we observe the restriction $\succeq_1$ of $\succeq$ to $\Omega \cup L(P)$—but we do not have any experimental evidence in that setting.

[5]As we discuss below, State Monotonicity is a *much* weaker assumption than Independence because it compares only lotteries whose outcomes are primitives—social states—rather than lotteries whose outcomes are themselves lotteries. Almost all decision-theoretic models that have been proposed as alternatives to Expected Utility obey State Monotonicity.

lotteries whose consequences are monetary outcomes for *self* alone).

- SOCIAL CHOICE The objects of choice are riskless social choices (deterministic divisions of money between *self* and one *other*).

- SOCIAL RISK The objects of choice are risky social choices (equiprobable binary lotteries whose consequences are divisions of money between *self* and *other*).

In the experimental setting, we present each decision problem as a choice from a budget line using a graphical interface developed by Choi et al. (2007b). The PERSONAL RISK domain is identical to the domain in the (symmetric) risk experiment of Choi et al. (2007a). The SOCIAL CHOICE domain is identical to the domain in the (linear) two-person dictator experiment of Fisman et al. (2007).[6] The SOCIAL RISK domain is new; it represents the choice problem over lotteries over pairs of consumption for *self* and for *other*.

This approach has a number of advantages. First, the choice of a bundle subject to a budget constraint provides more information about preferences than a typical binary choice. Second, because the interface is extremely user-friendly, it is possible to present each subject with *many* choices in the course of a single experimental session, yielding rich individual-level data. This makes it possible to analyze behavior at the level of the individual subject, without the need to pool data or assume that preferences are identical across subjects. Finally, the rich individual-level data allow us to make direct comparisons of choices across the three domains using an attractive nonparametric econometric approach that builds on classical revealed preference analysis.

To test the predictions of our theory, we need to know the *particular* personal state to which each social state is indifferent—not just the *fact* that every social state is indifferent to some personal state. For some of our subjects, this would require making additional assumptions about the form of the underlying preferences. However, for two important classes of subjects, the predictions of our theory *are* testable:

- For subjects who are *selfish*—those who, in the SOCIAL CHOICE domain give nothing to *other*—the theory predicts that choice behavior in the PERSONAL RISK domain should *coincide* with choice behavior in the SOCIAL RISK domain.

---

[6]Ahn et al. (2014) extended the work in Choi et al. (2007a) on risk to settings with ambiguity. Building on the experimental methodology and utilizing the CentERpanel (a nationally representative panel of households in the Netherlands), Choi et al. (2014) relate findings on individual-level behaviors from the experimental data with economic information and socio-demographic information on individuals. These datasets have also been analyzed by others, including Halevy et al. (2018) and Polisson et al. (Forthcoming). Fisman et al. (2015a,b, 2017), and Li et al. (2017) build on the work in Fisman et al. (2007) to study social preferences with different samples, including the American Life Panel (ALP) (a nationally representative U.S.-based sample). Because all experimental designs share the same graphical interface, we are building on expertise we have acquired in previous work.

- For subjects who are *impartial*—those who, in the SOCIAL CHOICE domain, treat *other* symmetrically to *self*— the theory predicts that choice behavior in the SOCIAL CHOICE domain should *coincide* with choice behavior in the SOCIAL RISK domain.[7]

Among our subjects, we find many who are completely (or at least extremely) selfish and a number who are completely (or at least extremely) impartial. For these subjects, the theory of revealed preference allows us to provide an individual-level nonparametric *permutation test* of these predictions, based on the observation that if preferences in two domains are the same then choice behavior in those two domains must also be the same, and hence any random selection of choices from the *union* of the choices in those domains should be *indistinguishable* from the actual choices in the two domains. This test shows that for the overwhelming majority of selfish or impartial subjects the theoretical predictions are well supported by the experimental data, but for a significant number the theoretical predictions are rejected. According to the theory, these subjects' preferences are *not consistent* across the various choice domains.

The analysis is complicated by the fact that individual choices frequently involve at least some errors: subjects may compute incorrectly, or execute intended choices incorrectly, or err in other less obvious ways. Because of these "mistakes" subjects' preferences need not be consistent *within* a choice domain—i.e., their choices need not conform perfectly with the Generalized Axiom of Revealed Preference (GARP)—so these "mistakes" must be taken into account in the permutation test for consistency *across* the various choice domains; the discussion in Section 7 explains how we do this.

The remainder of the paper is organized as follows. Section 2 provides the theoretical framework for our analysis. Section 3 contains our main theoretical result and proof. Section 4 provides the transition from the general theory to the implications for the experimental setting. Section 5 describes the experiment, Section 6 describes the data, and Section 7 presents the tests of the theory. Section 8 presents an extension of the theoretical analysis to the setting in which we can observe comparisons between social states and personal lotteries. Section 9 describes how the paper is related to prior research and provides some concluding remarks.

2 Framework

We consider a $\mathcal{DM}$ and a given set of outcomes Ω with a distinguished proper subset $P \subset \Omega$. For convenience, we refer to elements of Ω as *social states* and to elements of P

---

[7]The objects of choice in the SOCIAL CHOICE domain are payout pairs (x, y) where x is the payout to *other* and y is the payout to *self*. In the experimental setting, we offer choices from linear budget lines $px + qy = w$. We define behavior in the SOCIAL CHOICE domain to be *selfish* if the choice subject to the budget constraint $px + qy = w$ is always of the form $(0, y)$ (payout to *other* is 0). We define behavior in the SOCIAL CHOICE domain to be *impartial* if (a, b) is chosen subject to the budget constraint $px + qy = w$ if and only if (b, a) is chosen subject to the mirror-image budget constraint $qx + py = w$.

as *personal states*. In the interpretation discussed in the Introduction, a social state has consequences for society (of which the $\mathcal{DM}$ is a member) as a whole; a personal state has consequences only for the $\mathcal{DM}$. We stress, however, that this is only an interpretation: our abstract formalization is quite general and encompasses many other interpretations. We are agnostic about the specific natures of Ω, P in part because outside observers may differ with respect to their knowledge of the relevant social and personal states and with respect to what they observe.[8]

We assume Ω is finite; this avoids subtle issues about the topology of Ω and the continuity properties of preferences. We assume that $P \subset \omega$ contains at least two states that the $\mathcal{DM}$ does not find indifferent (there are states $A, B \in P$ with $A \succ B$) and that $\Omega \setminus P \neq \emptyset$; this avoids degeneracy. For any subset $\Theta \subset \omega$, we write $L(\Theta)$ for the set of lotteries over states in Θ. We frequently write $\sum_i p_i \omega_i$ for the lottery whose outcome is the state ω_i with probability p_i. We refer to lotteries in $L(P)$ as *personal lotteries* and to lotteries in $L(\Omega)$ as *social lotteries*.

We assume that the $\mathcal{DM}$ has a preference relation $\succeq$ on $L(\Omega)$ that satisfies the familiar requirements: Completeness, Transitivity, Continuity, Reduction of Compound Lotteries and the Sure Thing Principle. These imply that we can—and do—identify the lottery $\sum_i p_i \omega$ with the certain state Ω. Throughout, we also assume that $\succeq$ obeys the following requirement, which we call *State Monotonicity*.

State Monotonicity If $\omega_i, \omega_i' \in \omega$ for $i = 1, \ldots k$, $\omega_i \succeq \omega_i'$ for each i and $p = (p_1, \ldots, p_k)$ is a probability vector, then

$$\sum_{i=1}^{k} p_i \omega_i \succeq \sum_{i=1}^{k} p_i \omega_i'$$

State Monotonicity is *equivalent* to a condition that Grant et al. (1992) call Degenerate Independence.[9] For ease of comparison, recall that the familiar (von Neumann and Morgenstern, 1947) Independence Axiom is

Independence If $W_i, W_i' \in L(\Omega)$ for $i = 1, \ldots k$, $W_i \succeq W_i'$ for each i and $p = (p_1, \ldots, p_k)$ is a probability vector, then

$$\sum_{i=1}^{k} p_i W_i \succeq \sum_{i=1}^{k} p_i W_i'$$

---

[8]In earlier versions of this paper we were more specific about the nature of Ω, P. We assumed that Ω had a product structure: $\Omega \subset X \times Z$; where, given a state $\omega = (x, z)$ the component x represents the personal component of Ω and z represents the social component of Ω. We assumed a given reference social component $z_0 \in Z$ and identified P with the product $X \times \{z_0\}$. However, because the particular structure played no role in the actual analysis, we prefer to use the more abstract formulation given here.

[9]As we do here, Grant et al. (1992) asks to what extent *all* preference comparisons can be deduced from a *subset* of preference comparisons. However, because our intent is different from Grant et al. (1992) we face quite different issues so the differences are greater than the similarities. We elaborate on this in the discussion of the related literature in Section 9.

We have formulated State Monotonicity in terms of weak preference, rather than indifference, because the two formulations are not equivalent. We have formulated Independence in terms of weak preference, rather than indifference, despite the fact that the two formulations *are* equivalent, in oder to highlight the difference between State Monotonicity and Independence. Notice that the difference between these two Axioms is precisely that Independence posits comparisons between *lotteries over lotteries*, while State Monotonicity only posits comparisons between *lotteries over states*. As we discuss below, the difference is enormous. We note that almost all decision-theoretic models that have been proposed as alternatives to Expected Utility of which we are aware obey State Monotonicity, including Weighted Expected Utility (Dekel, 1986; Chew, 1989), Rank Dependent Utility (Quiggin, 1982, 1993), and (much of) Prospect Theory (Tversky and Kahneman, 1992).

To understand the relationship of our system of Axioms to others, suppose that Ω consists of three mutually non-indifferent states $\Omega = \{A, X, B\}$; without loss, assume that $A \succ X \succ B$ and picture the familiar Marschak-Machina triangle in which each point represents a lottery $aA + xX + bB$ over the states A, X, B ($a = 0$ on the horizontal edge, $x = 0$ on the hypotenuse, and $b = 0$ on the vertical edge). Continuity requires that X be indifferent to some lottery over A, B; say $X \sim \frac{1}{2}A + \frac{1}{2}B$.[10] Assuming the other Axioms, Independence implies that the preference relation $\succeq$ admits an Expected Utility representation, so that the indifference curves in the triangle are parallel straight lines. Hence knowledge that $X \sim \frac{1}{2}A + \frac{1}{2}B$ completely determines $\succeq$ on the entire triangle. In particular, $\frac{1}{2}A + \frac{1}{2}X \sim \frac{3}{4}A + \frac{1}{4}B$, $\frac{1}{2}X + \frac{1}{2}B \sim \frac{1}{4}A + \frac{3}{4}B$ and so forth (see Figure 1A).

Betweenness, which is a weaker axiom than Independence (and is the central Axiom in Weighted Expected Utility), implies that all indifference curves are again straight lines but they need not be parallel; in particular, it may be that $\frac{1}{2}A + \frac{1}{2}X \sim \frac{1}{8}A + \frac{7}{8}B$ and $\frac{1}{2}X + \frac{1}{2}B \sim \frac{1}{10}A + \frac{9}{10}B$ (see Figure 1B). In this example, the indifference curves "fan out," becoming steeper (corresponding to higher risk aversion) when moving northeast in the triangle. Rank Dependent Utility and Prospect Theory allow for indifference curves that are not straight lines and can "fan out" or "fan in," especially near the triangle boundaries (see Figure 1C).

Our system of axioms is weaker than any of these, but our system still has bite. To see this, consider a lottery $aA + xX + bB$ and the lottery $(a + \varepsilon)A + (x - \varepsilon)X + bB$ obtained by shifting ε of the probability mass from X to A. State Monotonicity—together with Transitivity and the Sure Thing Principle—entails that

$$
\begin{aligned}
aA + xX + bB \quad &\sim \quad aA + (x - \varepsilon)X + \varepsilon X + bB \\
&\preceq \quad aA + \varepsilon A + (x - \varepsilon)X + bB \\
&\sim \quad (a + \varepsilon)A + (x - \varepsilon)X + bB.
\end{aligned}
$$

Thus our axioms imply that the preference relation $\succeq$ is increasing (from bottom to top)

---

[10]Some caution is needed here. In the absence of the Independence Axiom, which we do not assume, the Continuity Axiom is stronger than the Archimedean Axiom.

along vertical lines. Similarly, $\succeq$ is decreasing (from left to right) along horizontal lines. Put differently, State Monotonicity implies that shifting probability mass from lower-ranked states to higher-ranked states is considered an improvement. Hence the indifference curves of $\succeq$ must be "upward sloping" (pointing northeast in the triangle) but can otherwise be quite arbitrary (see Figure 1D).

[Figure 1 here]

The $\mathcal{DM}$'s preference relation $\succeq$ over *all* lotteries $L(\Omega)$ is fixed, but not known (to us). We seek to *deduce* $\succeq$ but must base this deduction on observation/inference of only a *subset* of all comparisons. Our main focus is on a setting in which we observe only the $\mathcal{DM}$'s comparisons between social states and the $\mathcal{DM}$'s comparisons between personal lotteries— but not the $\mathcal{DM}$'s comparisons between social states and personal lotteries. That is, we observe the *incomplete relation* $\succeq_0$ whose graph is:

$$\text{graph}(\succeq_0) = \text{graph}(\succeq) \cap ([L(P) \times L(P)] \cup [\Omega \times \Omega])$$

In Section 8 we consider a setting in which we observe the $\mathcal{DM}$'s comparisons between social states and personal lotteries. That is, we observe the *sub-preference relation* $\succeq_1$ whose graph is:

$$\text{graph}(\succeq_1) = \text{graph}(\succeq) \cap ([\Omega \cup L(P)] \times [\Omega \cup L(P)])$$

Observing $\succeq_0$ means observing the restriction of $\succeq$ to ω *and* the restriction of $\succeq$ to $L(P)$, whereas observing $\succeq_1$ means observing the restriction of $\succeq$ to $\Omega \cup L(P)$. Our main focus is on the setting in which we observe $\succeq_0$ rather than $\succeq_1$ because the former seems more natural (see Footnote 11) and can more be easily presented in an experimental setting.

To discern the difference between observing $\succeq_0$ and $\succeq_1$, consider once again the setting in which Ω consists of the three states A, X, B with $A \succ X \succ B$ and picture the Marschak-Machina triangle in which each point represents a lottery (p_A, p_X, p_B) over the states A, X, B. Suppose that A, B are personal states and X is a social state (and hence is not equivalent to any personal state). If we observe $\succeq_0$ we observe the ordering $A \succ_0 X \succ_0 B$ and the ordering of lotteries between A, B (shaded gray in Figure 2A), but no others. State Monotonicity assures us that from these observations we can *infer* the ordering of lotteries between A, X and lotteries between X, B (see Figure 2B). Continuity assures us that X is indifferent to *some* lottery $aA + (1-a)B$, but we do not observe *which lottery*. If we observe $\succeq_1$ then we *do* observe which lottery—but that is all (see Figure 2C). However, if we observe $\succeq_1$ and we *assume* that $\succeq$ obeys Independence—and hence has an Expected Utility representation—then observing *which lottery* completely determines $\succeq$ (see Figure 2D).

[Figure 2 here]

This illustrates that the deducibility of the $\mathcal{DM}$'s entire preference relation $\succeq$ from an incomplete sub-relation depends not only on the amount that can be observed/inferred about $\succeq$ but *also* on the degree of rationality we ascribe to the $\mathcal{DM}$—in particular on whether the observer believes/assumes that the $\mathcal{DM}$'s preferences obey the Independence Axiom.

3   Deducing Preferences

We can now formalize our question in the following way: If we observe the incomplete relation $\succeq_0$ can we *deduce* the entire preference relation $\succeq$? In different words: is $\succeq$ the *unique* complete preference relation that extends the incomplete relation $\succeq_0$ and obeys the same axioms (Completeness, Transitivity, Continuity, Reduction of Compound Lotteries, the Sure Thing Principle and State Monotonicity)?

We show that a necessary and sufficient condition that it be possible to deduce the complete preference relation $\succeq$ from the incomplete relation $\succeq_0$ is that the $\mathcal{DM}$ finds every social state to be indifferent to some personal state. When this necessary and sufficient condition is *not* satisfied, there will be *many* lotteries in $L(\Omega)$ over which the preference ordering of the $\mathcal{DM}$ $\succeq$ *cannot* be deduced from $\succeq_0$.

Theorem 1 *Assume that the $\mathcal{DM}$'s preference relation $\succeq$ satisfies Completeness, Transitivity, Continuity, Reduction of Compound Lotteries, the Sure Thing Principle and State Monotonicity. In order that $\succeq$ can be deduced from $\succeq_0$ it is necessary and sufficient that the $\mathcal{DM}$ finds every social state $\omega \in \Omega \setminus P$ to be indifferent to some personal state $\tilde{\omega} \in P$.*

Remark   Before proceeding to the proof, it may be useful to provide some interpretation of the condition that the $\mathcal{DM}$ finds every social state $\omega \in \Omega \setminus P$ to be indifferent to some personal state $\tilde{\omega} \in P$. Suppose social states are allocations of (additional) money to society, of which the $\mathcal{DM}$ is a member, and that the personal states are those social states in which no (additional) money is allocated to others. (This is the setting in our experiment.) To say that the $\mathcal{DM}$ finds every social state to be indifferent to some personal state means that, for every possible allocation of money to society, there is some other allocation that gives *nothing* to others and that the $\mathcal{DM}$ finds equally desirable. Put loosely in political terms: no matter what is proposed for society, there is a "bribe" that the $\mathcal{DM}$ could be offered that would leave him/her indifferent between accepting the bribe and implementing the social proposal—and, implicitly, some "bribe" that the $\mathcal{DM}$ would strictly prefer.

Proof.   To see that this condition is sufficient, assume that every social state ω admits a personal state equivalent $\tilde{\omega}$. State Monotonicity implies that $\sum p_i \omega_i \sim \sum p_i \tilde{\omega}_i$ for every lottery $\sum p_i \omega_i \in L(\Omega)$. Hence given two lotteries $\sum p_i \omega_i, \sum q_j \omega_j$ it follows from transitivity that

$$\sum p_i \omega_i, \succeq \sum q_j \omega_j \Leftrightarrow \sum p_i \tilde{\omega}_i \succeq \sum q_j \tilde{\omega}_j \Leftrightarrow \sum p_i \tilde{\omega}_i \succeq_0 \sum q_j \tilde{\omega}_j.$$

That is, $\succeq$ can be deduced from $\succeq_0$.

To see that this condition is necessary, we suppose that there is some social state X that the $\mathcal{DM}$ does *not* find indifferent to any personal state and construct a preference relation that agrees with $\succeq$ on $L(P)$ and on Ω but not on all of $L(\Omega)$. Note that we must take some care to ensure that the preference relation we construct obeys Continuity and State Monotonicity.

Because $\succeq$ is continuous (by assumption) and $L(\Omega)$ can be identified with a finite-dimensional simplex, which is a separable metric space, we can use Debreu's representation theorem (Debreu 1954) to find a utility function $u : L(\Omega) \to \mathbb{R}$ that represents $\succeq$, that is

$$\text{for all } \Gamma, \Gamma' \in L(\Omega) : \Gamma \succeq \Gamma' \Leftrightarrow u(\Gamma) \geq u(\Gamma').$$

Without loss, assume that the range of u is contained in the interval $[0, 1]$. We construct a new utility function $U : L(\Omega) \to \mathbb{R}$ that agrees with u on $L(P)$ and induces the same ordering as u on ω but does not induce the same ordering as u on $L(\Omega)$.

To make the remainder of the proof easier to follow, suppose for the moment that $\Omega = \{A, X, B\}$, $P = \{A, B\}$ and that X is a social state that is not indifferent to either A or B. As our earlier discussion of State Monotonicity suggests, it is easy to construct a utility function U on $L(\Omega)$ with the desired properties in this setting—but it is less easy to do so in a way that generalizes to the general setting with more than three states. We may assume without loss that $A \succ B$, and it is convenient to distinguish three cases: (i) $A \succ X \succ B$, (ii) $X \succ A \succ B$, and (iii) $A \succ B \succ X$:

- **Case (i) $A \succ X \succ B$:**   Continuity guarantees that there is some $\gamma \in (0, 1)$ such that $X \sim \gamma A + (1 - \gamma)B$; equivalently, $u(X) = u(\gamma A + (1 - \gamma)B)$. We construct a continuous utility function U on $L(\Omega)$ that agrees with u on $L(P)$ and for which $U(A) > U(X) > U(B)$—so that U induces the same ordering as u on ω—but $U(X) \neq u(X) = u(\gamma A + (1 - \gamma)B)$—so that the preference relation $\succeq_U$ induced by U does not agree with $\succeq$ on $L(\Omega)$.

  To understand the idea behind the construction, consider once again the Marschak-Machina triangle (see Figure 1D). In order that $\succeq_U$ satisfy State Monotonicity, U must be strictly increasing (from bottom to top) along vertical lines and strictly decreasing (from left to right) along horizontal lines. To construct U we first define two auxiliary functions f, g:

  $$\begin{aligned} f(aA + xX + bB) &= u(aA + xA + bB) \\ g(aA + xX + bB) &= u(aA + xB + bB) \end{aligned}$$

  for every lottery $aA + xX + bB \in L(\Omega)$. Because u is continuous, both f and g are continuous. Moreover, f is constant on vertical lines and strictly decreasing on horizontal lines, while g is strictly increasing on vertical lines and constant on horizontal lines.

Hence for every $\lambda \in (0,1)$ the convex combination $\lambda f + (1 - \lambda)g$ is strictly increasing on vertical lines and strictly decreasing on horizontal lines. Choose λ so that

$$\lambda f(X) + (1 - \lambda)g(X) = \lambda u(A) + (1 - \lambda)u(B) \neq u(X)$$

Now define the utility function $U = \lambda f + (1 - \lambda)g$, and let $\succeq_U$ be the preference relation induced by U. It is evident that $\succeq_U$ satisfies Completeness, Transitivity, Reduction of Compound Lotteries and the Sure Thing Principle. Because u, f, g are continuous, so is U; hence $\succeq_U$ satisfies Continuity. By construction, U is strictly increasing along vertical lines and strictly decreasing along horizontal lines, so $\succeq_U$ satisfies State Monotonicity. Finally, note that U agrees with u on $L(P)$ and

$$U(A) = u(A) > U(X) = \lambda u(A) + (1 - \lambda)u(B) > u(B) = U(B)$$

so that $\succeq_U$ is an extension of $\succeq$. Finally, because we have chosen λ so that $U(X) \neq u(X) = u(\gamma A + (1 - \gamma)B)$ it follows that $\succeq_U \neq \succeq$; this completes the construction in Case (i).

- **Case (ii) $\mathbf{X} \succ \mathbf{A} \succ \mathbf{B}$:** This is easier than Case (i). Continuity guarantees that there is some $\nu \in (0,1)$ for which $A \sim \nu X + (1 - \nu)B$. Choose $\lambda > 0$ for which $u(\nu A + (1 - \nu)B) + \lambda \nu \neq u(A)$ and set

$$U(xX + aA + bB) = u(xA + aA + bB) + \lambda x$$

By construction, U agrees with u on $L(P)$ and $U(X) = u(A) + \lambda x > u(A) = U(A)$ so the preference relation $\succeq_U$ represented by U is indeed an extension of $\succeq$. It is easily checked that $\succeq_U$ satisfies all the required axioms; because $U(\nu X + (1 - \nu)B) = u(\nu A + (1 - \nu)B) + \lambda \nu \neq u(A)$ we conclude that $\succeq_U \neq \succeq$; this completes the construction in Case (ii).

- **Case (iii) $\mathbf{A} \succ \mathbf{B} \succ \mathbf{X}$:** The argument is almost the same as in Case (ii): we simply interchange the roles of A, B and change the sign of the linear term. Continuity guarantees that there is some $\eta \in (0,1)$ for which $B \sim \eta A + (1 - \eta)X$. Choose $\lambda > 0$ for which $u(\eta A + (1 - \eta)B) - \lambda \eta \neq u(B)$ and set

$$U(aA + bB + xX) = u(aA + bB + xB) - \lambda x$$

By construction, U agrees with u on $L(P)$ and $U(X) = u(B) - \lambda x < u(B) = U(B)$ so the preference relation $\succeq_U$ represented by U is indeed an extension of $\succeq$. It is easily checked that $\succeq_U$ satisfies all the required axioms; because $U(\eta A + (1 - \eta)X) = u(\eta A + (1 - \eta)B) - \lambda \eta \neq u(B)$ we conclude that $\succeq_U \neq \succeq$; this completes the construction in Case (iii).

We now turn to the general setting. Here we must take account of the possible presence of additional personal and social states and of the possible differences in the ranking of the distinguished social state with respect to the additional personal states, but the main idea remains the same. Suppose that there is some social state X that the $\mathcal{DM}$ does *not* find indifferent to any personal state. Write $\mathcal{A}$ for the set of states that are strictly preferred to X according to $\succeq$, $\mathcal{B}$ for the set of states that are strictly dis-preferred to X, and $\mathcal{X}$ for the set of states that are indifferent to X. Because X is not indifferent to any personal state, no member of $\mathcal{X}$ is indifferent to any personal state; moreover, at least one of $\mathcal{A}, \mathcal{B}$ is not empty.

If $\mathcal{A} \neq \emptyset$, let A be any $\succeq$-minimal element of $\mathcal{A}$; if $\mathcal{B} \neq \emptyset$ let B be any $\succeq$-maximal element of $\mathcal{B}$. (Such minimal and maximal elements exist because Ω is finite.) For each lottery $\Gamma = \sum p_i \omega_i \in L(\Omega)$ write

$$
\Gamma_{\mathcal{A}} = \sum_{\omega_i \in \mathcal{A}} p_i \omega_i \; ; \quad \Gamma_{\mathcal{B}} = \sum_{\omega_i \in \mathcal{B}} p_i \omega_i \; ; \quad \Gamma_{\mathcal{X}} = \sum_{\omega_i \in \mathcal{X}} p_i \omega_i \; .
$$

Evidently, $\Gamma = \Gamma_{\mathcal{A}} + \Gamma_{\mathcal{B}} + \Gamma_{\mathcal{X}}$.

We now distinguish three cases that are parallel to the three cases considered above and carry through constructions that are parallel to those in those cases:

- **Case (i) $\mathcal{A} \neq \emptyset$ and $\mathcal{B} \neq \emptyset$:** Continuity guarantees there is some $\gamma \in (0,1)$ such that $X \sim \gamma A + (1-\gamma)B$. As before, define auxiliary functions $f, g : L(\Omega) \to \mathbb{R}$ by

$$
\begin{aligned}
f(\Gamma) &= u(\Gamma_{\mathcal{A}} + x(\Gamma)A + \Gamma_{\mathcal{B}}) \\
g(\Gamma) &= u(\Gamma_{\mathcal{A}} + x(\Gamma)B + \Gamma_{\mathcal{B}})
\end{aligned}
$$

where $x(\Gamma) = \sum_{\omega_i \in \mathcal{X}} p_i$. Choose λ so that $\lambda f(X) + (1-\lambda)g(X) \neq u(\gamma A + (1-\gamma)B)$ and define $U = \lambda f + (1-\lambda)g$. It is easily checked that the preference relation $\succeq_U$ induced by U satisfies all the desired axioms; because $U(X) \neq U(\gamma A + (1-\gamma)B)$ we conclude that $\succeq_U \neq \succeq$. This completes the construction in Case (i).

- **Case (ii) $\mathcal{A} = \emptyset$ and $\mathcal{B} \neq \emptyset$:** Because there are at least two inequivalent personal states, we can choose $B' \in \mathcal{B}$ with $B \succ B'$. Continuity guarantees there is some $\nu \in (0,1)$ for which $B \sim \nu X + (1-\nu)B'$. Choose $\lambda > 0$ so that $u(\nu B + (1-\nu)B') - \lambda \nu \neq u(B)$ and set
$$
U(\Gamma) = u(x(\Gamma)B + \Gamma_{\mathcal{B}}) - \lambda x(\Gamma)
$$

It is easily checked that the preference relation $\succeq_U$ induced by U satisfies all the desired axioms, and that, because $U(\nu X + (1-\nu)B') = u(\nu B + (1-\nu)B') \neq u(B)$, we conclude that $\succeq_U \neq \succeq$. This completes the construction in Case (ii).

- **Case (iii) $\mathcal{A} \neq \emptyset$ and $\mathcal{B} = \emptyset$:** Because there are at least two inequivalent personal states, we can choose $A' \in \mathcal{A}$ for which $A' \succ A$. Continuity guarantees there is some

$\eta \in (0, 1)$ for which $A \sim \eta A' + (1 - \eta)X$. Choose $\lambda > 0$ for which $u(\eta A' + (1 - \eta)A) + \lambda \eta \neq u(A)$ and set

$$U(\Gamma) = u(\Gamma_{\mathcal{A}} + x(\Gamma)A) + \lambda x(\Gamma)$$

It is easily checked that the preference relation $\succeq_U$ induced by U satisfies all the desired axioms, and that because $U(\eta A' + (1 - \eta)X) = u(\eta A' + (1 - \eta)A) + \lambda \eta \neq u(A)$, we conclude that $\succeq_U \neq \succeq$. This completes the construction in Case (iii).

In each case we have constructed a preference relation that extends $\succeq_0$, satisfies all of our axioms and differs from $\succeq$; therefore the proof is complete. ∎

4   Testable Implications

This section provides a bridge between the general theory described above and our experiment, which is designed to test the implications of the theory. We first describe the choice domains in the experimental design and their theoretical properties. Then we develop a number of theoretical results in this setting that are testable on the basis of observed choices.

It is convenient to isolate the argument for sufficiency in Theorem 1 and extend it to a setting in which we consider only certain lotteries. To this end, fix a non-empty set Π of probability vectors. For each non-empty subset $\Theta \subset \omega$, let $L_{\Pi}(\Theta)$ be the set of lotteries of the form $p_1\theta_1 + \ldots + p_k\theta_k$, where $(p_1, \ldots, p_k) \in \Pi$ and $\theta_1, \ldots, \theta_k \in \Theta$. In view of the Sure Thing Principle, we may identify the lottery $p_1\theta + \ldots + p_k\theta$ with θ itself; thus $\Theta \subset L_{\Pi}(\Theta)$. If Π is the set of all probability vectors then $L_{\Pi}(\Theta) = L(\Theta)$; in particular, $L_{\Pi}(P) = L(P)$ and $L_{\Pi}(\Omega) = L(\Omega)$. Recall that observing $\succeq_0$ means observing the restrictions of $\succeq$ to Ω and to $L(P)$. Hence the following proposition generalizes the sufficient condition in Theorem 1.

Proposition 2 *Let Π be a non-empty set of probability vectors and let $\succeq$ be a preference relation on $L_{\Pi}(\Omega)$ that satisfies Completeness, Transitivity, Continuity, the Sure Thing Principle and State Monotonicity. In order that $\succeq$ can be deduced from its restrictions $\succeq_{\Omega}$ to Ω and $\succeq_{L_{\Pi}(P)}$ to $L_{\Pi}(P)$, it is sufficient that the $\mathcal{DM}$ finds social state $\omega \in \Omega \setminus P$ to be indifferent to some personal state $\tilde{\omega} \in P$.*

Proof. By assumption, for every social state there is some personal state $\tilde{\omega} \in P$ for which $\omega \sim_{\Omega} \tilde{\omega}$, which we do not require to be unique. State Monotonicity implies that if $(p_1, \ldots, p_k) \in \Pi$ and $\omega_1, \ldots, \omega_k \in \omega$ then

$$\sum p_i\omega_i \sim \sum p_i\tilde{\omega}_i$$

Hence given two lotteries $\sum p_i \omega_i, \sum q_j \omega_j \in L_\Pi(\Omega)$ it follows from State Monotonicity and Transitivity that

$$
\begin{array}{ccc}
\sum p_i \omega_i & \succeq & \sum q_j \omega_j \\
& \Updownarrow & \\
\sum p_i \tilde{\omega}_i & \succeq & \sum q_j \tilde{\omega}_j \\
& \Updownarrow & \\
\sum p_i \tilde{\omega}_i & \succeq_{L_\Pi(P)} & \sum q_j \tilde{\omega}_j
\end{array}
$$

That is, we can deduce $\succeq$ from $\succeq_\Omega$ and $\succeq_{L_\Pi(P)}$, as asserted. ∎

In the experiment there is a subject *self* (the $\mathcal{DM}$) and an (unknown) *other*. The set of social states Ω consists of *monetary* payout pairs (a,b), where $b \geq 0$ is the payout for *self* and $a \geq 0$ is the payout for *other*. However, the set of lotteries we can present to (human) subjects is limited; in fact we consider only *equiprobable binary* lotteries: $\frac{1}{2}(a,b) + \frac{1}{2}(c,d)$. In the framework of the Proposition 2, $\Pi = \{(\frac{1}{2}, \frac{1}{2})\}$. To simplify notation, let $\mathbb{L} = L_\Pi(\Omega)$ be the set of all such equiprobable binary lotteries.[11] Within $\mathbb{L}$ we distinguish three subsets:

$$
\begin{aligned}
\mathcal{PR} &= \{\tfrac{1}{2}(0,b) + \tfrac{1}{2}(0,d)\} = \mathbb{L}(P) \\
\mathcal{SC} &= \{\tfrac{1}{2}(a,b) + \tfrac{1}{2}(a,b)\} = \Omega \\
\mathcal{SR} &= \{\tfrac{1}{2}(a,b) + \tfrac{1}{2}(b,a)\} = \mathbb{L}(\text{Perm}(\Omega))
\end{aligned}
$$

where $\text{Perm}(\Omega)$ is the set of permutations of Ω. Notice that $\mathbb{L}$ is a 4-dimensional convex cone, which is is not easily presented to experimental subjects—but each of the three subsets $\mathcal{PR}, \mathcal{SC}, \mathcal{SR}$ is a 2-dimensional convex sub-cone and can be easily presented using our graphical experimental interface (details are below).

We can interpret choice in each domain—PERSONAL RISK, SOCIAL CHOICE, SOCIAL RISK—as choice in the corresponding subset above by making an obvious identification:

$$
\begin{array}{ll}
\text{PERSONAL RISK} & \langle x, y \rangle \mapsto \tfrac{1}{2}(0,x) + \tfrac{1}{2}(0,y) \\
\text{SOCIAL CHOICE} & (x,y) \mapsto \tfrac{1}{2}(x,y) + \tfrac{1}{2}(x,y) \\
\text{SOCIAL RISK} & [x,y] \mapsto \tfrac{1}{2}(x,y) + \tfrac{1}{2}(y,x).
\end{array}
$$

Hence in the PERSONAL RISK domain, the objects of choice are equiprobable personal lotteries (*other* receives nothing); in the SOCIAL CHOICE domain the objects of choice are deterministic payout pairs for *self* and *other* ; in the SOCIAL RISK domain, the objects of

[11] Fudenberg and Levine (2012) also study preference relations on the set $\mathbb{L}$ of equiprobable binary lotteries but their purpose is quite different: they are primarily interested in social fairness. They show that familiar theories of social fairness—Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Andreoni and Miller (2002), and Cox et al. (2008)—that are defined over riskless outcomes cannot be extended to lotteries—even equiprobable binary lotteries without violating either the Independence Axiom or suggested notions of fairness over risky outcomes.

choice are (in view the Sure Thing Principle) equiprobable social lotteries over symmetric pairs of payouts for *self* and for *other*. We have used the different brackets $\langle x, y \rangle$, (x, y), $[x, y]$ as a reminder that we are thinking of the pair x, y as representing a personal lottery in $\mathcal{PR}$, a social state in $\mathcal{SC}$, a social lottery in $\mathcal{SR}$, respectively.

Let $\succeq_{\mathbb{L}}$ be a preference relation on $\mathbb{L}$ and write $\succeq_{\mathcal{PR}}$, $\succeq_{\mathcal{SC}}$, $\succeq_{\mathcal{SR}}$ for its restrictions (sub-preference relations) to $\mathcal{PR}$, $\mathcal{SC}$, $\mathcal{SR}$, respectively. The restriction $\succeq_{\mathcal{PR}}$ of $\succeq_{\mathbb{L}}$ to $\mathcal{PR}$ prescribes preferences over personal lotteries and the restriction $\succeq_{\mathcal{SC}}$ of $\succeq_{\mathbb{L}}$ to $\mathcal{SC}$ prescribes preferences over social states. Thus to observe $\succeq_0$ in this setting is *exactly* to observe both $\succeq_{\mathcal{PR}}$ and $\succeq_{\mathcal{SC}}$ so Proposition 2 provides a sufficient condition that $\succeq_{\mathbb{L}}$. In particular, we can deduce the restriction $\succeq_{\mathcal{SR}}$ to $\mathcal{SR}$ from $\succeq_{\mathcal{PR}}$ and $\succeq_{\mathcal{SC}}$. That is, if we observe $\succeq_{\mathcal{PR}}$ and $\succeq_{\mathcal{SC}}$, and if every social state is indifferent to some personal state (a condition that is determined completely by $\succeq_{\mathcal{SC}}$), then we can deduce $\succeq_{\mathcal{SR}}$.

But to test this implication, it is not enough to know just the mere *fact* that every social state is indifferent to some personal state. For each social state we need to know a *particular* personal state to which that social state is indifferent. For some of our subjects, this would require making additional assumptions about the form, parametric or otherwise, of the underlying preferences. However, for two classes of subjects—*selfish* and *impartial*— we can construct a formal nonparametric test. We say that preferences $\succeq_{\mathcal{SC}}$ in the Social Choice domain are *selfish* if $(x, y) \sim_{\mathcal{SC}} (0, y)$ and *impartial* if $(x, y) \sim_{\mathcal{SC}} (y, x)$ for all $(x, y) \in \omega$. For a selfish $\mathcal{DM}$, we test whether choice behavior in the Personal Risk domain *coincides* with choice behavior in the Social Risk domain; For an impartial $\mathcal{DM}$, we test whether that choice behavior in the Social Choice domain *coincides* with choice behavior in the Social Risk domain (so an impartial $\mathcal{DM}$ is immune to social risk).

In our experiments, we present subjects with a sequence of consumer decision problems: selection of a bundle of commodities from a standard budget set. Without essential loss of generality, assume the $\mathcal{DM}$'s budget is normalized to 1. The set of budget lines is then

$$\mathcal{B} = \{(x, y) \in \Omega : p_x x + p_y y = 1\}$$

where $p_x, p_y > 0$; the $\mathcal{DM}$ can choose any allocation $(x, y) \geq 0$ that satisfies the budget constraint. A choice of the allocation (x, y) from the budget line represents an allocation between accounts x, y (corresponding to the usual horizontal and vertical axes). The actual payoffs of a particular choice in a specific domain are determined by the allocation to the x and y accounts, according to the particular domain – Personal Risk, Social Choice, Social Risk. With these preliminaries in hand, we can state the theoretical predictions for selfish preferences and impartial preferences in the experimental setting.

Proposition 3 *If preferences $\succeq_{\mathcal{SC}}$ in the Social Choice domain are selfish then preferences $\succeq_{\mathcal{PR}}$ in the Personal Risk domain coincide with preferences $\succeq_{\mathcal{SR}}$ in the Social Risk domain. In particular: if preferences $\succeq_{\mathcal{SC}}$ in the Social Choice domain are selfish then choice behavior in the Personal Risk domain and choice behavior in the Social Risk domain coincide. So if $B \in \mathcal{B}$ is a budget set then $\langle x, y \rangle \in \arg\max_B(\succeq_{\mathcal{PR}})$ in*

14

the PERSONAL RISK *domain if and only if* $[x, y] \in \arg\max_B(\succeq_{\mathcal{SR}})$ *in the* SOCIAL RISK *domain.*

Proof. $\langle x, y \rangle \in \arg\max_B(\succeq_{\mathcal{PR}})$ in the PERSONAL RISK domain if and only if $\langle x, y \rangle \succeq_{\mathcal{PR}} \langle \hat{x}, \hat{y} \rangle$ for every $\langle \hat{x}, \hat{y} \rangle \in B$. Unwinding the notation, this means that

$$\langle x, y \rangle = \tfrac{1}{2}(0, x) + \tfrac{1}{2}(0, y) \succeq_{\mathcal{PR}} \tfrac{1}{2}(0, \hat{x}) + \tfrac{1}{2}(0, \hat{y}) = \langle \hat{x}, \hat{y} \rangle$$

for every $\langle \hat{x}, \hat{y} \rangle \in B$. If preferences $\succeq_{\mathcal{SC}}$ are selfish $(x, y) \sim_{\mathcal{SC}} (0, y)$ for all (x, y); this is true if and only if

$$[x, y] = \tfrac{1}{2}(y, x) + \tfrac{1}{2}(x, y) \succeq_{\mathcal{SR}} \tfrac{1}{2}(\hat{y}, \hat{x}) + \tfrac{1}{2}(\hat{x}, \hat{y}) = [\hat{x}, \hat{y}]$$

for every $[\hat{x}, \hat{y}] \in B$. We conclude that preferences $\succeq_{\mathcal{PR}}$ in the PERSONAL RISK domain coincide with preferences $\succeq_{\mathcal{SR}}$ in the SOCIAL RISK domain. Given that preferences in the two domains coincide, choice behavior coincides as well. ∎

Proposition 4 *If preferences* $\succeq_{\mathcal{SC}}$ *in the* SOCIAL CHOICE *domain are impartial then preferences* $\succeq_{\mathcal{SC}}$ *in the* SOCIAL CHOICE *domain coincide with preferences* $\succeq_{\mathcal{SR}}$ *in the* SOCIAL RISK *domain. In particular: if preferences* $\succeq_{\mathcal{SC}}$ *in the* SOCIAL CHOICE *domain are impartial then choice behavior in the* SOCIAL CHOICE *domain and choice behavior in the* SOCIAL RISK *domain coincide—so if* $\mathcal{B}$ *is a budget set then* $(x, y) \in \arg\max_B(\succeq_{\mathcal{SC}})$ *in the* SOCIAL CHOICE *domain if and only if* $[x, y] \in \arg\max_B(\succeq_{\mathcal{SR}})$ *in the* SOCIAL RISK *domain.*

Proof. Assume preferences $\succeq_{\mathcal{SC}}$ in the SOCIAL CHOICE domain are impartial. Consider two choices $[x, y]$ and $[\hat{x}, \hat{y}]$ in the SOCIAL RISK domain and suppose $[\hat{x}, \hat{y}] \succ_{\mathcal{SR}} [x, y]$. Expressed explicitly in terms of lotteries, this means

$$\tfrac{1}{2}(\hat{x}, \hat{y}) + \tfrac{1}{2}(\hat{y}, \hat{x}) \succ_{\mathcal{SR}} \tfrac{1}{2}(x, y) + \tfrac{1}{2}(y, x)$$

State Monotonicity implies that either $(\hat{x}, \hat{y}) \succ_{\mathcal{SC}} (x, y)$ or $(\hat{y}, \hat{x}) \succ_{\mathcal{SC}} (y, x)$; Impartiality implies that if either of these is true then both are true. Hence we conclude that if $[\hat{x}, \hat{y}] \succ_{\mathcal{SR}} [x, y]$ in the SOCIAL RISK domain then $(\hat{x}, \hat{y}) \succ_{\mathcal{SC}} (x, y)$ in the SOCIAL CHOICE domain. Conversely if $(\hat{x}, \hat{y}) \succ_{\mathcal{SC}} (x, y)$ in the SOCIAL CHOICE domain then

$$\tfrac{1}{2}(\hat{x}, \hat{y}) + \tfrac{1}{2}(\hat{y}, \hat{x}) \succ_{\mathcal{SR}} \tfrac{1}{2}(x, y) + \tfrac{1}{2}(y, x)$$

That is: $[\hat{x}, \hat{y}] \succ_{\mathcal{SR}} [x, y]$ in the SOCIAL RISK domain. Putting these together we conclude that preferences $\succeq_{\mathcal{SC}}$ in the SOCIAL CHOICE domain coincide with preferences $\succeq_{\mathcal{SR}}$ in the SOCIAL RISK domain. Given that preferences in the two domains coincide, choice behavior coincides as well. ∎

To conclude, selfish subjects find any (x, y) to be indifferent to $(0, y)$ —we say that $(0, y)$ is the *selfish equivalent* of (x, y)— and impartial subjects find any (x, y) to be indifferent

15

to (y, x). Proposition thus 3 states that for selfish subjects preferences $\succeq_{\mathcal{SR}}$ in the Social Risk domain coincide with preferences $\succeq_{\mathcal{PR}}$ in the Personal Risk domain. Proposition 4 states that for impartial subjects preferences $\succeq_{\mathcal{SR}}$ in the Social Risk domain are immune to risk so these preferences coincide with the preferences $\succeq_{\mathcal{SC}}$ in the Social Choice domain.

5 The Experiment

We next describe our experiment, which is designed to test the theory described above. We conducted the experiment at the University of Bergen and NHH Norwegian School of Economics. The 276 subjects in the experiment were recruited from undergraduate classes in these institutions. The full experimental instructions, including the computer program dialog windows, are reproduced in the Appendix.

In our experiment, subjects choose a bundle from a budget line; the subjects can choose any allocation that satisfies this constraint. A choice of the allocation from the budget line represents an allocation between accounts x, y (corresponding to the usual horizontal and vertical axes). These budget lines are presented using the graphical interface introduced by Choi et al. (2007b).[12] Subjects make choices by using the computer mouse to move the pointer on the computer screen to the desired point, and were restricted to allocations on the budget constraint.[13]

The actual payoffs of a particular choice in a particular experimental treatment are determined by the allocation to the x and y accounts, according to the particular domain. In the experiment we consider three treatments, corresponding to the domains discussed above:

- In the Personal Risk treatment *self* (the subject) receives the tokens allocated to one of the accounts x or y, determined at random with equal probability; the tokens allocated to the other account are lost. This treatment involves only risk to *self*—*other* receives nothing—and is identical to the (symmetric) risk experiment of Choi et al. (2007a).

---

[12]Of course it is possible that presenting choice problems graphically biases choice behavior in some particular way, but there is no evidence that this is the case. For instance: behavior in the Social domain elicited graphically (Fisman et al., 2007) is quite consistent with behavior elicited by other means (Camerer, 2003), and behavior in the Personal Risk domain elicited graphically (Choi et al., 2007a) is quite consistent with behavior elicited by other means (Holt and Laury, 2002).

[13]In the two-person dictator experiment of Fisman et al. (2007), choices were not restricted to lie on the budget line. They report that most subjects had no violations of budget balancedness using a narrow confidence interval (those who did violate budget balancedness also had many GARP violations even among the subset of their choices that were on the budget constraint). All future experiments including Choi et al. (2007a, 2014), Ahn et al. (2014) and Fisman et al. (2015a,b, 2017) thus restricted choices to allocations on the budget constraint, which simplified the decision problem and made the computer program easier to use.

- In the SOCIAL CHOICE treatment, *self* receives the tokens allocated to y account, while *other* (an anonymous other subject, chosen at random from the group of other subjects in the experiment) receives the tokense allocated to the x account. This treatment involves only selfishness and altruism, and is identical to the (linear) two-person dictator experiment of Fisman et al. (2007).

- In the SOCIAL RISK treatment, *self* receives the tokens allocated to one of the accounts x or y, determined at random with equal probability; *other* receives the tokens allocated to the other account. This treatment is new: it involves risky social choices (whose consequences are not for *self* alone).

Each experimental subject faced 50 independent decision rounds in each of the three treatments. For each subject, the computer selected 50 budget lines randomly from the set of lines that intersect at least one axis at or above the 50 token level and intersect both axes at or below the 100 token level. Each subject faced exactly the same 50 budget lines in each treatment, but the order of presentation was randomized between treatments. The budget lines selected for each subject in his/her decision problems were independent of each other and of the budget lines selected for other subjects in their decision problems. In the PERSONAL RISK and SOCIAL RISK treatments, subjects were not informed of the account that was actually selected until the end of the experiment. This procedure was repeated until all 50 rounds were completed.[14]

The experimental subjects first faced the SOCIAL RISK treatment (because it is the centerpiece of the analysis). The order of the other experimental treatments—PERSONAL RISK and SOCIAL CHOICE – was counterbalanced across sessions to balance out treatment order effects.[15] At the beginning of the experiment subjects received only general instructions on the experimental procedures and the use of the computer interface. At the beginning of each treatment, subjects received specific instructions for that treatment but not for subsequent treatments. Each part of the experiment ended after all subjects had made all their decisions.

At the end of the experiment, the computer randomly selected one of the 50 decision rounds from each of the three treatments of the experiment to carry out for payoffs. The round selected from each treatment depended solely on chance. In the SOCIAL CHOICE and SOCIAL RISK treatments, each subject then received the tokens that he/she allocated to *self* in the round and the subject with whom he/she was matched received the tokens

---

[14] The x- and y-axes were scaled from 0 to 100 tokens. The resolution compatibility of the budget lines was 0.2 tokens, and the appearance and behavior of the pointer were set to the Windows mouse default. At the beginning of each decision round, the subject was presented with a budget line, with the pointer positioned randomly on the line. At the end of each decision round, the experimental program dialog window went blank, after which the entire setup reappeared for the next decision round.

[15] We also had an OBSERVER treatment where each subject faced the same menu of 50 budget lines representing monetary payoffs for two (anonymous) *others*, but that treatment does not provide testable implications of our theory so we make no use of it here.

that she allocated to *other*. The computer program ensured that no two subjects were ever paired as both *self-other* and *other-self*.[16] Payoffs were calculated in terms of tokens and then converted into money.[17]

6   Data Description

We next provide an overview of the basic features of the experimental data. The experiments provide us with a very rich data set. For each subject we observe a choice from each of 50 budget lines in each of the experimental treatments—PERSONAL RISK, SOCIAL CHOICE   CHOICE, SOCIAL RISK—and this yields a rich data set that is well-suited to analysis at the level of the individual subject without the need to pool data or assume that preferences are identical across subjects. Most importantly, the changes in relative prices are such that budget lines cross frequently. This means that our data lead to high power tests of revealed preference conditions Choi et al. (2007b).

6.1   Aggregate Behavior

In this section, we provide an overview of some important features of the experimental data, which we summarize by reporting the distribution of allocations in a number of ways. The dark gray histogram in Figure 3 depicts the distribution (across individuals) in the SOCIAL CHOICE treatment of the average across all choices of the number of tokens kept by *self* as a fraction of the total of tokens allocated to *self* and *other*; that is, the average across all choices of the fraction $y/(x+y)$. On the horizontal axis we show bins of the average $y/(x+y)$; on the vertical axis we show the fraction of subjects whose average is in each bin. As might be expected, there were very few subjects whose averages are much below the midpoint of 0.5; of our 276 subjects, only six (2.2%) kept on average fewer than 0.45 of the tokens and of these only two kept fewer than 0.4.

[Figure 3 here]

For our purposes, we classify a subject as *selfish* if the average $y/(x+y) > 0.95$ and as *impartial* if the average $y/(x+y) \in (0.45, 0.55)$. By this criteria, of our 276 subjects, 103 (37.3%) are classified as selfish and 19 (6.9%) are classified as impartial. Of the 19 impartial

---

[16]As is customary in social preference experiments (Andreoni and Miller, 2002; Fisman et al., 2007, , among others) each subject received two groups of tokens, one based on his/her own decision to allocate tokens and another based on the decision of another random subject to allocate tokens. A concern with this payout method is that it may create a sense of reciprocity among subjects. But in both Andreoni and Miller (2002) and Fisman et al. (2007) the fraction given to *other* is about 20%, similar to the average reported by Camerer (2003) in a summary of dictator games. The computer arranges pairings between subjects so that if subject i could receive tokens from subject j then subject j could not receive tokens from subject i.

[17]Each token was worth 1.2 Norwegian Krone (NOK) (approximately 0.2 USD). A 100 NOK participation fee and subsequent earnings, which averaged about 270 NOK, were paid in private at the end of the session.

subjects, three subjects allocated all their tokens to *self* when $p_y < p_x$ and to *other* when $p_y > p_x$, which is consistent with utilitarian preferences (with respect to money), and three subjects always made approximately equal allocations, which is consistent with Rawlsian preferences (with respect to money). As the histogram in Figure 3 shows, there is a great deal of heterogeneity among the subjects who are neither selfish nor impartial.

Because the Personal Risk and Social Risk treatments are symmetric (the two accounts x and y were equally likely) and budget lines are drawn from a symmetric distribution, reporting the distribution of the average $y/(x + y)$ would not be very informative. Instead, the light gray histograms in Figure 3 depict the the distributions in the Personal Risk and Social Risk treatments of the fraction of tokens allocated to the *cheaper* account (that is, to x when $p_x < p_y$ and to y otherwise). The distributions are quite similar: both have a mode near the midpoint of 0.5, fall off sharply above the midpoint, and have no observations below the midpoint of 0.5.[18] 41 subjects (14.9%) allocated more than 0.95 of the the (available) tokens to the cheaper account in the Personal Risk treatment; this is consistent with risk neutrality. 30 subjects (10.9%) allocated more than than 0.95 of the tokens to the cheaper account in the the Social Risk treatment; this is consistent with utilitarianism (in money). Only 9 subjects (3.3%) allocated less than 0.55 of the tokens to the cheaper account in the Personal Risk treatment, which is consistent with infinite risk aversion. And only 9 subjects (3.3%) allocated less than 0.55 of the tokens to the cheaper account in the Social Risk treatment, which is consistent with Rawlsianism (in money). Among the remaining subjects, there is considerably heterogeneity in choice behavior in both the Personal Risk and Social Risk treatments.

6.2 Individual Behavior

The aggregated data above tells us little about the choice behavior of individual subjects. To get some idea of the wide range of behavior observed, we display scatterplots of choices of four subjects (see Figure 4). For ease of exposition, we have chosen subjects who we classified as selfish on the basis of their choices in the Social Choice treatment. In these scatterplots, each entry shows the subject's relative demand $y/(x + y)$ at a given log-price ratio $\ln(p_y/p_x)$ in the Personal Risk and Social Risk treatments. (We show all 50 choices for each subject in each of these treatments.) We chose these particular subjects because their behavior corresponds to one of several prototypical preference relations and because their behavior illustrates both the striking regularity *within* subjects and the heterogeneity *across* subjects that is characteristic of all our data. These scatterplots also

---

[18]Because the two accounts are equally likely, any decision to allocate fewer tokens to the cheaper account would violate monotonicity with respect to first-order stochastic dominance (FOSD), which is regarded as an error—a failure to recognize that some allocations yield payoff distributions with unambiguously lower returns (Hadar and Russel, 1969). This principle seems compelling and is generally accepted in decision theory. All the preference relations usually considered satisfy monotonicity with respect to FOSD. Indeed, various theories of choice under uncertainty have been amended explicitly to avoid violations of monotonicity with respect to FOSD (Quiggin, 1990; Wakker, 1993).

demonstrate the sensitivity of decisions to changes in relative prices in terms of token shares in all treatments. The scatterplots for all subjects (in all treatments) are available upon request.

[Figure 4 here]

Because all four of these subjects are selfish in the Social Choice treatment, the prediction of the theory is that their choice behavior in the Personal Risk treatment should coincide with their choice behavior in the Social Risk treatment. This was true for three of the subjects but not the fourth. ID 511 (see Figure 4A) allocated all the tokens to the cheaper account in the Personal Risk and Social Risk treatments; this behavior is consistent with risk neutrality in the Personal Risk treatment and utilitarianism in the Social Risk treatment. ID 635 (see Figure 4B) chose nearly equal expenditures ($p_x x = p_y y$) in the Personal Risk and Social Risk treatments; this behavior is consistent with maximizing the utility function $\log x + \log y$ in both treatments.[19] ID 317 (see Figure 4C) allocated all the tokens to the cheaper account for extreme price ratios but chose equal allocations for intermediate price ratios. In the Personal Risk treatment, this subject seems to be 'switching' between risk neutrality and infinite risk aversion. Each of these behaviors is consistent with Expected Utility, but not with the same underlying felicity function. This subject's choices are suggestive of *disappointment aversion* (Dekel, 1986; Gul, 1991) where the safe allocation $x = y$ is the reference point.[20] Interestingly, this subject displays the same choice behavior in the Social Risk treatment, as the theory predicts. The fourth subject, ID 645, almost always allocated all the tokens to the cheaper account in the Personal Risk treatment but not in the Social Risk treatment (see Figure 4D).

As noted, all four of the subjects in Figure 4 are selfish; the first three display the same choice behaviors in the Personal Risk and Social Risk treatments, as the theory predicts; ID 645 does not. Of course, these are special cases, for which the regularities in the data are very clear. Choice behavior is much less clear for many other subjects, and there is no obvious taxonomy that allows us to classify all subjects unambiguously. Furthermore, an inspection of scatterplots cannot provide an adequate test of the theory for most subjects. This is the purpose of our individual-level revealed preference tests below.

---

[19]No selfish subject made nearly equal allocations in the Personal Risk and Social Risk treatments; this behavior would be consistent with infinite risk aversion in the Personal Risk treatment and Rawlsianism in the Social Risk treatment.

[20]The utility function in Gul (1991) takes the form $\min\{\alpha u(x) + u(y), u(x) + \alpha u(y)\}$, where $\alpha \geq 1$ is a parameter measuring disappointment aversion where the safe allocation $x = y$ is taken to be the reference point. If $\alpha > 1$ there is a kink at the 45-degree line, which corresponds to an allocation with a certain payoff. Expected Utility is the special case when $\alpha = 1$. See Choi et al. (2007a) for more information on this representation.

6.3 Testing Rationality

Because subjects' consistency (or lack of it) *within* a treatment must be taken into account testing for consistency *across* treatments, we begin by measuring the extent to which subjects' behavior in each of the three treatments is consistent with utility maximization. Afriat's (1967) Theorem tells us that a finite number of individual choices can be rationalized by a well-behaved utility function if and only if the data satisfies the Generalized Axiom of Revealed Preference (GARP).[21] Because our subjects make choices in a wide range of budget sets, our data provide a strong test of utility maximization.

Let $\{(\mathbf{p}^i, \mathbf{x}^i)\}_{i=1}^{50}$ be the data generated by some individual's choices, where $\mathbf{p}^i$ denotes the i-th observation of the price vector and $\mathbf{x}^i$ denotes the associated allocation.[22] An allocation $\mathbf{x}^i$ is *directly revealed preferred* to $\mathbf{x}^j$ denoted $\mathbf{x}^i R^D \mathbf{x}^j$ if $\mathbf{p}^i \mathbf{x}^i \geq \mathbf{p}^i \mathbf{x}^j$ and *strictly directly revealed preferred* if the inequality is strict. The relation *indirectly revealed preferred* denoted $\mathbf{x}^i R \mathbf{x}^j$ is the transitive closure of the directly revealed preferred relation. GARP requires that if $\mathbf{x}^i$ is indirectly revealed preferred to $\mathbf{x}^j$, then $\mathbf{x}^j$ is not strictly directly revealed preferred to $\mathbf{x}^i$.

We assess how well individual choice behavior complies with GARP by using Afriat's (1972) Critical Cost Efficiency Index (CCEI), which measures the fraction by which each budget constraint must be tightened in order to remove all violations of GARP. Formally: for $0 \leq e \leq 1$, define the direct revealed preference relation $R^D(e)$ as

$$\mathbf{x}^i R^D(e) \mathbf{x}^j \iff e\mathbf{p}^i \mathbf{x}^i \geq \mathbf{p}^i \mathbf{x}^j,$$

and define $R(e)$ to be the transitive closure of $R^D(e)$. The CCEI is the largest value of e such that the relation $R(e)$ satisfies GARP. By definition, the CCEI is between 0 and 1; and index closer to 1 means the data are closer to perfect consistency with GARP and hence to perfect consistency with utility maximization.

Mean CCEI's across all subjects are 0.959, 0.951, and 0.901 in the PERSONAL RISK, SOCIAL CHOICE and SOCIAL RISK treatments, respectively. Figure 5 depicts the distributions of CCEI scores in our three treatments. The horizontal axis presents bins of CCEI ranges; the vertical axis indicates the percent of subjects whose CCEI is in each bin. The fact that for most subjects, choices are sufficiently consistent to be considered utility-generated in all three treatments is a striking result in its own right (more below). Nevertheless, the distribution of CCEI scores is generally further to the left for the SOCIAL RISK treatment. This might be expected, because the SOCIAL RISK treatment seems more complicated and less familiar than the other treatments. Of our 276 subjects, the CCEI scores of 248 (89.9%) and 237 (85.7%) subjects were above 0.90 in the PERSONAL RISK

---

[21] See Varian (1982, 1983) for more details. For excellent overviews of the literature, see Chambers and Echenique (2016) and the papers by Afriat (2012), Diewert (2012), Varian (2012) and Vermeulen (2012) published in a special volume of the *Economic Journal* on the foundations of Revealed Preference.

[22] The data generated by an individual's choices are $\left\{\left(\bar{x}^i, \bar{y}^i, x^i, y^i\right)\right\}_{i=1}^{50}$, where $(\bar{x}^i, \bar{y}^i)$ are the endpoints of the budget line. Thus, the i'th budget line is given by $\frac{x^i}{\bar{x}^i} + \frac{y^i}{\bar{y}^i} = 1$.

and Social Choice treatments, respectively, while only 158 (59.2%) were as high in the Social Risk treatment.

[Figure 5 here]

We interpret the CCEI scores as confirmation that subject choices are generally consistent with utility maximization but there is no natural threshold for determining whether subjects are close enough to satisfying GARP. To provide additional evidence, we follow Bronars (1987), which builds on Becker (1962), and compare the behavior of our actual subjects to the behavior of simulated subjects who randomize uniformly on each budget line. Mean CCEI's for 100,000 simulated subjects are only 0.585. Figure 6 compares the distributions of the minimum and maximum CCEI scores in the three treatments for the actual subjects to the distribution of the CCEI scores generated by the simulated subjects. This provides a clear graphical illustration of the extent to which subjects did worse than choosing consistently and the extent to which they did better than choosing randomly. Of our 276 subjects, 159 (57.6%) subjects have a minimum CCEI score above 0.90, while only a very few simulated subjects have CCEI's that high.[23]

[Figure 6 here]

We refer the interested reader to Choi et al. (2007b, 2014) for more details on the use of GARP to test for consistency and a discussion of various alternative measures that have been proposed for this purpose by Varian (1990, 1991), Echenique et al. (2011) and Houtman and Maks (1985).[24] The subjects' CCEI scores, and the alternative consistency scores, in the three treatments are available from the authors upon request. In practice, all indices yield similar conclusions.

---

[23]The Bronars (1987) test rules out the possibility that consistency is the accidental result of random behavior, but it cannot tell whether utility maximization is the correct model. To this end, Choi et al. (2007b) propose generating a sample of hypothetical subjects who maximize a utility function with an idiosyncratic preference shock that has a logistic distribution. Specifically, the hypothetical subjects implement (with error) the power utility function (commonly employed in the empirical analysis) in the Personal Risk treatment. Fisman et al. (2007) generate a sample of hypothetical subjects who implement (with error) a constant elasticity of substitution (CES) utility function for giving in the Social Choice treatment. Their analysis provides a clear benchmark of the extent to which subjects do worse than choosing consistently and the extent to which they do better than different levels of bounded rationality, and demonstrates that if utility maximization is not in fact the correct model, then our experiment is powerful enough to detect it.

[24]Varian (1990, 1991) refined Afriat's CCEI to provide a score that reflects the minimum adjustment required to eliminate the violations of GARP associated with each budget constraint. The score of Echenique et al. (2011) is based on the idea that an individual who violates GARP can be exploited as a "money pump." The discrepancies between the CCEI and the Varian (1990, 1991) score and the money pump score are discussed in Echenique et al. (2011). Houtman and Maks (1985) finds the largest subset of choices that is consistent with GARP.

7   Testing the Theory

As explained in Section 4, to test the predictions of our theory for a given subject, we need not only to *know* that the subject finds each social state to be indifferent to a personal state, but also to *identify* which social states are found indifferent to which personal state. For some of our subjects, this would require postulating a parametric form for the underlying utility function. However, as we have shown, for *selfish* and for *impartial* subjects the predictions of our theory *are* testable:

- If the subject's preferences $\succeq_{\mathcal{SC}}$ in the Social Choice domain are selfish then preferences in the Personal Risk and in the Social Risk domains coincide: $\succeq_{\mathcal{PR}}=\succeq_{\mathcal{SR}}$ (Proposition 3).

- If preferences $\succeq_{\mathcal{SC}}$ in the Social Choice domain are impartial then preferences in the Social Choice domain and preferences in the Social Risk domain coincide: $\succeq_{\mathcal{SC}}=\succeq_{\mathcal{SR}}$ (Proposition 4).

Of our 276 subjects, we classify 103 (37.3%) as selfish and 19 (6.9%) as impartial. Among the selfish subjects there is substantial heterogeneity in choice behaviors in the Personal Risk and in the Social Risk treatments, and among the impartial subjects there is substantial heterogeneity in choice behaviors in the Social Choice and in the Social Risk treatments; this facilitates a serious test of the implications of the theory for these subjects (more below).

To test whether preferences, and hence choice behavior, in two domains coincide, it might seem that the "obvious" approach would be to compare the two choices (one in each domain) from each budget line. However, such an obvious approach obscures the obvious truth. Different choices might arise from different preferences—but they might also arise from indifference; there is no reason to believe optimal choices are unique. And even if we were to assume that optimal choices are unique, different choices might arise from different realizations of errors. Moreover, simply measuring the distance between the two choices on a given budget line will not do, because there is no clear notion of how far apart choices should be to be regarded as different.

An alternative approach would be to impose parametric forms for the underlying utility functions in the different domains, derive the associated demand functions, fit these to the data, and test to see if they conform to the special restrictions imposed by the theory. The inherent shortcomings of this approach are precisely that it *is* parametric: the utility functions postulated must be good approximations of the "true" underlying preferences (a hypothesis that is not directly testable) and the conclusions will be sensitive to the functional forms, the estimation technique, and the manner in which the error term is introduced.

Instead we create an individual-level nonparametric *permutation* test (Good, 2005). This approach builds on the revealed preference techniques used above to test the consis-

tency of choice behavior *within* each treatment to test for consistency of choice behaviors *across* treatments. It is *nonparametric*, making no assumptions about the form of the subject's underlying utility functions in the three treatments—Personal Risk, Social Choice and Social Risk—and allowing for the reality that subjects' behavior is not *perfectly* consistent with utility maximization.

The basis of our test is the following observation: If a subject has a complete and transitive preference ordering $\succeq_{\mathcal{I}}$ in some treatment $\mathcal{I}$ then the set of choices in that treatment should satisfy GARP. If preferences $\succeq_{\mathcal{I}}$ and $\succeq_{\mathcal{J}}$ in the treatments $\mathcal{I}$ and $\mathcal{J}$ (respectively) are the same—that is, $\succeq_{\mathcal{I}}=\succeq_{\mathcal{J}}$—then the union of the sets of choices in these two treatments (and *a fortiori*, any 50-element subset of this union) should also satisfy GARP. However, our actual test cannot be so simple because, as we have already noted, for many subjects, choices in a given treatment do not satisfy GARP *exactly*, so we should certainly not expect that choices across two treatments should satisfy GARP *exactly*. Instead, we view the observed choice from a given budget line $B \in \mathcal{B}$ in treatment $\mathcal{I}$ as a random draw from some distribution function $F_{\mathcal{I}}^B$ over all allocations that satisfy the budget constraint. If preferences are the same in the two treatments $\mathcal{I}$ and $\mathcal{J}$, then choices in the two treatments should be independent draws from the *same* distribution function—that is, $F_{\mathcal{I}}^B = F_{\mathcal{J}}^B$. (Recall that subjects see the *same* 50 budget lines in each treatment.) This is the null hypothesis we test.

Formally, let $\{(\mathbf{p}^i, \mathbf{x}_{\mathcal{I}}^i\}_{i=1}^{50}$ and $\{(\mathbf{p}^i, \mathbf{x}_{\mathcal{J}}^i)\}_{i=1}^{50}$ be the data generated by some individual's choices in the two treatments $\mathcal{I}$ and $\mathcal{J}$, where $\mathbf{p}^i$ denotes the i-th observation of the price vector and $\mathbf{x}_{\mathcal{I}}^i$ and $\mathbf{x}_{\mathcal{J}}^i$ denote the associated choices in treatments $\mathcal{I}$ and $\mathcal{J}$, respectively. There are 2^{50} possible distinct 50-element subsets of this union; each such subset is formed by drawing the choice from treatment $\mathcal{I}$ or $\mathcal{J}$ for each of the 50 budget lines $\mathbf{p}^i$. Clearly we cannot examine all 2^{50} possible subsets; instead we draw 10,000 subsets at random, where each draw is made independently and with equal probability from the choice in $\mathcal{I}$ or from the choice $\mathcal{J}$. Note that the actual datasets from treatments $\mathcal{I}$ and $\mathcal{J}$ are simply the particular realizations in which each choice happened to be drawn from the same treatment.

Similarly, the actual CCEI scores in the two treatments $\mathcal{I}$ and $\mathcal{J}$, denoted by $e_{\mathcal{I}}$ and $e_{\mathcal{J}}$ respectively, are simply realizations from the distribution of CCEI scores calculated for each of the permuted datasets $\{(\mathbf{p}^i, \mathbf{x}^i)\}_{i=1}^{50}$ we draw. If $e_{\mathcal{I}}, e_{\mathcal{J}} = 1$ (so actual choices within each treatment $\mathcal{I}$ and $\mathcal{J}$ are perfectly consistent), we should expect the CCEI of scores of the permuted data sets to be equal to 1—or at least very close; if these scores are substantially below 1 then we should reject the null that preferences in two treatments $\mathcal{I}$ and $\mathcal{J}$ coincide—that is, $\succeq_{\mathcal{I}}=\succeq_{\mathcal{J}}$.[25]

To obtain a distribution function F for the test statistic under the null hypothesis, for

---

[25]In the permutation test, we focus on the CCEI, which offers a straightforward interpretation. Calculating the CCEI or any of the other scores we have mentioned—Varian (1990, 1991), Echenique et al. (2011) and Houtman and Maks (1985)—for each subject for each of our 10,000 permuted datasets is computationally intensive, and would be entirely impractical if, roughly speaking, there were a large number of GARP violations.

each subject we randomly draw 10,000 datasets $\{(\mathbf{p}^i, \mathbf{x}^i)\}_{i=1}^{50}$, calculate the CCEI score for each of these datasets, and compare the distribution of CCEI scores to the actual CCEI scores $e_{\mathcal{I}}$ and $e_{\mathcal{J}}$. Set $e^- = \min\{e_{\mathcal{I}}, e_{\mathcal{J}}\}$ and $e^+ = \max\{e_{\mathcal{I}}, e_{\mathcal{J}}\}$ and let p^- and p^+ be the corresponding p-values. Under the null,

$$p^- = (1 - \widehat{F}^-(t - \epsilon))^2 \text{ and } p^+ = 1 - (\widehat{F}^+(t - \epsilon))^2$$

where F^- (resp. F^+) is the distribution function of the minimum (resp. maximum) of two draws from the permutation distribution function F, $\widehat{F}$ is the estimate of the permutation distribution function F and $\epsilon > 0$ is small (introduced to account for discrete bunching in the permuted CCEI scores illustrated in Figure 7 below). To counteract the problem of multiple comparison, in addition to p^- and p^+ we also use the Bonferroni correction

$$\min\{2\min\{p^-, p^+\}, 1\}$$

as our p-value. This is known to be conservative but difficult to improve on without imposing further structure.

To illustrate our test, we consider the four selfish subjects identified in Figure 7. Each panel presents a histogram of the permuted CCEI scores and the two actual CCEI scores from the PERSONAL RISK and SOCIAL RISK treatments for a subject whose choice behavior in the SOCIAL CHOICE treatment is selfish. For ID 505 (see Figure 7A) $e^- = 1$ and all the permuted CCEI scores are also equal to 1 so we do not reject the null that this subject's preferences in PERSONAL RISK and SOCIAL RISK treatments coincide: $\succeq_{\mathcal{PR}} = \succeq_{\mathcal{SR}}$. However, for ID 729 see Figure 7B), $e^- < e^+ < 1$ and all permuted CCEI scores are below e^- so we do reject the null that $\succeq_{\mathcal{PR}} = \succeq_{\mathcal{SR}}$.

It is more difficult to draw clear conclusions about subjects ID 514 and ID 502 (see Figures 7C and 7D, respectively) because many of the permuted CCEI scores for these subjects lie (weakly) between e^- and e^+. For ID 514, e^+ is sufficiently far into the (right) tail of the permuted CCEI scores that we the Bonferroni correction allows us to reject the null that $\succeq_{\mathcal{PR}} = \succeq_{\mathcal{SR}}$. For ID 502, e^- and e^+ are not extreme with respect to the distribution of the permuted CCEI scores and we cannot reject the null $\succeq_{\mathcal{PR}} = \succeq_{\mathcal{SR}}$. Diagrams for all subjects are available upon request.

[Figures 7 here]

In Table 1, we tabulate the percent of subjects we classify as selfish and impartial based on their choices in the in the SOCIAL CHOICE treatment for whom we can reject null that the preferences coincide—in PERSONAL RISK and SOCIAL RISK $\succeq_{\mathcal{PR}} = \succeq_{\mathcal{SR}}$ for the selfish subjects and in SOCIAL CHOICE and SOCIAL RISK $\succeq_{\mathcal{SC}} = \succeq_{\mathcal{SR}}$ for the impartial subjects. We present the results using the Bonferroni correction, p^- and p^+ for the conventional levels 1% (top line), 5% (middle line) and 10% (bottom line) levels. We conclude that for the large majority of selfish and impartial subjects the theoretical predictions are well supported by the experimental data.

Finally, we generate a benchmark with which we can compare our finding that, for most selfish and impartial subjects, the preferences coincide, exactly as Propositions 3 and 4 predict. We focus on Proposition 3 and add noise to the actual choices of our 103 (37.3%) selfish subjects. Specifically, we assign a probability μ of replacement, and for each choice in each of the 10,000 randomly drawn datasets from the PERSONAL RISK and the SOCIAL RISK treatments, with probability μ we replace the actual choice with a choice drawn independently from the uniform distribution over all allocations on the budget line that allocate more tokens to the cheaper account.[26]

We then calculate the CCEI score for each of these datasets and retest, using the Bonferroni correction. to see whether $\succeq_{\mathcal{PR}}=\succeq_{\mathcal{SR}}$, as the theory predicts for selfish subjects. Rejecting the null that $\succeq_{\mathcal{PR}}=\succeq_{\mathcal{SR}}$ when we replace only a small fraction of actual choices with random choices will demonstrate that the experiment is sufficiently powerful to detect if our theory is not in fact the correct model. Table 2 report the fraction of selfish subjects for whom we reject the null that $\succeq_{\mathcal{PR}}=\succeq_{\mathcal{SR}}$ at the 1% (top panel) 5% (middle panel) and 10% (bottom panel) levels when we replace the actual choices with random choices with probabilities $\mu = 0, 0.05.0.1, 0.15, 0.2$. (Keep in mind that $\mu = 0$ means we are using actual choices.) The results show there is a very high probability that we reject the null even when a few individual choices are replaced with random choices. A similar test of Proposition 4 with the impartial subjects yields the same conclusion.

8 Observing More

In our main theoretical result—and in the experiment—we assumed that we can observe the restrictions $\succeq_0$ of $\succeq$ to Ω and to $L(P)$. That is, we observe the $\mathcal{DM}$'s comparisons between social states and the $\mathcal{DM}$'s comparisons between personal lotteries – but not the $\mathcal{DM}$'s comparisons between social states and personal lotteries. For completeness, we now discuss the setting in which we can observe the restriction $\succeq_1$ of $\succeq$ to $\Omega \cup L(P)$. That is, in

---

[26]In the PERSONAL RISK andSOCIAL RISK treatments, any decision to allocate fewer tokens to the cheaper account is a violation of monotonicity with respect to first-order stochastic dominance (FOSD) because there are other feasible allocations that yield unambiguously higher monetary payoffs. FOSD (Hadar and Russel, 1969) is compelling and generally accepted in decision theory. Overall, the choices made by subjects in our experiment show very low rates of FOSD violations, so we restrict the random choices to allocations that allocate more tokens to the cheaper account (positions on the longer side of the budget line relative to the 45-degree line). Note that any choice that allocates fewer tokens to the cheaper account would violate monotonicity with respect to FOSD but need not lead to a violation of GARP, whereas any choice to allocate more tokens to the cheaper account would not violate monotonicity with respect to FOSD. See Choi et al. (2014) for more information on testing for consistency *only* with FOSD and Polisson et al. (Forthcoming) for testing for consistency with GARP *and* FOSD.

addition to observing the $\mathcal{DM}$'s comparisons between social states and between personal lotteries, we *also* observe the $\mathcal{DM}$'s between social states and personal lotteries.

Perhaps surprisingly, Theorem 5 below shows that (in the absence of strong assumptions about the $\mathcal{DM}$'s preferences) observing comparisons between social states and personal lotteries is no help at all: exactly as with $\succeq_0$, in order to deduce the complete preference relation $\succeq$ from the incomplete relation $\succeq_1$ it is necessary and sufficient that the $\mathcal{DM}$ finds every social state to be indifferent to some personal state.

However, if we are willing to make the very strong assumption that the $\mathcal{DM}$'s preferences $\succeq$ obey the Independence Axiom—and hence have an Expected Utility representation—then observing comparisons between social states and personal lotteries makes a big difference: in order to deduce the complete preference relation $\succeq$ from the incomplete relation $\succeq_1$ it is necessary and sufficient only that the $\mathcal{DM}$ finds every social state to be indifferent to some personal *lottery*. (In light of the Continuity Axiom, this is equivalent to the $\mathcal{DM}$ finding every social state to be ranked (weakly) between two personal states.)

Theorem 5 *Assume that the $\mathcal{DM}$'s preference relation $\succeq$ satisfies Completeness, Transitivity, Continuity, Reduction of Compound Lotteries, the Sure Thing Principle and State Monotonicity. In order that $\succeq$ can be deduced from $\succeq_1$ it is necessary and sufficient that the $\mathcal{DM}$ finds every social state $\omega \in \Omega \setminus P$ to be indifferent to some personal state $\tilde{\omega} \in P$..*

Proof. In view of Theorem 1, this condition is sufficient that $\succeq$ can be deduced from $\succeq_0$, so it is certainly sufficient that $\succeq$ can be deduced from $\succeq_1$, which provides more information. To see that this condition is necessary we use the same argument as in the proof of Theorem 1, but with a twist. We suppose that there is some social state X that the $\mathcal{DM}$ does *not* find indifferent to any personal state and construct a preference relation that agrees with $\succeq$ on $L(P) \cup \omega$ but not on all of $L(\Omega)$.

As in the proof of Theorem 1, we will need to ensure that the preference relation we construct obeys Continuity and State Monotonicity. As before, we use Debreu's representation theorem (Debreu, 1954) to find a utility function $u : L(\Omega) \to \mathbb{R}$ that represents $\succeq$. Without loss of generality, assume that the range of u is contained in the interval $[0, 1]$. We construct a new utility function $U : L(\Omega) \to [0, 1]$ that agrees with u on $L(P) \cup \omega$—and hence induces the same preference ordering as $\succeq_1$ on $L(P) \cup \omega$—but does not induce the same ordering as $\succeq$ on $L(\Omega)$.

As before, define $\mathcal{A}, \mathcal{B}, \mathcal{X}$ to be the sets of social states that are strictly preferred to X, strictly dis-preferred to X and indifferent to X, respectively. If $\mathcal{A} \neq \emptyset$, let $A \in \mathcal{A}$ be a minimal element; if $\mathcal{B} \neq \emptyset$, let $B \in \mathcal{B}$ be a maximal element. For each lottery $\Gamma = \sum p_i \omega_i \in L(\Omega)$ define $\Gamma_{\mathcal{A}}$, $\Gamma_{\mathcal{B}}$ and $\Gamma_{\mathcal{X}}$ exactly as before and note, as before, that $\Gamma = \Gamma_{\mathcal{A}} + \Gamma_{\mathcal{B}} + \Gamma_{\mathcal{X}}$. An as before, set $x(\Gamma) = \sum_{\omega_i \in \mathcal{X}} p_i$

We distinguish the same three cases as in Theorem 1; the arguments for Case (ii) and Case (iii) are identical to those in the proof of Theorem 1, but the argument for Case (i) requires a twist because we must be careful to construct the utility function U to preserve

the relationship between the social state X (and all those social states in $\mathcal{X}$ that are indifferent to X) and personal *lotteries*, as well as personal *states*. (No additional care is required in Case (ii) because X is preferred to all personal lotteries, or in Case (iii) because X is dis-preferred to all personal lotteries.)

Suppose therefore that neither $\mathcal{A}$ nor $\mathcal{B}$ is empty. Use Continuity to choose $\gamma, \zeta \in (0, 1)$ such that

$$X \sim \gamma A + (1 - \gamma)B$$

and

$$(1/2)A + (1/2)X \sim \zeta A + (1 - \zeta)B.$$

As before, define auxiliary functions $f, g : L(\Omega) \to \mathbb{R}$ by

$$f(\Gamma) = u(\Gamma_{\mathcal{A}} + x(\Gamma)A + \Gamma_{\mathcal{B}})$$
$$g(\Gamma) = u(\Gamma_{\mathcal{A}} + x(\Gamma)B + \Gamma_{\mathcal{B}})$$

We are no longer free to choose U to be an arbitrary convex combination $U = \lambda f + (1 - \lambda)g$ because we require

$$U(X) = U(\gamma A + (1 - \gamma)B) = u(\gamma A + (1 - \gamma)B),$$

which would completely determine λ—and for this λ it might happen that the preference relation $\succeq_U$ coincides with $\succeq$. To get around this, we need an additional parameter. For each $\alpha, \beta \in (0, 1)$ define the function $H_{\alpha,\lambda} : [0, 1] \times [0, 1] \to [0, 1]$ by

$$H_{\alpha,\beta}(s, t) = \tfrac{1}{2} \left[\alpha s^2 + (1 - \alpha)t^2 \right]^{1/2} + \tfrac{1}{2} \left[\beta s + (1 - \beta)t \right]$$

and define a utility function $U_{\alpha,\beta} : L(\Omega) \to [0, 1]$ by

$$U_{\alpha,\beta}(\Gamma) = H_{\alpha,\beta} \left(f(\Gamma), g(\Gamma) \right)$$

Finally, let $\succeq_{\alpha,\beta}$ be the preference relation on $L(\Omega)$ induced by $U_{\alpha,\beta}$. Because f, g and $H_{\alpha,\beta}$ are all continuous, so is $U_{\alpha,\beta}$; hence $\succeq_{\alpha,\beta}$ satisfies the Continuity Axiom. Note that $H_{\alpha,\beta}$ is the identity on the diagonal $H_{\alpha,\beta}(s, s) = s$ so that $U_{\alpha,\beta}(\Gamma) = u(\Gamma)$ for every $\Gamma \in L(\mathcal{A} \cup \mathcal{B})$; in particular, $U_{\alpha,\beta}(\Gamma) = u(\Gamma)$ for every $\Gamma \in L(P)$. Note also that $H_{\alpha,\beta}(s, t)$ is strictly increasing in s, t separately; it is then easily checked that $\succeq_{\alpha,\beta}$ satisfies State Monotonicity. Finally, because

$$f(X) = u(A) > u(X) = u(\gamma A + (1 - \gamma)B > u(B) = g(X)$$

and

$$f(X) = u(A) > u((1/2)A + (1/2)X) > u((1/2)A + (1/2)B) = g(X) > u(B)$$

and the parameters α, β are independent, we can choose the parameters α, β so that

$$
\begin{aligned}
U_{\alpha,\beta}(X) &= H_{\alpha,\beta}(f(X), g(X)) \\
&= u(X)
\end{aligned}
$$

and

$$
\begin{aligned}
U_{\alpha,\beta}((1/2)A + (1/2)X) &= H_{\alpha,\beta}(u(A), u((1/2)A + (1/2)B)) \\
&\neq u(\zeta A + (1-\zeta)B) \\
&= u((1/2)A + (1/2)X).
\end{aligned}
$$

For such a choice of α, β the preference relation $\succeq_{\alpha,\beta}$ is an extension of $\succeq_1$ that differs from $\succeq$, as desired. ∎

Theorem 6 *Assume that the $\mathcal{DM}$'s preference relation $\succeq$ satisfies Completeness, Transitivity, Continuity, Reduction of Compound Lotteries, the Sure Thing Principle and Independence (and hence admits an Expected Utility representation).. In order that $\succeq$ can be deduced from $\succeq_1$ it is necessary and sufficient that the $\mathcal{DM}$ finds every social state $\omega \in \Omega \setminus P$ to be indifferent to some personal lottery $\sum_i p_i \omega_i \in L(P)$.*

Proof. To see that this condition is sufficient, consider lotteries $\sum p_i \omega_i$ and $\sum q_j \omega_j$. The condition guarantees we can find personal lotteries Φ_i and Ψ_j such that $\omega_i \sim \Phi_i$ and $\omega_j \sim \Psi_j$ for each i, j. Independence guarantees that $\sum p_i \omega_i \sim \sum p_i \Phi_i$ and $\sum q_j \omega_j \sim \sum q_j \Psi_j$ so

$$
\sum p_i \omega_i \succeq \sum q_j \omega_j \Leftrightarrow \sum p_i \Phi_i \succeq \sum q_j \Psi_j \Leftrightarrow \sum p_i \Phi_i \succeq_1 \sum q_j \Psi_j.
$$

Thus $\succeq$ can be deduced from $\succeq_1$.

To see, that this condition is necessary, assume there is a social state $X \in \omega$ that is not indifferent to any personal lottery. If there were personal states $A, B \in P$ such that $A \succeq X \succeq B$ then Continuity would imply that $X \sim \gamma A + (1-\gamma)B$ for some $\gamma \in (0,1)$, which would contradict the assumption that X is not indifferent to any personal lottery. It follows that either $X \succ A$ for all personal states $A \in P$ or $X \prec A$ for all personal states $A \in P$.

Choose a continuous utility function $u : \Omega \to \mathbb{R}$ that yields an Expected Utility representation of $\succeq$ and write Eu for the expected utility extension of u to $L(\Omega)$. Note that the range $Eu(L(P))$ of Eu on $L(P)$ is precisely the convex hull of the range $u(P)$ of u on P; because P is finite, $Eu(L(P))$ is thus a closed interval. Let $f : \mathbb{R} \to \mathbb{R}$ be any strictly increasing function that is the identity on $Eu(L(P))$. Then $f \circ u$ defines a utility function u_f on Ω. write Eu_f for its expected utility extension to $L(\Omega)$.

Because f is the identify on $Eu(L(P))$, Eu_f agrees with Eu on $L(P)$; because f is strictly increasing, the ordering induced by u_f on Ω agrees with the ordering induced by

29

u. Hence the preference ordering $\succeq_f$ represented by Eu_f agrees with $\succeq$ on $[\Omega \cup L(P)] \times [\Omega \cup L(P)]$; that is, $\succeq_f$ is an extension of $\succeq_1$.

Because X does not have a personal lottery equivalent , $u(X) \notin Eu(L(P))$; because $Eu(L(P))$ is an interval this means either $u(X) < u(A)$ for all personal states $A \in P$ or $u(X) > u(A)$ for all personal states $A \in P$. In what follows we treat the case in which $u(X) > u(A)$ for all $A \in P$ (X is preferred to every personal state); the argument in the reverse case is similar and left to the reader.

By assumption there are personal states $A, B \in P$ with $A \succ B$, so that $u(A) > u(B)$. Choose f so that $f(u(X)) \neq u(X)$, and set $\lambda = [u(A) - u(B)]/[u(X) - u(B)]$; note that $\lambda \in (0, 1)$. The expected utility property guarantees that

$$Eu(\lambda X + (1 - \lambda)B) = \lambda u(X) + (1 - \lambda)u(B) = u(A)$$

and also that

$$Eu_f(\lambda X + (1 - \lambda)B) = \lambda u_f(X) + (1 - \lambda)u_f(B) = \lambda u_f(X) + (1 - \lambda)u(B)$$

Because $f(u(X)) \neq u(X)$ we conclude that $Eu_f(\lambda X + (1 - \lambda)B) \neq u(A) = u_f(A)$. Thus, $A \sim \lambda X + (1 - \lambda)B$ but $A \not\sim_f \lambda X + (1 - \lambda)B$; in particular, $\succeq_f \neq \succeq$. This completes the proof. ∎

9 Concluding Remarks

It is often said that private choices should remain private. As Paul Krugman has written "... I'm talking about professional mistakes. The other kinds of mistakes ... are none of your business." This point of view seems reasonable when applied to Krugman, who is not a candidate for a public office. But, to the extent that choices are not mistakes but rather are the consequences of attitudes toward risk and attitudes toward personal risk are indicative of attitudes toward social risk, then this point of view would seem mistaken when applied to candidates for public office. Accordingly, we *should* care about the personal choices of those individuals who might be in a position to make choices that have consequences for others—at least to the extent that those choices involve risk.

From a purely technical point of view, our paper poses a problem in decision theory: under what circumstances is a preference relation over some set of lotteries completely determined by its restriction to a subset of lotteries? Grant et al. (1992), which is closest to the present work, pose the problem in the context of lotteries whose outcomes are commodity bundles and lotteries whose outcomes are monetary payoffs. Given fixed prices for commodities, they seek conditions guaranteeing that preferences over lotteries whose outcomes are commodity bundles are completely determined by the restrictions of those preferences to lotteries whose outcomes are monetary payoffs; the sufficient condition they identify is one we call State Monotonicity (and they call Degenerate Independence).

But because our intent is different from Grant et al. (1992), we pose different questions and face quite different issues. In particular, although prices play a crucial role for Grant et al. (1992) (prices mediate between monetary outcomes and consumption bundles), prices play no role at all in our setting. More subtly, the central issue in our setting is whether all choices in a larger set (social choices) have equivalents (are viewed as indifferent to) choices in a smaller set (personal choices). In Grant et al. (1992) it is *assumed* that all choices in the larger set have equivalents in the smaller set; the central issue is whether this condition is strong enough to determine preferences over lotteries.

There is a large experimental literature on risk and social preferences, which we will not attempt to review here. However, we call attention to three recent papers that seem most relevant to our study: Krawczyk and Le Lec (2010, 2016), Brock et al. (2013) and Cappelen et al. (2013) These papers report experiments designed to examine social preferences in the presence of risk. In particular, these papers seek to understand whether giving behavior is driven by considerations of ex-ante fairness or ex-post fairness. The overall takeaway from these studies is that a theory based solely on ex-ante comparisons or solely on ex-post comparisons cannot fully account for the experimental data.

In political science, a substantial literature argues that personal character is an important predictor of Presidential conduct. The argument is made most famously and forcefully in a classic book *The Presidential Character*, by Barber (1972), who writes: "Character is the force, the motive power, around which the person gathers his view of the world, and from which his style receives its impetus. The issues will change; the character of the president will not." Barber argues in particular that a presidential candidate's character provides "a realistic estimate of what will endure into a man's White House years."[27] The political science literature does not define character but Barber (1972) and others (explicitly or implicitly) argue that character is revealed by personal choices. As Barber puts it "the personal past foreshadows the presidential future."Such an argument would seem coherent—and of use to voters—only if the candidate's attitude with respect to social policy—and in particular toward social risk—after achieving office can be deduced from the candidate's attitude with respect to personal choices – and in particular toward personal risk—before achieving office.[28]

However, it can be dangerously easy to err and infer too much from observations that are too imperfect. During (and after) the 1992 presidential campaign, stories were widely told about Bill Clinton's personal choices, which—entirely aside from its moral content— were surely quite risky, and many pundits—and no doubt many voters—used these stories

---

[27]A number of Presidents agreed with Barber (1972). Richard Nixon, for instance, said "... with all the power that a President has, the most important thing to bear in mind is this: You must not give power to a man unless, above everything else, he has character. Character is the most important qualification the President of the United States can have."

[28]Some voters might care about a candidate's personal choices on purely moral grounds, independent of the implications for choices the candidate might make or policies the candidate might follow when he or she actually assumes office—but that is not the argument being made by Barber (1972) and others. Nelson (2018) provides excellent discussions of the arguments.

as the basis for predictions about his choices in the public domain. Such predictions did not stop with Clinton's election; a 1994 article in Newsweek, for instance, concluded that "...it may well be that this is one case where personal behavior does give an indication of how a politician will perform in the arena."History is yet to write its judgement of that prediction, but voters have already done so: Clinton left office with the highest approval rating of any President in recent history.

References

Afriat, S. N. (1967). "The construction of a utility function from expenditure data," *International Economic Review*, 8(1): 67–77.

Afriat, S. N. (1972). "Efficiency estimation of production functions," *International Economic Review*, 13(3): 568–598.

Afriat, S. N. (2012). "Afriat's theorem and the index number problem," *Economic Journal*, 122(560): 295–304.

Ahn, David, Syngjoo Choi, Douglas Gale, and Shachar Kariv (2014). "Estimating ambiguity aversion in a portfolio choice experiment," *Quantitative Economics*, 5(2): 195–223.

Andreoni, James and John Miller (2002). "Giving according to GARP: An experimental test of the consistency of preferences for altruism," *Econometrica*, 70(2): 737–753.

Barber, James David (1972). *The Presidential Character: Predicting Performance in the White House*, Englewood Cliffs, NJ: Prentice-Hall.

Becker, Gary S. (1962). "Irrational behavior and economic theory," *Journal of Political Economy*, 70(1): 1–13.

Bolton, Gary E. and Axel Ockenfels (2000). "ERC: A theory of equity, reciprocity, and competition," *American Economic Review*, 90(1): 166–193.

Brock, J. Michelle, Andreas Lange, and Erkut Y. Ozbay (2013). "Dictating the risk: Experimental evidence on giving in risky environments," *American Economic Review*, 103(1): 415–437.

Bronars, Stephen G. (1987). "The power of nonparametric tests of preference maximisation," *Econometrica*, 55(3): 693–698.

Camerer, Colin F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Cappelen, Alexander W., James Konow, Erik Ø. Sørensen, and Bertil Tungodden (2013). "Just luck: An experimental study of risk taking and fairness," *American Economic Review*, 103(3): 1398–1413.

Chambers, Christopher P. and Federico Echenique (2016). *Revealed Preference Theory*, Cambridge University Press.

Charness, Gary and Matthew Rabin (2002). "Understanding social preferences with simple tests," *Quarterly Journal of Economics*, 117(3): 817–869.

Chew, S. H. (1989). "Axiomatic utility theories with the betweenness property," *Annals of Operations Research*, 19(1): 273–298.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv (2007a). "Consistency and heterogeneity of individual behavior under uncertainty," *American Economic Review*, 97(5): 1921–1938.

Choi, Syngjoo, Raymond Fisman, Douglas M. Gale, and Shachar Kariv (2007b). "Revealing preferences graphically: An old method gets a new tool kit," *American Economic Review*, 97(2): 153–158.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman (2014). "Who is (more) rational?" *American Economic Review*, 104(6): 1518–1550.

Cox, James C., Daniel Friedman, and Vjollca Sadiraj (2008). "Revealed altruism," *Econometrica*, 76(1): 31–69.

Debreu, Gerard (1954). "Representation of a preference ordering by a numerical function," in R.M. Thrall, C. H. Coombs, and R. L. Davis (eds.), "Decision Processes," Wiley.

Dekel, Eddie (1986). "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 40(2): 304–318.

Diewert, W. Erwin (2012). "Afriat's theorem and some extensions to choice under uncertainty," *Economic Journal*, 122(560): 305–331.

Echenique, Federico, Sangmok Lee, and Matthew Shum (2011). "The money pump as a measure of revealed preference violations," *Journal of Political Economy*, 119(6): 1201–1223.

Fehr, Ernst and Klaus M. Schmidt (1999). "A theory of fairness, competition and cooperation," *Quarterly Journal of Economics*, 114(3): 817–868.

Fisman, Raymond, Pamela Jakiela, and Shachar Kariv (2015a). "How did distributional preferences change during the Great Recession?" *Journal of Public Economics*, 128: 84–95.

Fisman, Raymond, Pamela Jakiela, and Shachar Kariv (2017). "Distributional preferences and political behavior," *Journal of Public Economics*, 155: 1–10.

Fisman, Raymond, Pamela Jakiela, Shachar Kariv, and Daniel Markovits (2015b). "The distributional preferences of an elite," *Science*, 349(6254): aab0096.

Fisman, Raymond J., Shachar Kariv, and Daniel Markovits (2007). "Individual preferences for giving," *American Economic Review*, 97(5): 1858–1876.

Fudenberg, Drew and David K. Levine (2012). "Fairness, risk preferences and independence: Impossibility theorems," *Journal of Economic Behavior and Organization*, 81(2): 606–612.

Good, Phillip (2005). *Permutation, Parametric, and Bootstrap Tests of Hypotheses*, Springer, 3 edition.

Grant, Simon, Atsushi Kajii, and Ben Polak (1992). "Many good choice axioms: When can many-good lotteries be treated as money lotteries?" *Journal of Economic Theory*, 56(2): 313–337.

Gul, Faruk (1991). "A theory of disappointment aversion," *Econometrica*, 59(3): 667–686.

Hadar, Josef and William R. Russel (1969). "Rules for ordering uncertain prospects," *American Economic Review*, 59(1): 25–34.

Halevy, Yoram, Dotan Persitz, and Lanny Zrill (2018). "Parametric recoverability of preferences," *Journal of Political Economy*, 126(4): 1558–1593.

Holt, Charles A. and Susan K. Laury (2002). "Risk aversion and incentive effects," *American Economic Review*, 92(5): 1644–1655.

Houtman, Martijn and J. A. H. Maks (1985). "Determining all maximal data subsets consistent with revealed preference," *Kwantitatieve Methoden*, 6(19): 89–104.

Krawczyk, Michal and Fabrice Le Lec (2010). "'Give me a chance!' An experiment in social decision under risk," *Experimental Economics*, 14(3): 500–511.

Krawczyk, Michal and Fabrice Le Lec (2016). "Dictating the risk: Experimental evidence on giving in risky environments: Comment," *American Economic Review*, 106(3): 836–839.

Li, Jing, William H. Dow, and Shachar Kariv (2017). "Social preferences of future physicians," *Proceedings of the National Academy of Sciences*, 114(48): E10291–E10300.

Nelson, Michael (2018). *The Presidency and the Political System*, Thousand Oaks, CA: CQ Press, 11th edition.

von Neumann, John and Oskar Morgenstern (1947). *Theory of Games and Economic Behavior*, Princeton: Princeton University Press, 2nd edition.

Polisson, Matthew, John K. H. Quah, and Ludovic Renou (Forthcoming). "Revealed preferences over risk and uncertainty," *American Economic Review*.

Quiggin, John (1982). "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 3(4): 323–343.

Quiggin, John (1990). "Stochastic dominance in regret theory," *Review of Economic Studies*, 57(3): 503–511.

Quiggin, John (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*, Kluwer Academic Publishers.

Tversky, Amos and Daniel Kahneman (1992). "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 5(5): 297–323.

Varian, Hal R. (1982). "The nonparametric approach to demand analysis," *Econometrica*, 107(4): 945–972.

Varian, Hal R. (1983). "Non-parametric tests of consumer behavior," *Review of Economic Studies*, 50(1): 99–110.

Varian, Hal R. (1990). "Goodness-of-fit in optimizing models," *Journal of Econometrics*, 46(1-2): 125–140.

Varian, Hal R. (1991). "Goodness of fit for revealed preference tests," Technical Report 13, Michigan – Center for Research on Economic & Social Theory.

Varian, Hal R. (2012). "Revealed preference and its applications," *Economic Journal*, 122(560): 332–338.

Vermeulen, Frederic (2012). "Foundations of revealed preference: Introduction," *Economic Journal*, 122(560): 287–294.

Wakker, Peter (1993). "Savage's axioms usually imply violation of strict stochastic dominance," *Review of Economic Studies*, 60(2): 487–493.
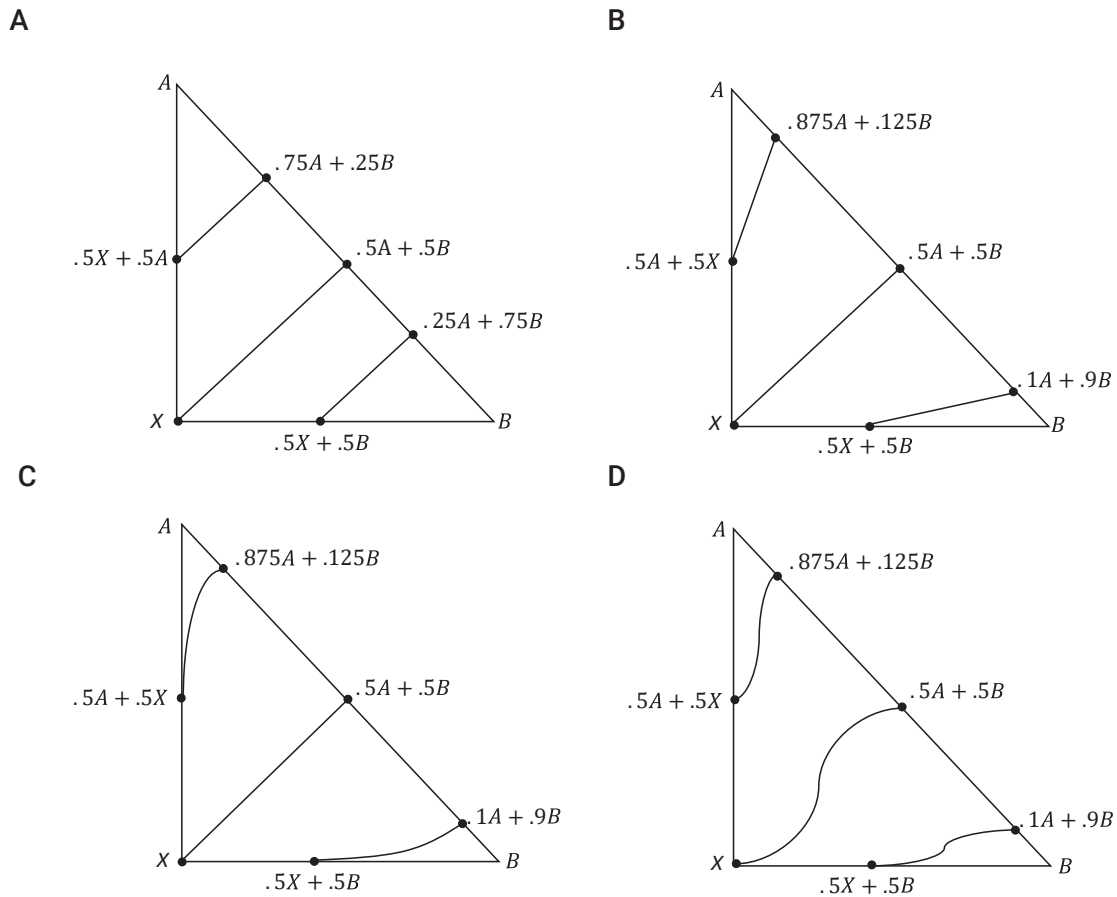
A

B

C

D

Figure 1: An illustration of the axioms in the Marschak–Machina triangle

Note: (1A) The preference relation $\succeq$ admits an Expected Utility representation. The indifference curves in the triangle are parallel straight lines. (1B) The preference relation $\succeq$ only admits a Weighted Expected Utility representation. The indifference curves are straight lines but they need not be parallel. (1C) The indifference curves are not straight lines as in Rank Dependent Utility and Prospect Theory. (1D) The preference relation $\succeq$ satisfy only State Monotonicity. The indifference curves are "upward sloping" but can otherwise be quite arbitrary.
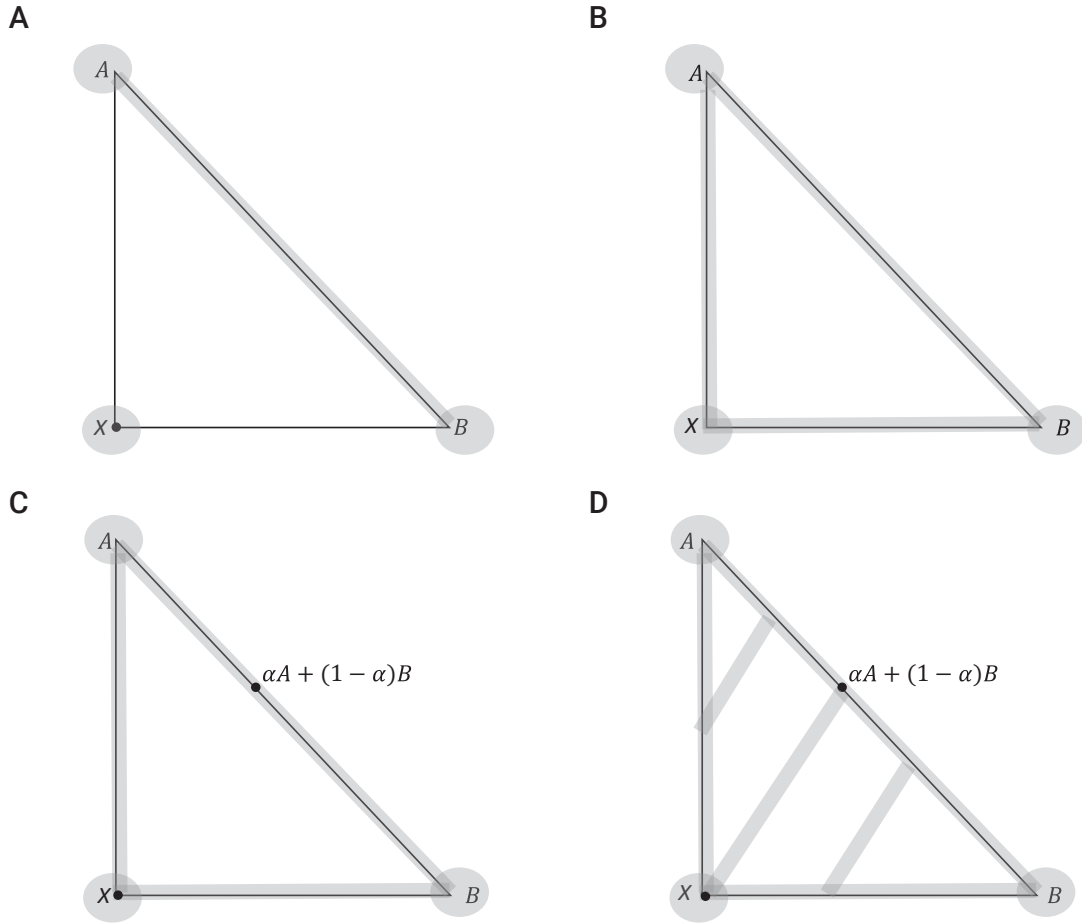
Figure 2: The difference between $\succeq_0$ and $\succeq_1$ in the Marschak-Machina triangle

Note: (2A) If we observe $\succeq_0$ we observe the ordering $A \succ_0 X \succ_0 B$ and the ordering of lotteries between A, B – but no others. (2B) If we observe $\succeq_0$ and we assume State Monotonicity we can *infer* the ordering of lotteries between A, X and lotteries between X, B. (2C) Continuity assures us that X is indifferent to *some* lottery $aA + (1 - a)B$, but if we observe only $\succeq_0$ then we do not observe *which lottery*, but if we observe $\succeq_1$ then we *do* observe which lottery. (2D) If we observe $\succeq_1$ and assume that $\succeq$ obeys Independence then observing *which lottery* completely determines $\succeq$.
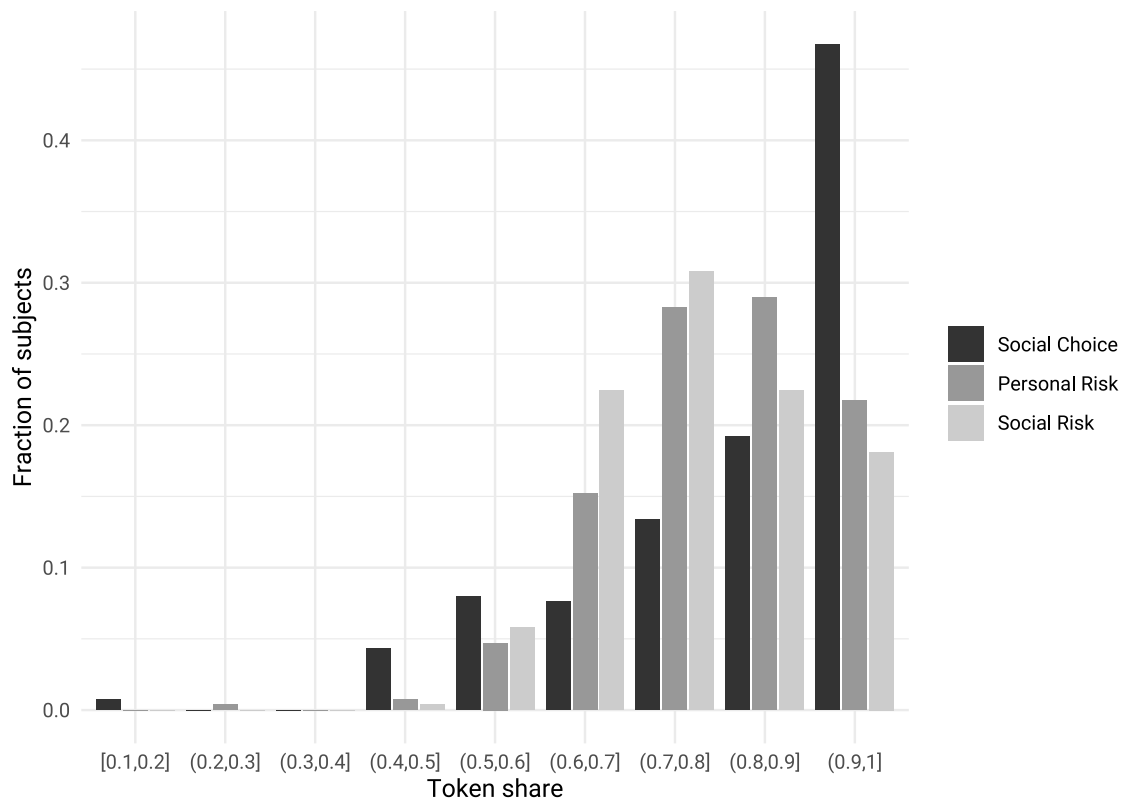
Figure 3: The distribution of aggregated behavior across subjects

Note: SOCIAL CHOICE: The distribution of tokens allocated to account y (kept by *self*) as a fraction of the tokens allocated to both accounts x and y (the total of tokens allocated to *self* and *other*). PERSONAL RISK and SOCIAL RISK : The fraction of tokens allocated to the *cheaper* account (the account with lower price – x when $p_x < p_y$ and y otherwise).
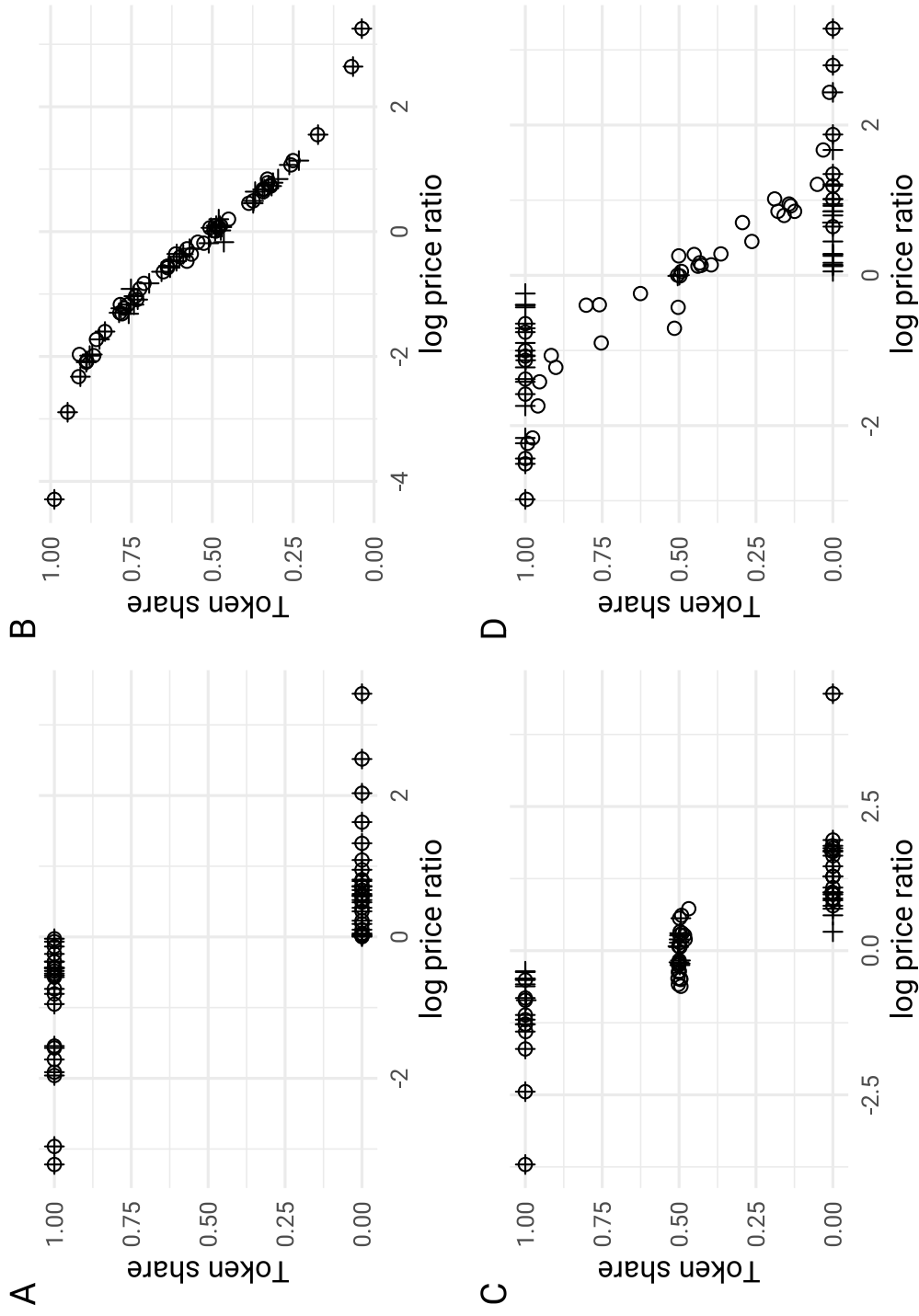
Figure 4: The choice behavior of four individual subjects

Note: The four subjects (IDs 511, 635, 317, and 645) are classified as selfish on the basis of their choices in the SOCIAL CHOICE treatment. Each entry shows the subject's relative demand $y/(x+y)$ (vertical axis) at a given log-price ratio $\ln(p_y/p_x)$ (horizontal axis) in the PERSONAL RISK (+) and SOCIAL RISK (○) treatments.
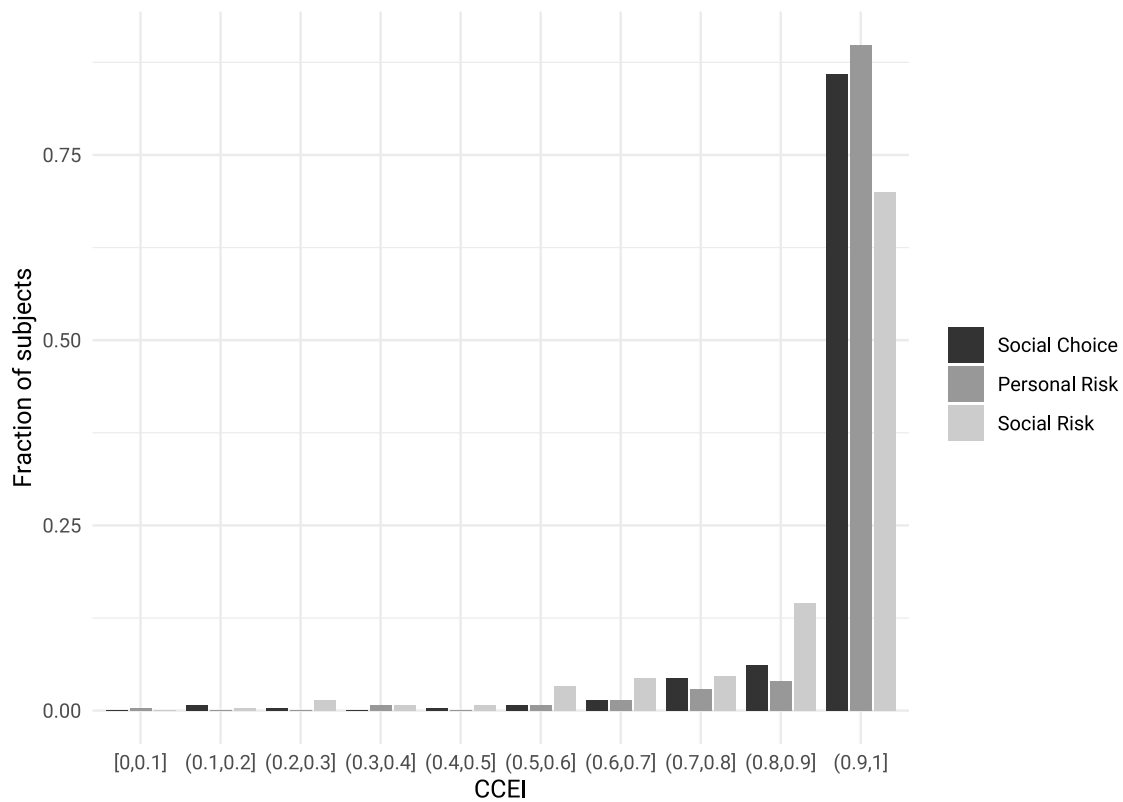
Figure 5: The distributions of CCEI scores for the three treatments

Note: The CCEI is between 0 and 1: indices closer to 1 indicate the data are closer to perfect consistency with GARP and hence to perfect consistency with utility maximization.
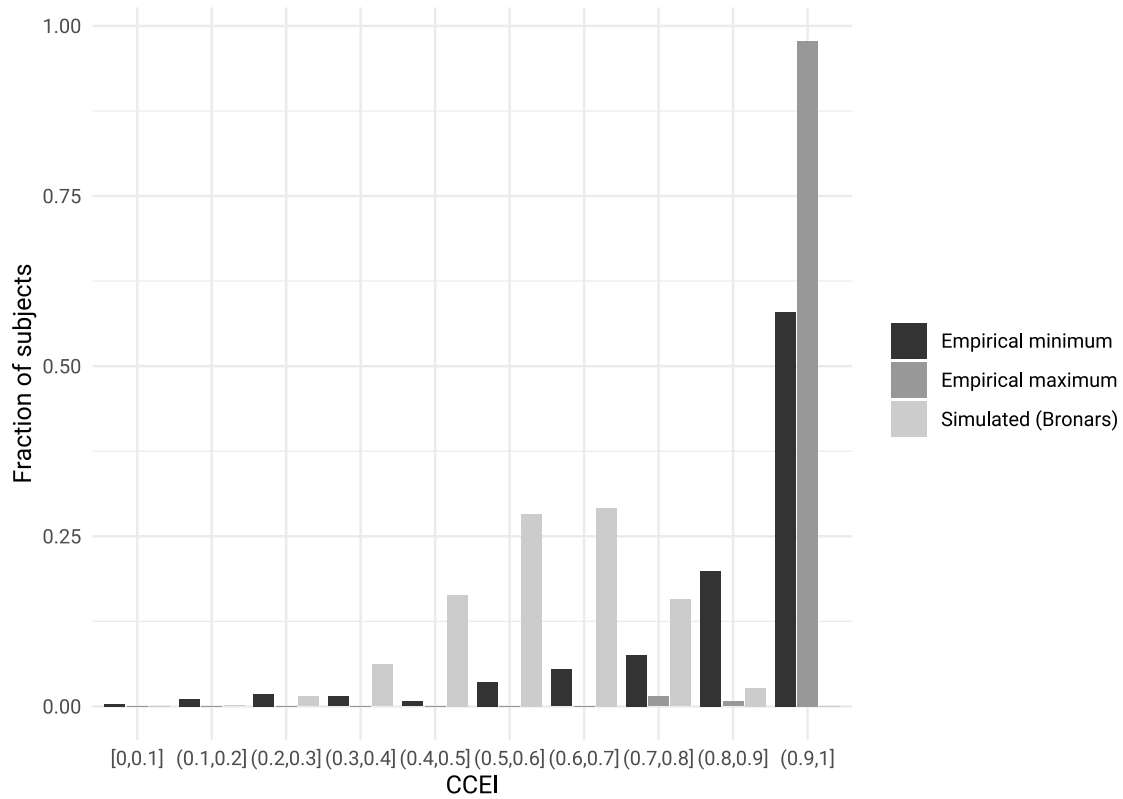
Figure 6: The power of the GARP test

Note: The distributions of the minimum and maximum CCEI scores for the three treatments (PERSONAL RISK, SOCIAL CHOICE, SOCIAL RISK) for the actual subjects and the distribution of the CCEI scores generated by 100,000 simulated subjects who randomize uniformly on each budget line (Bronars, 1987). Each of the 100,000 random subjects makes 50 choices from randomly generated budget sets in the same way as human subjects.
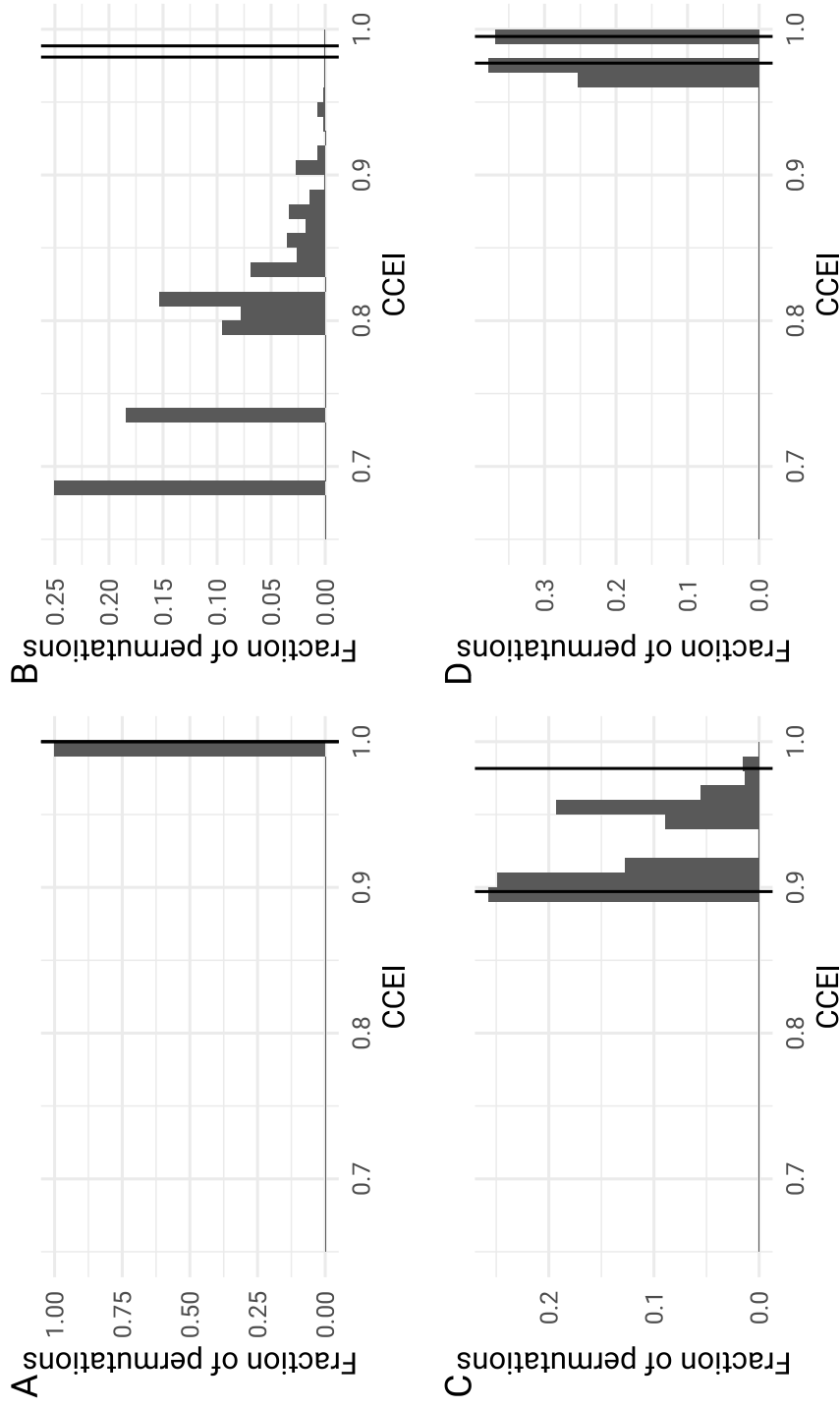
Figure 7: The test of Proposition 3 for four selfish subjects

Note: The four subjects – IDs 505, 729, 514 and 502 – are classified as selfish on the basis of their choices in the SOCIAL CHOICE treatment. Each panel presents the distribution of the permuted CCEI scores and the two actual CCEI scores from the PERSONAL RISK and SOCIAL RISK treatments (the vertical lines) – that is, $e^- = \min\{e_{\mathcal{PR}}, e_{\mathcal{SR}}\}$ and $e^+ = \max\{e_{\mathcal{PR}}, e_{\mathcal{SR}}\}$.

Table 1: The test of Propositions 3 and 4

Significance level	Rejection criterion		
	Combined Bonferroni-corrected	p^-	p^+
1%	0.107	0.049	0.107
5%	0.197	0.082	0.180
10%	0.205	0.115	0.230

Note: Proposition 3: If the subject's preferences $\succeq_{\mathcal{SC}}$ in the SOCIAL CHOICE domain are selfish, then the preferences in the PERSONAL RISK and in the SOCIAL RISK domains coincide: $\succeq_{\mathcal{PR}}=\succeq_{\mathcal{SR}}$. Proposition 4: If preferences $\succeq_{\mathcal{SC}}$ in the SOCIAL CHOICE domain are impartial, then preferences in the SOCIAL CHOICE domain and preferences in the SOCIAL RISK domain coincide: $\succeq_{\mathcal{SC}}=\succeq_{\mathcal{SR}}$. We present the fraction of participants with significant results using the Bonferroni correction, p^- and p^+ for the 1%, 5% and 10% levels.

Table 2: A benchmark for the test of Proposition 3

Significance level	Probability of random choice replacement (μ)				
	0	5%	10%	15%	20%
1%	0.097	0.165	0.388	0.553	0.718
5%	0.204	0.243	0.592	0.786	0.854
10%	0.214	0.408	0.748	0.874	0.922

Note: The fraction of selfish subjects for whom we reject the null that $\succeq_{\mathcal{PR}}=\succeq_{\mathcal{SR}}$ at the 1%, 5% and 10% levels when we replace the actual choices with random choices with different probabilities μ.