Disentangling Risk and Other-Regarding Preferences*

Paul Feldman†
Department of Agricultural Economics
Texas A&M University

Kristian López Vargas‡
Economics Department
University of California Santa Cruz

March, 2023

Abstract

We present a theoretical framework to study risk and other-regarding preferences jointly. The model can explain the main behavioral patterns in the combined domain and is the first to disentangle elementary attitudes towards risk, altruism, social substitution, and ex-ante or ex-post inequality. We devise an experiment with four decision environments based on convex choice sets. Our findings are consistent with the model, showing that most subjects adjust their risk attitudes due to ex-post inequality concerns and exhibit ex-ante fairness-seeking behavior. We estimate the model at the individual level and characterize subjects' heterogeneity of preferences over five fundamental dimensions.

JEL classification: C72, D03, D63, D64, D81.

Keywords: Risk Attitudes, Social Preferences, Other-Regarding preferences, Fairness, Inequality, Ex-Ante Fairness, Ex-Post Fairness, Epstein-Zin preferences.

*This research was sponsored by the National Science Foundation Dissertation Grant #1824121. We thank Zachary Breig, Shachar Kariv, Dan Friedman, Marco Palma, and Gerelt Tserenjigmid for their helpful comments. We also thank the feedback from audiences at the 2022 Econometric Society Meeting, the 2022 ESA World Meeting in Boston, the 2021 North American Meeting in Tucson, the Workshop of Experimental and Behavioral Economics of the Americas, the 2022 Galapagos Experimental Economics Conference; the 2022 LACEA-LAMES conference, among others. We thank Rachel Chen and Eli Pandolfo for their excellent programming work at the UCSC LEEPS Laboratory. We also thank Diego Tocre, Paula Armas, Martín Sanchez, and Grecia Barandiarán for their excellent research assistance.

†Email: paul.feldman@tamu.edu
‡Email: klopezva@ucsc.edu (corresponding author)

1 Introduction

Most decisions we make have uncertain consequences and social implications –they impact others we care about or with whom we compare ourselves while making the choices. Parents make risky investments affecting their children's future; the current generation's environmental policies aim to increase the odds of positive outcomes for future generations; entrepreneurs take on endeavors that might imply different risks for themselves and their partners; when allocating a prize or assigning a chore that cannot be shared, people often consider flipping a coin.

All of the examples above have *risks* and *others* as essential considerations in the decision; however, in economics, we often study them separately. In this paper, we document how they are intertwined to the degree that studying each in isolation leaves prevalent and relevant behavioral patterns unexplained. Thus, we propose an integrated approach to modeling other-regarding preferences and risk attitudes. To make a case for this approach, let us start with the four primary behavioral regularities concerning risk and social preferences and discuss intuitively the properties a model should have to account for them.

The first regularity is the well-documented *aversion to personal risks*: for most, five dollars for sure is strictly preferred to a coin-flip over zero and 10 dollars.[1] The second regularity is that many people exhibit non-trivial *other-regarding preferences*. This means they exhibit some degree of altruism or dislike for inequality where they typically sacrifice some personal material benefit to help others or to stick to the egalitarian allocation. For example, suppose the ordered pair (x, y) denotes a material payoff for a decision maker (DM) and another person, respectively. People often prefer the outcome pair $(\$15, \$15)$ over $(\$20, \$0)$. These first two regularities have been extensively researched in isolation in the standard approach. The next two intrinsically interlace attitudes toward *risks* and *others* and have been largely under-examined.

The third regularity is the concern for *ex-post inequality*: people tend to dislike unequal outcomes even if everyone faced the same opportunities ex-ante. Consider two coin-flip risks, A and B, which pay the DM zero or 30 dollars. Under A, the partner always receives the same payment as the DM. Under B, their fortune is reversed: the partner receives zero dollars if the DM receives 30 dollars and vice-versa. That is, $A := \frac{1}{2}(0, 0) + \frac{1}{2}(30, 30)$ and $B := \frac{1}{2}(0, 30) + \frac{1}{2}(30, 0)$. If the DM dislikes ex-post unequal outcomes, then the DM will strictly prefer prospect A to B, $A \succ B$, despite both prospects exposing each agent to the same personal risk. Indeed, as we document below, typically, people tolerate risks significantly more in prospects that lead to egalitarian payoffs than in prospects with unequal payoffs. In short, the anticipation of inequality lowers people's risk tolerance.

The fourth regularity is the tendency to prefer *ex-ante fairness* or equalized risks, often associated with the pursuit of *equal opportunity*. Suppose you must make a decision that impacts you

[1] This regularity also includes peoples' aversion to risks faced by others.

and a friend. There are two fixed, mutually exclusive outcomes: A, advantageous to you, and B, advantageous to your friend. For example, you own only one concert ticket for an artist both of you like. If you could decide the probability of getting the ticket (outcome A) or the other person (outcome B), which probability would you choose? As documented below, typically, a person chooses to give the other person some chance to get the prize at the expense of her own prospects.[2] For example, suppose the two outcomes are represented by $A = (40, 0)$ and $B = (0, 40)$. You might prefer to flip a coin (the 50-50 randomization) over the best of the two outcomes. Formally, $\frac{1}{2}A + \frac{1}{2}B \succ A \succ B$. As detailed below, this propensity to strictly prefer the randomized lottery over the two allocations is a fairly prevalent behavior. Throughout the paper, we use the notion of *fairness*, which refers to egalitarian outcomes or prospects; therefore, we use the terms *unfair* and *unequal* interchangeably.

What models can explain these empirical regularities? The neoclassical model with selfish preferences explains only the first regularity of *aversion to personal risks*. Then, over the last four decades, economists have addressed the second regularity by developing models of *other-regarding preferences*. Nevertheless, they mainly consider only risk-free settings and, as such, are unable to make predictions regarding the last two regularities. To account for the third regularity, we can extend social-preferences models of *inequality aversion* (Fehr and Schmidt, 1999; Charness and Rabin, 2002, e.g.) to lotteries using the Expected Utility Theory (EUT).

Modeling the fourth regularity is more challenging. In the example above, the propensity of the DM to randomize who gets the prize is directly incompatible with EUT, the cornerstone of most economics models. In fact, nearly all models that weaken EUT's core assumption of *independence* still fail to explain the concerns for ex-ante inequality. The Cumulative Prospect Theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), the models of Rank-Dependent Expected Utility (Quiggin, 1982), the models of Disappointment Aversion (Gul, 1991), and even more recent theoretical work linking social preferences and attitudes toward risks (Zame et al., 2020), are all unable to explain people's propensity to randomize for ex-ante fairness motives.[3]

This paper proposes a joint model of risk attitudes and other-regarding preferences that accommodates *all* four regularities. Our model disentangles preferences in the risk and social domains in an approach that parallels to some extent the literature that untangles risk and time preferences (Kreps and Porteus, 1978; Epstein and Zin, 1989; Dillenberger et al., 2020) but adds elements that are inherent to social considerations. When there are no social comparisons because all states

[2]Although we fully acknowledge the value of using neutral pronouns, we often use she/her pronouns for the decision-maker in order to avoid ambiguity.

[3]In the case of non-expected utility preferences described above as long as the underlying preferences are defined over social outcomes and not social lotteries, these models cannot accommodate a strict preference for our ex-ante fairness example above. In the case of linking preferences, state monotonicity is the axiom that rules out the strict preference for the coin toss.

yield to equality, only basic risk attitudes are at play, and the framework admits any theory of choice under uncertainty. When there is no risk, any model of deterministic social preferences is admissible. When choices involve both social considerations and risk, our model extends the notion of *inequality aversion* to lotteries in a manner that is amenable to decomposing the dislike of inequality into ex-post and ex-ante components. Further, the our formal notion of ex-ante fairness is novel and based on the idea that individuals compare their risks with those faced by their partners. In particular, ex-ante fairness is modeled as a utility discount that increases with the distance between the marginal distributions of payoffs faced by the DM and the other person. When social considerations are turned off, by imposing equality in outcomes across states, our model simplifies to EUT and can accommodate any Bernoulli utility function. Similarly, our model can accommodate arbitrary risk-less social preferences.

Importantly, we propose a parameterization of our model that provides the necessary structure to characterize each individual according to their fundamental attitudes toward risk, altruism, fairness, ex-ante inequality, and ex-post inequality. Our parametric instance encompasses CRRA utility, and generalizes most existing utility models of social preferences and fairness concerns, including Fehr and Schmidt (1999); Charness and Rabin (2002); Fisman et al. (2007).

To our knowledge, no other model can accommodate the four empirical regularities. However, a close noteworthy relative to our model is the Expected Inequality Aversion (EIA) model, introduced by Fudenberg and Levine (2012) and axiomatized by Saito (2013). The EIA model was advanced to explain ex-post and ex-ante fairness concerns and postulates that the DM uses a linear convex combination of the expected utility and the utility of expected values with a root utility given by Fehr and Schmidt (1999).[4] Nonetheless, the EIA model makes a strong prediction regarding the relation between ex-ante motives and risk tolerance: the more ex-ante fairness-seeking behavior, the higher the tolerance to risks. This prediction holds even if we generalize the model to add flexibility. In contrast, the parametric instance of our model (with the same number of parameters) does not impose nor restrict any correlation between ex-ante motives and risk tolerance.

We study the implications of our model in the laboratory by leveraging experimental and econometric techniques introduced in Andreoni and Miller (2002); Fisman et al. (2007); Choi et al. (2007); Andreoni and Sprenger (2012a,b); Brock et al. (2013). We design four choice environments, each allowing us to identify each dimension or slice of an individual's preferences. Together, choices in these environments identify all parameters in the model. We estimate the parametric instance of the model at the individual level and, in this way, study the distribution of the underlying fundamental preferences. We use an intuitive graphical interface for each choice environment that presents each decision-maker with a convex budget set. This approach makes it possible to collect numerous choices per participant and to characterize individual preferences

[4]Formally, $U(L) = \delta \mathbb{E} u(x, y) + (1 - \delta) u(\mathbb{E}x, \mathbb{E}y)$, where $u(\cdot)$ is given by Fehr and Schmidt (1999).

in detail. In each session, we pair participants at random and anonymously. In each of the following four experimental choice environments presented below, the DM faces a convex set of joint distributions $F_{X,Y}(x, y)$ over payoffs for themselves (x) and another participant (y).[5]

Our first environment is **deterministic giving**. For each choice in this environment, the DM must choose an allocation for herself (x) and her counterpart (y) that lie on a budget constraint $m = p_x x + p_y y$. Here, m represents the DM's wealth, and p_y/p_x is the cost of giving one currency unit to the other person. This environment identifies social preferences associated with choices without risk.

Our second environment consists of **ex-post fair risks**. There are two equally likely states, A and B. The DM selects a bundle (x_A, x_B) of securities A and B that pay one currency unit if the associated state realizes, subject to a budget constraint, $m = p_A x_A + p_B x_B$. Here, m represents the DM's wealth, and p_A (p_B) is the price of one unit of security A (B). These securities entail ex-post-fair risks because each security pays the counterpart the same amount it pays the DM ($y_A = x_A$ and $y_B = x_B$). Note ex-post fairness implies ex-ante fairness because state-by-state equality also implies equal personal risks. Therefore, this task allows us to identify the basic risk attitudes –since choices are free of any inequality concern.

Our third choice environment consists of **ex-ante fair but ex-post unfair risks**. This choice environment is the same as the second, except both agents have their fortunes reversed. As before, under state A (B) DM gets x_A (x_B). However, now the counterpart gets $y_A = x_B$ in state A, and $y_B = x_A$ in state B. From the personal risk view (own marginal distribution of payoffs), the prospects in this environment and the previous one are identical. The only difference between the two environments is that now taking any risk involves ex-post inequality. Therefore, differences in behavior between these environments can only be attributed to ex-post inequality concerns.

Lastly, we present the participants with ex-ante dilemmas. We refer to these tasks as **giving or sharing in chances** because we present the DM with two states A and B, each with a fixed pair of monetary outcomes (x_A, y_A) and (x_B, y_B), and the DM must choose the probabilities of each state so that they add up to one. Moreover, outcomes under A and B are not materially Pareto-ranked ($x_A > x_B$ if and only if $y_B > y_A$). In this environment, the DM can only give to others (or seek fairness) through probabilities. Therefore, choices here identify the strength of the ex-ante motives. In all four tasks, participants face 50 budget lines (decisions) except in the sharing in chances task, where they face 28 choices.

Our empirical analysis is two-fold. We first advance three broad hypotheses that emerge from the general features of the conceptual framework. (1) We conjecture that people exhibit ex-post fairness-seeking behavior, which can be captured in the different behavioral patterns of risk-taking between the second and third-choice environments. Our study, to our knowledge, is the first to

[5]The four choice environments match the four regularities discussed above.

precisely measure the extent to which risk-taking changes with fairness concerns at the aggregate and individual levels. We find that ex-post inequality concerns impact risk attitudes substantially. People's frequency of choosing the safe lottery goes up by over 50% when they face ex-post unfair risks relative to when they face ex-post fair ones. (2) We hypothesize that the majority of people exhibit ex-ante fairness-seeking behavior. To our knowledge, ours is the first detailed measurement of the strength of ex-ante fairness attitudes at the individual and aggregate levels. By presenting subjects with only a few ex-ante dilemmas, previous studies were not designed to measure the pervasiveness of this behavior but to show that it exists.[6] Using different methods, we document that the share of participants that consistently exhibit ex-ante fairness-seeking behavior is in the range of 50%-60% of our sample. (3) We test one of the main predictions of the leading alternative model: that individuals with higher ex-ante fairness motivations tolerate more risks. We do not find evidence that supports this implication of the alternative model.

Our main analysis is a structural estimation at the individual level. We derive closed-form solutions for the optimal behavior in our model using each of the four choice environments described above and use robust minimum-distance estimation techniques to fit the model's parameters at the individual level. We characterize each person's risk and social preferences, disentangling their fundamental components: basic risk attitudes, selfishness, and inequality aversion, and measure the strength of fairness motives (ex-ante and ex-post). To our knowledge, we are the first to implement this detailed characterization of elementary attitudes in the combined domain of risks and concern for others. Results from this structural exercise also confirm the first two broad hypotheses.

Finally, we compare the predictive power of our model and the enhanced version of the alternative model –a generalization of the EIA model with an equal number of parameters as ours. We conduct a series of individual-level out-of-sample prediction exercises. As suggested by the theory, our model outperforms the alternative model in tasks involving risks.

The rest of the paper has the following structure. Section 2 discusses the existing literature. Section 3 describes the model. Section 4 presents the experimental design, hypotheses, and implementation. Section 5 reports the results of the experiments, and Section 6 concludes. We present propositions' proofs, derivations, experiment instructions, and supplementary results in the Appendices.

2 Related Literature

In the standard model, risk preferences are considered stable and determined only by the choice context, e.g., prices and income, and independent from social considerations (Stigler and Becker, 1977). Similarly, other-regarding preferences are mostly modeled under risk-less environments

[6]These studies did not explore regions of the space of outcome pairs where this behavior could emerge, and so underestimated the actual prevalence.

(Camerer, 2003; Fehr and Schmidt, 2006; Meier, 2007). Indeed, some early experimental studies claimed no empirical interaction between other-regarding preferences and risk attitudes (Brennan et al., 2008; Bolton and Ockenfels, 2010). Others claimed individual risk attitudes were only marginally affected by others' risks and that people preferred risks to be independent across individuals in society rather than correlated (Rohde and Rohde, 2011).

More recent studies have examined the relationship between fairness concerns and risk attitudes more closely and found non-trivial results. We identify two branches of this literature. The first focuses on ex-ante fairness or equal opportunity; the second focuses on ex-post fairness –i.e., inequality post uncertainty. Only a few models following this literature have begun incorporating both notions of fairness in settings with uncertainty.

In the literature on ex-post fairness, a central question is whether individuals avoid risks with a negative correlation (between their outcomes and others') more than risks with a positive correlation. Adam et al. (2014) and López Vargas (2015) find that subjects are more risk averse when paired in a negatively correlated lottery than in a positively correlated one. Gaudeul (2015) shows evidence that supports a preference for positively correlated payoffs over uncorrelated and these over negatively correlated ones. Finally, Trautmann and Vieider (2012); Schwerter (2013); Fafchamps et al. (2015); Gamba et al. (2017); Jaramillo and López Vargas (2019) study risk attitudes keeping others' payoff fixed or using it as a reference point. This evidence suggests that exogenously manipulating others' payoffs does alter risk-taking behavior. These papers document the existence of non-trivial preference differences. Still, none present experimental designs that purposefully measure the degree to which risk attitudes are altered due to fairness concerns in the extensive and intensive margins as our paper.

In the branch focused on ex-ante fairness, Karni and Safra (2002) presents a dual-self model where the decision-maker combines an ex-ante-fair self with an egotistic self. Similarly, Borah (2013) presents a model with two components of utility: an individual's expected utility over social outcomes and an ex-ante component that depends only on the risks the other agent faces. Interestingly, nearly all other theories of fairness or altruism in probabilistic environments assume people compare expected outcomes. For example, Bolton et al. (2005) extend the ERC model (Bolton and Ockenfels, 2000), assuming, as in the original model, that people care about relative payoffs, except these are replaced by relative expected values. Similarly, Trautmann (2009) extends Fehr and Schmidt (1999) to the uncertainty case by simply making the Fehr-Schmidt inequality discounts depend on expected outcomes. Krawczyk and Le Lec (2010) propose a formulation that linearly combines an egoistic expected utility component and a fairness component that depends, in turn, on the (subjective) expected outcomes of the DM and another agent. The preference for randomization for ex-ante fairness purposes is directly incompatible with EUT and with nearly all models that weaken EUT's core assumption of *independence*[7] The Cumulative Prospect Theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), the models of Rank-Dependent Expected Utility (Quiggin, 1982), the models of Disappointment Aversion (Gul, 1991) that rely on the axiom of *betweenness*[8] and most relative models are all unable to explain people's propensity to randomize for ex-ante fairness motives. In the experimental literature on ex-ante fairness, Krawczyk and Le Lec (2010) and Brock et al. (2013) document evidence from a dictator-like game, where individuals must decide the probabilities of two mutually exclusive, undominated outcomes. In both papers, the authors find social preferences for randomization (exhibiting ex-ante fairness) is a persistent behavior–at least one-third of subjects assigned positive probabilities to unfavorable outcomes. Notably, none of these papers expound on the role of risk attitudes in social situations. Furthermore, with only a few choices elicited through ex-ante dilemmas, their experiments were not designed to properly measure the prevalence of ex-ante fairness-seeking behavior.

To our knowledge, the only existing model to incorporate and explicitly balance ex-post and ex-ante fairness, introduced by Fudenberg and Levine (2012) and Saito (2013), is the Expected Inequality Aversion (EIA) model. In this model, preferences are represented by a linear combination of the utility of expected outcomes and their expected utility. The utility over expectations is meant to capture ex-ante concerns, and the expected utility is meant to capture ex-post concerns. Moreover, in this model, the root utility is the standard and deterministic Fehr and Schmidt (1999) utility (FS). This model has been used to organize experimental data in Brock et al. (2013), and Gaudeul (2013). However, Gaudeul (2013) replaces each outcome in the FS utility with a power function of the corresponding outcome.

A significant recent contribution to the literature connecting social preferences and risk attitudes is advanced in Zame et al. (2020). They provide the conditions under which identifying deterministic social preferences and attitudes toward personal risks suffice to identify the preference relation over all social prospects completely. Notably, their theory, in its core axiom, *state monotonicity*[9], rules out ex-ante fairness and, therefore, does not encompass the utility models we propose in our paper.

Our paper is also connected to the large literature that studies entwined risky and inter-temporal preferences and aims to separate those theoretically and empirically in a principled manner. See Kreps and Porteus, 1978; Epstein and Zin, 1989 for the seminal ideas, and Andreoni and Sprenger (2012b) and Dillenberger et al. (2020) for more recent empirical and theoretical contributions, respectively. Our paper shares with this literature the principle of shutting down one motive (no

[7]*Independence* states that if A, B, and C are lotteries, $\alpha \in [0,1]$, and $A \succeq B$, then, $\alpha A + (1-\alpha)C \succeq \alpha B + (1-\alpha)C$.

[8]*Betweenness* states that if A and B are two lotteries, $\alpha \in [0,1]$, and $A \succeq B$, then $A \succeq \alpha A + (1-\alpha)B \succeq B$.

[9]*State Monotonicity* states that if we consider k pairs of outcomes $(\omega_i, \omega_i')_{i\in\{1,\ldots,k\}}$, we have $\omega_i \succeq \omega_i'$ for all i, and $p = (p_1, \ldots, p_k)$ is a probability vector, then $\sum_{i=1}^k p_i\omega_i \succeq \sum_{i=1}^k p_i\omega_i'$.

risk or time-invariant consumption or perfect equality) to identify the other motive.

Other studies are less directly related to our paper. Cappelen et al. (2013) ask what people's typical fairness views regarding risk-taking behavior are. Fairness views are cleverly associated with whether or not an external observer redistributes post-uncertainty payoffs in a society where subjects can choose different degrees of risk. They find significant heterogeneity in people's fairness views. Karni et al. (2008) empirically study the predictions of Karni and Safra (2002). Chakravarty et al. (2011) find that individuals making decisions for a stranger exhibit less aversion towards risks faced by the stranger than towards their risks. Van Koten et al. (2013) study risk attitudes restricted to uncertain pie sizes in bargaining games. Harrison et al. (2012) study how risk attitudes towards social outcomes vary with information regarding risk preferences of the other agent, finding that learning others' risk preferences makes individuals more risk averse.

3 The Model

Consider the space of two-dimensional payoff outcomes, $\mathcal{X} \times \mathcal{Y} \subset \mathbb{R}^2_+$. An element of this space, (x, y), represents a pair of payoffs for the DM and her counterpart, respectively. For simplicity, we assume $\mathcal{X} \times \mathcal{Y}$ is finite of size K. Let $\mathcal{L}$ denote the space of lotteries over outcomes in this space, $\Delta(\mathcal{X} \times \mathcal{Y})$. Each lottery, $L \in \mathcal{L}$, has a discrete cumulative distribution function (CDF), $F_{X,Y}(x, y)$, associated with it. Any arbitrary L can also be described by an array $(p_k, x_k, y_k)_{k \in \{1,2,\ldots,K\}}$ where p_k is the probability of the kth outcome according to L. We often use the CDF and array representation of a lottery in the same expression for notational ease. We present our theory using a utility representation form for conciseness. Future work will present the axiomatic build of this theory. The general representation of our utility model is given by:

$$U(L) = W\left[(p_k, g\left[x_k, y_k, D\left(F_X, F_Y\right)\right])_k\right]. \tag{1}$$

F_X and F_Y denote the marginal CDFs of x and y, respectively, given the joint, $F_{X,Y}$, associated with L. $W[\cdot]$ is a *risk aggregator* capturing the *basic risk attitudes*, as illustrated below. We use a general form for $W[\cdot]$; different theories of risk preferences (e.g., prospect theory or disappointment aversion) will impose different conditions on it. With a slight abuse of notation, let $g_k = g\left[x_k, y_k, D\left(F_X, F_Y\right)\right]$. We assume that $W[(p_k, g_k)_k]$ is continuous and supermodular in each (p_k, g_k) [10] as well as increasing with respect to g_k for all k. Additionally, it is symmetric with respect to any two pairs (p_k, g_k) and $(p_{k'}, g_{k'})$. The function $g[\cdot]$ is a *social aggregator* that models other-regarding preferences. We only require $g[\cdot]$ to be strictly increasing in x, and non-increasing in D. $D(\cdot)$ is a metric in the space of discrete CDFs with finite support in $\mathcal{X} \cup \mathcal{Y}$. It is a measure of ex-ante inequality since it compares the personal risks the two agents bear modeled

[10]This means better social outcomes and higher likelihoods for them are complements.

by their marginal CDFs F_X and F_Y. The following condition enables preference separation of risk and other-regarding preferences in the Epstein-Zin sense.

Assumption 1. $g[z, z, D(F_Z, F_Z)] = z$

This assumption says that in the class of perfectly ex-post fair lotteries, the scale of the social aggregator is the same as the scale of consumption. More importantly, any other property of $g[\cdot]$ does not influence its shape over this class of lotteries.

To identify basic risk attitudes, we shut down social considerations by focusing precisely on lotteries where $P[x = y] = 1$. Being a metric, $D(\cdot) = 0$ for any lottery of this class. By Assumption 1, the utility simplifies to $U(L) = W[(p_k, z_k)_k]$. That is only the properties of $W[\cdot]$ shape preferences in the domain of ex-post fair lotteries.

To identify deterministic preferences, we shut down the risk dimension by only considering lotteries that are degenerate. With degenerate marginal CDFs, $D(F_X, F_Y)$ induces a metric in $\mathcal{X} \cup \mathcal{Y}$ that we denote by $d(x, y)$. Since $W[.]$ is assumed to be symmetric, supermodular in each entry (p_k, g_k) and strictly increasing with respect to all g_k, we can take the appropriate increasing transformation so that deterministic preferences (utility over the degenerate lotteries) can be represented by $g[x, y, d(x, y)]$.[11] In sum, the utility components that determine preferences over *deterministic social preferences* and *risk attitudes* over completely egalitarian lotteries can be *independent* of each other. This framework encompasses EUT and most theories of risk attitudes as well as near-arbitrary theories of deterministic social preferences as long they assume rationality and continuity. The following example illustrates how this model encompasses EUT.

Example 3.1. *Expected Utility Theory:*
In the subdomain of lotteries with equal social outcomes, $Pr[x = y] = 1$, for the underlying preferences to be consistent with EUT, we need $W(\cdot)$ to take the form $W[(p_k, z_k)_k] = \sum_k p_k \times w(z_k)$ for some increasing (Bernoulli) function $w(\cdot)$. If we want the model to be consistent with EUT over the whole set, $\mathcal{L}$, then we need an additional condition: $g[\cdot]$ must be invariant with respect to its third argument, $D(\cdot)$. Otherwise, $U(L)$ will not be linear in outcome probabilities.

Overall, our model provides a framework to study choices that involve both motivations simultaneously: evaluating risks and social considerations. Critically, we want this model to explain the third and fourth behavioral regularities presented in the Introduction: that individuals prefer ex-post equality of outcomes and ex-ante fair joint distributions. To formalize both regularities, we need to define two properties of $g[\cdot]$:

[11] Let $g_k := g(x_k, y_k, d(x_k, y_k))$ and consider two degenerate lotteries with certain outcomes $(x_{k'}, y_{k'})$ and $(x_{k''}, y_{k''})$. Note that $g_{k''} > g_{k'}$ if and only if $W[..., (1, g_{k''}), ..., (0, g_{k'}), ...] > W[..., (1, g_{k'}), ..., (0, g_{k''}), ...] = W[..., (0, g_{k''}), ..., (1, g_{k'}), ...]$ where the strict inequality comes from the supermodularity of $W[.]$ and the equal sign from its symmetry.

1. *Ex-post Fairness:* $g[x, y, \cdot]$ is supermodular with respect to (x, y).

2. *Ex-ante Fairness:* $g[\cdot, \cdot, D(F_X, F_Y)]$ is strictly decreasing in its third argument, $D(\cdot)$.

Suppose we have outcomes $A = (x_A, y_A)$ and $B = (x_B, y_B)$ with $x_A > x_B$ and $y_B > y_A$. Then, Property 1 (ex-post fairness) implies $\frac{1}{2}(x_A, y_B) + \frac{1}{2}(x_B, y_A) \succ \frac{1}{2}A + \frac{1}{2}B$. In particular, $\frac{1}{2}(10, 10) + \frac{1}{2}(0, 0) \succ \frac{1}{2}(10, 0) + \frac{1}{2}(0, 10)$ as in the example in the Introduction. Property 2 states that the DM dislikes it when individual risks (measured by the marginal CDFs of payoffs) between the two agents are different. In particular, this property is the main necessary condition of ex-ante fairness-seeking behavior: there exist outcomes A and B, and $\alpha \in (0, 1)$ such that $\alpha A + (1 - \alpha)B \succ A \succeq B$.[12]

The model in Equation 1 provides a rather general framework on how social and risk-related motives operate and interact. We need further structure to identify and disentangle these motives and characterize preferences. We therefore introduce a parametric instance of the model in Equations 2, 3 and 4.

$$U(L) := \mathbb{E}_{F_{X,Y}} \left[\frac{1}{\gamma} \left[a x^{\rho} + \bar{a} y^{\rho} - s_{\rho} \theta \left(\bar{\delta}\, d(x, y) \, + \, \delta\, D(F_X, F_Y) \right) \right]^{\frac{\gamma}{\rho}} \right] \tag{2}$$

where

$$d(x, y) := |x^{\rho} - y^{\rho}| \tag{3}$$

and

$$D(F_X, F_Y) := \int_{-\infty}^{\infty} (F_{X^{\rho}}(t) - F_{Y^{\rho}}(t))^2 \, dt, \tag{4}$$

We assume that $\mathcal{X}$ is a bounded grid graph in $\mathbb{R}_+$ and $\mathcal{X} = \mathcal{Y}$ for simplicity.[13] $F_{X^{\rho}}$ and $F_{Y^{\rho}}$ are the marginal CDFs of X^{ρ} and Y^{ρ}, respectively. Parameter γ captures *basic risk attitudes* since $1 - \gamma$ is the constant relative risk aversion coefficient exhibited when lotteries are free from any inequality (when there are payoffs that are state-by-state egalitarian). Parameter $a \geq 0$ captures the degree of selfishness; conversely, $\bar{a} := 1 - a$ reflects the degree of altruism. Parameter $\rho \in (-\infty, 1]$ captures the degree of substitution between x and y, since $\frac{1}{1-\rho}$ is the constant elasticity of social substitution (ESS) in deterministic preferences everywhere except at the $x = y$ loci. Intuitively, this parameter captures the tolerance to unfair/unequal outcomes. For notational ease, we define $s_{\rho} := \text{sign}(\rho)$. Parameter $\theta \in [0, a]$ captures the aversion to inequality in the ex-ante and the ex-post sense together. In this parametric framework, either type of inequality causes a form of utility discount. Our model can distinguish the two types of inequality concerns: ex-post and ex-ante. $d(\cdot)$ captures the discount for ex-post inequality, and $D(\cdot)$ captures the discount for ex-ante inequality. Parameter $\delta \in [0, 1]$ is the weight of ex-ante concerns in relation to all inequality

[12] The sufficient condition for this behavior requires a slightly more technical condition.

[13] We assume this to match the experiment environment described in the next section.

concerns. Similarly, $\bar{\delta} := 1 - \delta$ is the weight of ex-post inequality concerns.[14]

In terms of the general framework given in Equation 1, the parametrization given in Equation 2 sets: $W[(p_k, g_k)_k] = \sum_k p_k \frac{1}{\gamma} g_k^{\gamma}$; and $g[x, y, D] = \left[ax^{\rho} + \bar{a}y^{\rho} - s_{\rho}\theta\left(\bar{\delta}d(x, y) + \delta D(F_X, F_Y)\right)\right]^{\frac{1}{\rho}}$.

As elaborated below, this parameterization of the model is flexible yet tractable. Notably, it encompasses CRRA for the case of ex-post egalitarian lotteries and most models of deterministic other-regarding preferences and fairness concerns (Fehr and Schmidt, 1999; Andreoni and Miller, 2002; Charness and Rabin, 2002; Fisman et al., 2007; Cox et al., 2007, e.g.).[15]

Example 3.2. *Fehr and Schmidt (1999) proposed the following utility model (FS model, hereafter):*

$$u(x, y) = x - \beta\, max\{0, x - y\} - \alpha\, max\{0, y - x\} \tag{5}$$

defined over the space of riskless payoff outcomes and with $\beta \in [0, 1)$ and $\alpha \geq \beta$. Any admissible parameterization of the FS model can be obtained in our model by setting: $\rho = 1$, $a = 1 + \frac{\alpha - \beta}{2}$ and $\theta = \frac{\alpha+\beta}{2}$. The parameters γ and δ can take any value because there is no risk; therefore, ex-post and ex-ante inequality coincide. See Proof A.1 in Appendix A.

Similarly, Charness and Rabin (2002) also presented a piece-wise linear utility with inequality aversion. Our proposed model also nests utility models representing riskless preferences where, unlike the FS model, giving choices respond more smoothly to the price of giving, as documented in Andreoni and Miller (2002), Fisman et al. (2007), and Cox et al. (2007). In these last three papers, the utility is represented by variations of constant elasticity of substitutions (CES), and there is no kink at the 45-degree line of the (x, y) plane, except in the limit case when ρ goes to $-\infty$. Therefore, to nest these models, we only need $\theta = 0$, and all other parameters are free.

Although our model contemplates the possibility of upward-slopping indifference curves when $y > x$, as in Fehr and Schmidt (1999), for simplicity in the characterization of preferences and the experimental design, we assume all indifference curves for deterministic social preferences are downward slopping. This requires $\theta \leq \min\{a, \bar{a}\}$ (see Proof A.2 to Proposition 3.1 in Appendix A).

3.1 Model Predictions and Identification

In this subsection, we characterize the properties of the model and its predicted behavior in four convex choice set environments. The model's closed-form prediction in each environment identifies a "slice" of preferences in that a subset of the model's parameters governs it. Jointly, the predictions in the four environments identify all parameters. This theoretical analysis is linked to

[14]For simplicity, we assume that inequality aversion operates in relation to equal payoffs $(x = y)$. This model could be extended to more general views of what constitutes a fair outcome for a decision-maker.

[15]In the case of Charness and Rabin (2002) and Cox et al. (2007), our model excludes their reciprocity components. However, extending our model to incorporate this component is straightforward.

the design of the experiment, as these four classes of choices correspond to the four groups of tasks in the experiment, and the analytical representation of the optimal behavior is used to estimate the parameters structurally at the individual level.

Deterministic Giving

First, consider choices where there is no risk. That is, the DM knows or chooses x and y with certainty. In this case, all CDFs are degenerate, then $D(F_X, F_Y) = d(x, y) = |x^\rho - y^\rho|$. Since inequality occurs with certainty, ex-post and ex-ante motivations cannot be separated, then δ does not play a role in these choices. Furthermore, as discussed above, we can represent the same deterministic preferences using a utility function that does not depend on γ. Specifically, the preferences represented by the utility function in Equation 2, can also be represented by the following CES utility with a kink on the 45-degree line ($x = y$):

$$u(x, y) = [(a - s_\Delta\theta)x^\rho + (\bar{a} + s_\Delta\theta)y^\rho]^{\frac{1}{\rho}} \tag{6}$$

where $s_\Delta = sign(x - y)$ and for $\rho \neq 0$. When $\rho = 0$, $u(x, y)$ becomes $x^{a - s_\Delta\theta} y^{\bar{a} + s_\Delta\theta}$.[16] As described above, this utility generalizes most existing models of deterministic social preferences. Next, we define a choice environment and derive the corresponding prediction governed by parameters a, ρ, and θ.

Definition 3.1. ***Choice Environment: DetGiv***
DM must choose a bundle, (x, y)*, from the set* $\{x, y : m = p_x x + p_y y\}$ *for* $m, p_x, p_y > 0$.

Proposition 3.1. *DM's optimal behavior in the choice environment DetGiv satisfies:*

$$\frac{x}{y} = \left[min\left\{max\left\{1, \frac{p_x}{p_y}\frac{\bar{a} - \theta}{a + \theta}\right\}, \frac{p_x}{p_y}\frac{\bar{a} + \theta}{a - \theta}\right\}\right]^{\frac{1}{\rho - 1}} \tag{7}$$

See Proof A.2 in Appendix A. From this proposition, we can see that the size of the kink at $x = y$, $\frac{\lim_{x \downarrow y} |MRS|}{\lim_{x \uparrow y} |MRS|}$, is given by $\frac{(a+\theta)(\bar{a}+\theta)}{(\bar{a}-\theta)(a-\theta)}$, which is strictly greater than one if and only if $\theta > 0$.

Ex-Post Fair Risks

Now, we consider the class of lotteries that are *completely fair* or *ex-post fair* (EPF) in the sense they yield egalitarian outcomes in every state. Formally, $Pr[X = Y] = 1$, which implies $d(\cdot) = 0$

[16]Visual comparative statics of the indifference curves are presented in Figures 16c, 16d and 16e in Appendix J.

and $D(\cdot) = 0$.[17] It is immediate to see that this utility model collapses to Equation 8.

$$U(L) = \frac{1}{\gamma}\mathbb{E}_{F_X}[x^\gamma] \tag{8}$$

That is, γ is the only decision-relevant parameter when risks are ex-post fair. The following proposition characterizes the DM's optimal behavior in ex-post fair choice problems, where the DM must select a portfolio of Arrow securities that pay both agents the same amount in every state.

Definition 3.2. ***Choice Environment: EPF*** *There are two equally likely states of the world, A and B. The DM selects a portfolio, (x_A, x_B), of Arrow securities A and B from the set $\{x_A, x_B : m = p_A x_A + p_B x_B\}$, for $m, p_A, p_B > 0$. Either security pays both the DM and the DM's counterpart the exact same amount in both states: $y_A = x_A$ and $y_B = x_B$.*

Proposition 3.2. *In the choice environment EPF, DM's utility-maximizing behavior is given by:*

$$\frac{x_A}{x_B} = \left(\frac{p_A}{p_B}\right)^{\frac{1}{\gamma-1}} \tag{9}$$

when $\gamma < 1$. If $\gamma = 1$, $x_s = 1_{\{p_s < p_{s'}\}} \frac{m}{p_s}$ for $s, s' \in \{A, B\}$.

The proof is immediate.

Ex-Post Unfair (but Ex-Ante Fair) Risks

Next, we concentrate on the class of *ex-ante fair and ex-post unfair* lotteries. In this case, according to our notion of ex-ante fairness, the DM's personal risks (modeled by F_X) and those of her counterpart (F_Y) are the same ($F_X = F_Y$). However, the actual realizations of payoffs do not need to be egalitarian. For tractability, we concentrate on the sub-class where payoffs are perfectly negatively correlated, and there are only two equally-likely states. That is, their fortunes are always opposite unless no risk is taken. Then, we must have $y_A = x_B$ and $y_B = x_A$; otherwise, the two marginal CDFs are not equal.

As we are limiting our analysis to ex-ante fair risks ($D = 0$) and because the Bernoulli utility depends only on the outcomes, the parametric model is consistent with EUT in this environment.[18]

Since in this environment, we have $F_X = F_Y$ and, therefore, $D = 0$, our parametric model obeys EUT.[19] Therefore, preferences can be fully represented by the indifference map in the space of state-contingent payoffs for the DM (x_A, x_B). If we also assume weak risk aversion, it can be

[17] Note, since ex-post equality implies ex-ante equality.

[18] Our model can deviate from EUT only if $g(\cdot)$ varies with D. Clearly, when we allow Bernoulli utility functions to vary across distributions, the model can accommodate deviations from the independence axiom.

[19] In our parameterization, violations to the independence axiom emerge only if $g(\cdot)$ varies with D, but in this case, D is constant.

shown that the |MRS| in the space of personal payoffs takes the following form almost everywhere (i.e., outside the loci of $x_A = x_B$):

$$|MRS|_{x_A,x_B} = \left(\frac{x_A}{x_B}\right)^{\rho-1} \frac{V\left(\frac{x_A}{x_B}\right)(a - s_\Delta\theta\bar{\delta}) + (\bar{a} - s_\Delta\theta\bar{\delta})}{V\left(\frac{x_A}{x_B}\right)(\bar{a} + s_\Delta\theta\bar{\delta}) + (a + s_\Delta\theta\bar{\delta})} \tag{10}$$

where $s_\Delta := \text{sign}(x_A - x_B)$ and

$$V\left(\frac{x_A}{x_B}\right) = \left[\frac{\left(\frac{x_A}{x_B}\right)^{\rho}(a - s_\Delta\theta\bar{\delta}) + (\bar{a} + s_\Delta\theta\bar{\delta})}{\left(\frac{x_A}{x_B}\right)^{\rho}(\bar{a} - s_\Delta\theta\bar{\delta}) + (a + s_\Delta\theta\bar{\delta})}\right]^{\frac{\gamma}{\rho}-1} \tag{11}$$

From this, it follows that there is a convex kink along the 45-degree line given by:

$$\frac{\lim_{x_A \downarrow x_B} |MRS|}{\lim_{x_A \uparrow x_B} |MRS|} = \left[\frac{1 + 2\theta(1-\delta)}{1 - 2\theta(1-\delta)}\right]^2 \geq 1 \tag{12}$$

Which is strictly greater than one if and only if $\theta(1-\delta) > 0$. In words, there is a kink if the DM exhibits ex-post fairness concerns.[20]

The following definition and proposition characterize the optimal behavior in the EPU portfolio choice problem.

Definition 3.3. ***Choice Environment: EPU*** *Consider two equally likely states, A and B, and two Arrow securities associated with each state. The DM chooses a portfolio (x_A, x_B) from the set $\{x_A, x_B : m = p_A x_A + p_B x_B\}$ for $m, p_A, p_B > 0$. If state A realizes, security A pays x_A to the DM and pays x_B to her counterpart. If state B realizes, security B pays x_B to the DM and pays x_A to her counterpart. That is, $y_A = x_B$, and $y_B = x_A$.*

Proposition 3.3. *(i) The optimal behavior in the choice environment EPU satisfies the following implicit equation:*

$$\frac{x_A}{x_B} = \left[\max\left(\min\left(1, \frac{p_A}{p_B}\frac{V\left(\frac{x_A}{x_B}\right)(\bar{a} + \theta\bar{\delta}) + (a + \theta\bar{\delta})}{V\left(\frac{x_A}{x_B}\right)(a - \theta\bar{\delta}) + (\bar{a} - \theta\bar{\delta})}\right), \frac{p_A}{p_B}\frac{V\left(\frac{x_A}{x_B}\right)(\bar{a} - \theta\bar{\delta}) + (a - \theta\bar{\delta})}{V\left(\frac{x_A}{x_B}\right)(a + \theta\bar{\delta}) + (\bar{a} + \theta\bar{\delta})}\right)\right]^{\frac{1}{\rho-1}} \tag{13}$$

where $V(\cdot)$ is given by Equation 11.
(ii) For $\frac{p_A}{p_B} \in [1 - \theta\bar{\delta}, 1 + \theta\bar{\delta}]$ the optimal choice is the safe portfolio $x_A = x_B$.

[20]See the detail in the Proof to Proposition 3.3 in Appendix A and a depiction of comparative statics in Figures 16a and 16b, also in the Appendix.

(iii) If $\theta\bar{\delta} = 0$ and $\rho = \gamma$, Equation 13 becomes Equation 9. The optimal behavior in the EPU context coincides with the optimal behavior in the EPF context.

See Proof A.3 in Appendix A.

Proposition 3.3 has three parts. The first item describes the optimal portfolio. The second and third items follow from the first claim and are important characterizations of this portfolio in the vicinity of the safe choice and away from it. Item 2 states that as the degree of concern for ex-post inequality, $\theta\bar{\delta}$, increases, so does the size of the kink: the range of prices for which the DM chooses the safe option ($x_A = x_B$) is larger. In EPF, instead, there was no kink because we assumed a simple CRRA form for the basic risk attitudes. If we instead had used disappointment or loss aversion, we would have also had a kink in the optimal behavior of EPF. Still, the kink in EPU would have been larger as long as ex-post inequality concerns were present. In the limit, when $\theta\bar{\delta} \to 0$ (concerns for ex-post inequality vanish), the kink disappears, and the |MRS| becomes one from the left and right as in EPF. Item 3 speaks of behavior away from the loci of safe choices. Suppose ex-post motives vanish ($\theta\bar{\delta} = 0$). Then the differences in optimal behavior between EPU and EPF depend on the relative strength of social substitution, ρ, and constant relative risk aversion, γ. When they coincide, Equations 13 becomes the (much simpler) Equation 9. That is, the prediction for the EPU context coincides with the one for EPF.

Sharing in Chances

We now present and discuss a choice problem in which the DM faces a direct ex-ante fairness dilemma. The following definition provides a formal description of this context.

Definition 3.4. ***Choice Environment: SiC***

The DM is presented with two fixed, mutually exclusive outcomes $A = (x_A, y_A)$ and $B = (x_B, y_B)$. Outcome A gives a higher payoff to the DM and outcome B to the other agent ($x_A > x_B$ and $y_B > y_A$). In this environment, the DM must choose the probabilities of these two outcomes (q_A, q_B) such that $q_A + q_B = 1$.

Let us denote the chosen probability of outcome A as simply q. In this choice context, it can be shown that:

$$U(q) = qw(g_A(q)) + (1 - q)w(g_B(q)) \tag{14}$$

where

$$w(z) = (1/\gamma)z^{\gamma} \tag{15}$$

$$g_s(q) = (ax_s^\rho + \bar{a}y_s^\rho - s_\rho\theta\bar{\delta}|x_s^\rho - y_s^\rho| - s_\rho\theta\delta D(q))^{\frac{1}{\rho}} \text{ for } s \in \{A, B\} \tag{16}$$

and with

$$\begin{aligned} D(q) =& s_\rho((\bar{y_A}^\rho - y_A^\rho + x_A^\rho - \bar{x_A}^\rho)q^2 + (\bar{x_B}^\rho - x_B^\rho + y_B^\rho - \bar{y_B}^\rho)(1-q)^2 \\ &+ (\bar{x_A}^\rho - \bar{x_B}^\rho)(1-2q)^2 + s_\Delta(\bar{y_B}^\rho - \bar{x_A}^\rho)) \end{aligned} \tag{17}$$

where $\bar{y}_B = \min(y_B, \max(x_A, x_B, y_A))$, $\bar{x}_B = \max(x_B, \min(x_A, y_A, y_B))$, $\bar{y}_A = \max(y_A, \min(x_A, x_B, y_B))$, $\bar{x}_A = \min(x_A, \max(x_B, y_A, y_B))$, and $s_\Delta = sign(y_B - x_B)$. See detail in the first part of the proof to Proposition 3.4 in Appendix A.

The predicted behavior of our model says that a DM who is ex-ante fairness-oriented prefers strict randomization over any of the available sure outcomes in at least some choice problems. The following proposition formally characterizes optimal behavior in this context.

Proposition 3.4. *(i) In the choice environment of sharing in chances (SiC), the DM's optimal behavior satisfies the following implicit equation:*

$$q = \begin{cases} 0 & \text{if } U_q(0) < 0 \\ 1 & \text{if } U_q(1) > 0 \\ \frac{w(g_B(q)) - w(g_A(q)) - w'(g_B(q))g'_B(q)}{w'(g_A(q))g'_A(q) - w'(g_B(q))g'_B(q)} \in (0,1) & \text{otherwise} \end{cases} \tag{18}$$

where $U_q(\cdot)$ is the derivative with respect to q of $U(\cdot)$ given by Equation 14; $w(\cdot)$ is given by Equation 15 and $w'(.)$ is its derivative; and $g_A(q)$ and $g_B(q)$ are given by Equation 16 and $g'_A(q)$, $g'_B(q)$ are their corresponding derivatives with respect to q.
(ii) If $s_\rho > 0$ ($s_\rho < 0$) and $D(q) \leq \min\{|\bar{x_A}^\rho - \bar{y_A}^\rho|, |\bar{x_B}^\rho - \bar{y_B}^\rho|\}$ ($D(q) \geq \min\{|\bar{x_A}^\rho - \bar{y_A}^\rho|, |\bar{x_B}^\rho - \bar{y_B}^\rho|\}$) then $U(q) > \min\{U(A), U(B)\}$ for all q in (0,1).
(iii) If $\theta\delta = 0$, the DM obeys EUT.

See Proof A.4 in Appendix A.

The proposition states that if $\theta\delta > 0$, there will be some choice problems in which the DM strictly prefers randomization over either outcome A or B with certainty ($0 < q^* < 1$). The first item provides a condition when preferences are differentiable. The second item provides a geometric condition that guarantees the local quasi concavity of preferences over social outcomes–ex-ante fairness preferences. The condition is intuitive, as long as the mixture over social events A and B does not generate more ex ante unfairness than ex post fairness, an ex ante oriented DM must strictly prefer to randomize their choices. Otherwise, as stated in the third item, $\theta\delta = 0$ will imply $g_s(.)$ is constant with respect to q. Then, $U(q)$ is strictly linear with respect to q, and utility is maximized at either $q = 0$ or $q = 1$ everywhere except under indifference between A and B.

3.2 An Alternative Parametrization

Consider the following alternative decision utility in the lottery space:

$$U(L) = \delta' u(\mathbb{E}[x], \mathbb{E}[y]) + (1 - \delta')\mathbb{E}[u(x, y)] \tag{19}$$

where the expectation is taken with respect to the joint CDF $F_{X,Y}$ and $u(.)$ is the utility function proposed in Fehr and Schmidt (1999) (the FS model) displayed in Equation 5.

This is the Expected Inequality Aversion (EIA) model that was introduced by Fudenberg and Levine (2012) and axiomatized and named bySaito (2013). The parameter δ' has the same interpretation as our parametric model: it represents the relative importance the individual gives to ex-ante fairness relative to overall fairness considerations. As long as $\delta' \in (0, 1)$, this model is consistent with ex-post and ex-ante fairness concerns. However, the underlying FS model gives the EIA model unrealistic comparative statics to changes in the price of giving and makes it difficult to accommodate risk attitudes. Following the discussion in López Vargas (2015), we propose a generalized EIA (henceforth GEIA). Specifically, we replace the underlying FS utility function with a more flexible form given in Equation 20, which also nests the FS model.

$$u(x, y) = \gamma^{-1} \left[ax^\rho + \bar{a}y^\rho - s_\rho \theta |x^\rho - y^\rho|\right]^{\frac{\gamma}{\rho}} . \tag{20}$$

For comparability, we use the same Greek letters in this model as in our model since their interpretations are roughly parallel. These Greek letters, however, are associated with our model everywhere else in the paper beyond this sub-section.

Interestingly, the GEIA model can be seen as an instance of our general framework in Equation 1, with $W[(p_k, g_k)_k] = \sum_k p_k g_k$ and $g[x, y, F_X, F_Y] = (1-\delta')u(x, y) + \delta' u(\mathbb{E}[x], \mathbb{E}[y])$ for $u(.)$ given by Equation 20.[21] However, the GEIA model does not satisfy the separating Assumption 1. To see this, notice that for lotteries where $Pr[x = y] = 1$ we have $g[x, x, F_X, F_X] = (1 - \delta')\frac{x^\gamma}{\gamma} + \delta' \frac{(\mathbb{E}[x])^\gamma}{\gamma} \neq x$. In sum, in the GEIA model, preferences cannot be disentangled in the sense of Epstein-Zin. Even when we shut down all social considerations, the social parameter δ' still influences the risk attitudes. Further, the specifics of how GEIA entangles risk attitudes and other-regarding preferences imply a testable prediction: more ex-ante fairness-oriented individuals (higher δ') will exhibit higher risk tolerance. We present a thorough study of this model in Appendix I, where we provide a formal expression of the ideas discussed above, an empirical test of its prediction, and an assessment of its predictive power.

[21] Admittedly, this requires extending the definition of our framework to allow functional $D(\cdot)$ to be a more general mapping of the marginal CDFs.

4 Experiment Design

We designed and deployed an experiment implementing the four choice environments studied in the theory section. We collect numerous decisions per subject using graphical interfaces. The method allows us to collect a larger number of choices than standard form-based elicitation, and the quality of the decisions –as measured by common rationality notions– is comparable with prior studies. The detailed view the experiment provides on individuals' behavior makes it possible to (i) measure more accurately the prevalence of ex-post and ex-ante fairness motives in a sample and (ii) conduct a structural, individual analysis identifying all parameters of the model and therefore, characterizing subjects' heterogeneity according to their five elementary attitudes. The detailed choice data for each participant also allows us to contrast our model's predictive power with alternative models. We first describe the design of the choice tasks deployed for each environment. Then, we formulate three broad hypotheses and document the rest of the experimental procedures.

4.1 Choice Task Design

In each session of the experiment, an even number of participants are anonymously paired with one another. Then subjects make multiple decisions in each of the four choice environments we study. The order in which the environments are presented and the order of the choice sets within each environment are randomized. At the end of the session, the computer randomly chooses one member of each pair and calculates payoffs according to her decisions.

Deterministic Giving (DetGiv)

This is the implementation of the choice environment *DetGiv* or *deterministic giving* where deterministic other-regarding preferences are elicited in dictator-like decisions varying the budget and the prices of *giving*. This environment has been studied previously Andreoni and Miller (2002), Charness and Rabin (2002), Fisman et al. (2007), among others. Formally, in each round of this environment, a subject is asked to choose the allocation (x, y) of payoff for themselves and their partner from those satisfying a budget constraint: $p_y y + p_x x = m$. Participants made choices in 50 decision rounds of this kind, where, across rounds, the price of giving p_y/p_x and the size of the budget m varied. The slopes of the budget lines change in steps of 5%, 10%, and 20%, with smaller steps around price equality. The use of small steps around the unit price ratio helps with the identification of individuals' inequality aversion (the strength of the kink at the 45-degree line). The graphical interface is shown in Figure 8b in Appendix B. The set of 50 budget constraints $(p_x/p_y, m)$ used in the experiment is represented in Figure 1 and listed in Table A3 of the Appendix. The predicted behavior is characterized in Proposition 3.1.

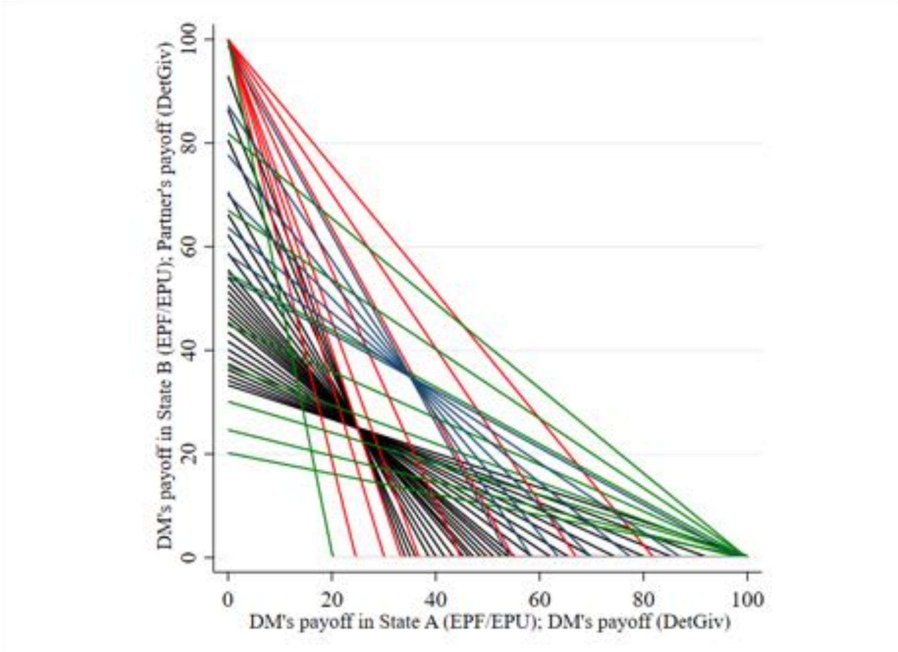


Figure 1: All budget constraints for EPF, EPU, and DetGiv. We designed these budgets to have a wide range of slopes systematically spaced on the logarithm scale. We did so because, in the three tasks, the predictions are piece-wise linear or near-linear (implicit or explicit) functions of the $ln(p_x/p_y)$ for DetGiv and $ln(p_A/p_B)$ in EPF and EPU. There are four subgroups of budget lines (BLs). Black BLs pass through $(25, 25)$, and $ln(p_1/p_2)$ changes in steps of 0.05 if $ln(p_1/p_2) \in [-0.2, 0.2]$, and in steps of 0.1 if $ln(p_1/p_2) \in [-1.1, -0.2] \cup [0.2, 1.1]$. Blue BLs pass through $(35, 35)$, and $ln(p_1/p_2)$ changes in steps of 0.2 if $ln(p_1/p_2) \in [-0.6, 0.6]$. Green BLs pass through $(100, 0)$, and $ln(p_1/p_2)$ changes in steps of 0.2 if $ln(p_1/p_2) \in [-1.6, -0.2]$. Red BLs pass through $(0, 100)$, and $ln(p_1/p_2)$ changes in steps of 0.2 if $ln(p_1/p_2) \in [0.2, 1.6]$. In EPF and EPU, $p_1/p_2 \equiv p_A/p_B$. In DetGiv $p_1/p_2 \equiv p_x/p_y$.

Ex-Post Fair Risks (EPF)

The implementation of the choice environment Ex-Post Fair (EPF) in an experimental setting is novel and elicits risk attitudes free from inequality concerns. The participant selects a portfolio of risky assets that pay her and their counterpart the same amount in every state of the world. Specifically, there are two equally likely states of Nature, A and B, and the decision maker selects a bundle (x_A, x_B), of securities A and B subject to a budget constraint $m = p_A x_A + p_B x_B$. m represents the participant's income in the round, and p_A (p_B) represents the security A (B) price. Security A (B) pays the decision maker and her counterpart x_A (x_B) if state A (B) happens. In short, $y_s = x_s$ for $s \in \{A, B\}$. Decision makers in the experiment face 50 tasks of this kind, where we vary the round income m and the relative prices p_A/p_B across decision rounds.

The graphical interface is shown in Figure 7a in Appendix B. The set of 50 budget constraints $(p_x/p_y, m)$ used in the experiment is represented in Figure 1 and listed in Table A3 of the Appendix. The predicted behavior is characterized in Proposition 3.2.

Ex-Post Unfair Risks (EPU)

This is the implementation of the choice environment with ex-post unfair risks (EPU), and it is also a novel task in experimental settings. This environment aims to identify the strength of ex-post fairness concerns by shutting down all ex-ante inequality concerns (setting $F_X = F_Y$). Everything is the same as in the EPF environment, except now the luck of the two agents is reversed. Suppose the participant chooses a portfolio (x_A, x_B). If state A realizes, the decision maker receives x_A, and her counterpart receives x_B. Similarly, if state B realizes, the decision maker receives x_B, and her counterpart receives x_A. In short, $y_A = x_B$ and $y_B = x_A$. Decision makers also face 50 tasks of this kind, varying income m and the relative prices p_A/p_B across decision rounds. Since the ex-post fairness behavior is identified by comparing choices in this environment and EPF, the budgets used for this EPU environment are the same used in the EPF environment.

The graphical interface is shown in Figure 7b in Appendix B. The set of 50 budget constraints $(p_x/p_y, m)$ used in the experiment is represented in Figure 1 and listed in Table A3 of the Appendix. The predicted behavior is characterized in Proposition 3.3.

Sharing in Chances (SiC)

This task is the implementation of the Sharing in Chances (SiC) choice environment, also known as giving in probabilities. This environment elicits ex-ante fairness motives directly. Each decision round of this task type presents the DM with two fixed, mutually exclusive (materially undominated) outcomes, $A = (x_A, y_A)$ and $B = (x_B, y_B)$. Without relevant loss of generality, we concentrate on cases where DM's payoff is always higher in outcome A, while her partner's payoff is higher in outcome B ($x_A > x_B$ and $y_A < y_B$). For example, $A = (5, 85)$ and $B = (85, 5)$. Participants are asked to allocate probabilities between the two fixed outcomes. Formally, for two given outcomes (x_A, y_A) and (x_B, y_B), each participant is asked to choose a lottery from this set: $\{q\,(x_A, y_A) + (1-q)\,(x_B, y_B) : q \in [0, 1]\}$. Participants in the experiment face 28 decision rounds of this type. The list of outcomes is depicted in Figure 9 and listed in Table A5 in Appendix C . The graphical interface is displayed in Figure 8a of the Appendix B.

4.2 Broad Predictions

Our experiment is designed to facilitate structural analysis at the individual level. However, the model explains important behavioral patterns that we conjecture are prevalent in the population that previous experiments did not measure properly. Consequently, we formulate broad hypotheses regarding these conducts. We also test the broad implication of the GEIA model. Specifically, the first two hypotheses below are validations of our model, while Hypothesis 3 is a broad prediction of the GEIA model.

H1: *A majority of participants exhibit ex-post fairness-seeking behavior: they bear more risk in the EPF environment than in the EPU.*

This hypothesis emerges from an implication of Propositions 3.2 and 3.3: when people exhibit ex-post motives $\theta(1-\delta) > 0$, they stick more to the safe choices in EPU than in EPF ($x_A = x_B$). Away from the $x_A = x_B$ line, the degree of risk tolerance depends on the relative sizes of γ and ρ. We conjecture the net effect will be such that there is more tolerance to risks in EPF than in EPU around the kink and away from it.[22] We argue that prior studies were not exhaustive in measuring risk bearing in a sufficiently wide range of risk-return situations as our experiment.

H2: *A majority of people exhibit ex-ante fairness-seeking behavior.*

We argue that prior experiments (Krawczyk and Le Lec, 2010; Brock et al., 2013, e.g.) were not properly designed to assess the extent to which people seek ex-ante fairness since the set choices presented in their studies were limited. Our experiment presents subjects with one order of magnitude more choices, with a much broader range of incentives. A more precise experiment, like ours, will yield a clearer picture of the prevalence of ex-ante equity motives. And we hypothesize that the actual prevalence is larger than in previous experiments.[23]

H3: Individuals with higher ex-ante fairness motivations (measured by frequent interior choices in the sharing in chances environment (SiC) will tolerate more risk. That is, the DM exhibits behavior closer to risk neutrality.
This is a broad prediction of the alternative model (the GEIA model), by which the ex-ante fairness motives are necessarily negatively correlated with risk aversion. See Proposition I.1 in Appendix I.

Regarding the structural estimation, we hypothesized that the individual evidence is consistent with Propositions 3.1, 3.2, 3.3, 3.4. Finally, we hypothesize that our proposed model predicts observed behavior more accurately than the GEIA.

4.3 Experiment Procedures

We presented participants with 50 decisions in each of the first three choice environments (DetGiv, EPF, and EPU). In the fourth environment, Sharing in Chances (SiC), subjects were presented with 28 decision problems. The order of choice environments was randomized across sessions. Each person received the budget lines in random order, independently from the order presented to other participants. The experiment software was developed at the University of California Santa

[22]This also implies that holding the marginal CDFs equal across players, there is a preference for positively correlated risks over negatively correlated risks.

[23]Naturally, we advise caution regarding the corresponding results since they are from a different subject pool.

Cruz LEEPS Lab, using oTree (Chen et al., 2016). The experiment was conducted online at Universidad del Pacífico's **E2Lab** (Lima). A total of 158 subjects participated. At the beginning of the session, experiment instructions were provided on the computer screen and read aloud. Instructions were extensively discussed and piloted (with students and colleagues) to balance clarity and length. Instruction reading took approximately 15 to 20 minutes. The on-screen instruction page of every choice environment included an interactive trial of the decision box with the explicit message that those choices did not have real consequences for their payoffs. Players were not allowed to communicate with each other, nor did they learn the identity or characteristics of others. Session earnings were based on a randomly chosen decision and payments followed the standard confidential procedures.

5 Results

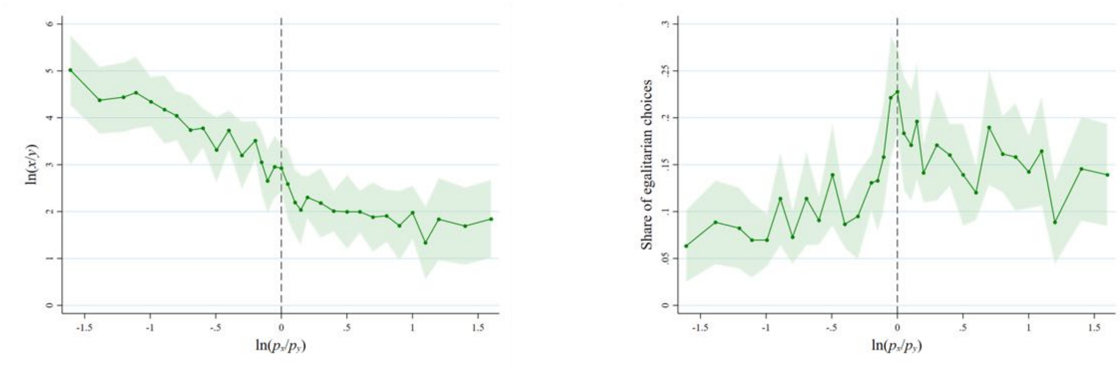
In this section, we present the results of the experiment. We begin with a descriptive analysis, followed by testing the models' broad or aggregate predictions. We then turn to the structural and individual analysis by fitting the proposed parametric model. Last, we carry out a test of predictive power.

5.1 Descriptive Analysis

Deterministic Giving (DetGiv)

As described in sub-section 4.1, in each decision round, we present the DM with a budget line on the two-dimensional space of DM's payoff (x) and their partner's payoff (y). Then, the DM must choose her preferred bundle (x, y). The model's prediction for this task is in sub-sections 3.1.

Figure 2a depicts the average relative demand (x/y, DM's payoff over partner's payoff) as a function of the relative prices (p_x/p_y); both log-transformed. A higher x/y ratio is a more selfish choice. The relative price indicates how much the DM can increase their partner's payoff if they reduce their payoff by one currency unit – i.e. the social benefit of giving one dollar. As expected, the relative demand decreases (DMs give more to their partners) as the social benefit of giving increases. However, as reported in previous papers, most subjects are not impartial but imperfectly selfish: they give themselves most of the available resources: and the average $ln(x/y)$ is clearly above zero (the equality level). Still, a sizable share of choices is egalitarian. Figure 2b presents the share of egalitarian decisions against the ratio of prices (in logs). In general, when the social benefit of giving one currency unit is low ($p_x/p_y < 1$), the share of egalitarian choices is low but rises as the social benefit of giving goes up, with a sharp increase at the point in which the benefit of giving to self and the other is the same, surpassing 20%. When the social benefit of giving is high ($p_x/p_y > 1$), the share of egalitarian choices is higher and remains around 15%.



(a) Ratio of the DM's and partner's payoffs

(b) Share of egalitarian choices

Figure 2: Behavior under Deterministic giving. Panel(a) plots the ratio of $ln(\frac{x}{y})$ against $ln(\frac{p_x}{p_y})$. We impute $\frac{x}{y} = 0.001$ when $x = 0$,and $\frac{x}{y} = 1000$ when $y = 0$. Panel (b) plots the share of egalitarian choices. A choice is considered egalitarian when the difference between the DM's payoff and their partner's is 2 or less.

Ex-Post Fair (EPF) and Ex-Post Unfair (EPU) Risks

The experimental tasks for ex-post fair (EPF) and ex-post unfair (EPU) risk consist of choosing a portfolio of securities whose outcomes are pairs of payoffs for the DM and their counterpart. In EPF, risks are completely fair (there is equality in every state), and in EPU, risks are only ex-ante fair ($F_X = F_Y$) but ex-post unfair. We describe these tasks and their models predictions in sub-sections 3.1 and 4.1.

We use four measures to characterize participants' behavior and inspect how those change with the securities price ratios, p_A/p_B. As the price ratio moves away from one, the opportunity for getting a higher return by taking more risks increases. Although we presented choice problems with price rations below and above one, for ease of exposition, we transformed the data so that security A always has a lower expected return than security B (i.e., $p_A \geq p_B$). In the transformed data, the higher the ratio p_A/p_B, the higher the opportunity of getting a high return (and risk) from buying more security B.[24] The four metrics allow us to describe behavior in three margins where they might have changed: the prevalence of safe choices (kink choices), the responsiveness along the budget lines, and the incidence of high risk-taking (corner choices).

Figure 3 presents the central tendency choices for the four metrics. In Panel (a), we plot the median ratio of relative security holdings, x_A/x_B, against the price ratio. Since a ratio of one means that the choice is safe (no risk is taken), the farther away choices are from one, the more risks the median individual takes. As expected, the more attractive security B becomes relative to A, the lower the relative demand x_A/x_B.

More importantly, EPF tasks generally elicit more risk tolerance than EPU, which is consistent

[24]We do not find evidence of systematic asymmetry in choices with respect to the name of the assets, A and B.

with Hypothesis 1. Note that the more striking difference in behavior comes from the median participant sticking to the safe choice ($x_A/x_B = 1$) only under EPU and for a wide range of relative prices. Our model predicts this as the main behavioral difference between EPF and EPU.

In Panel (b), we present a nonparametric measure of risk tolerance defined as the distance, along the budget line, from the safe choice to the chosen portfolio as a proportion of the rational segment of the budget line.[25] Again, the median risk tolerance is higher under EPF than under EPU for all price ratios.

In Panel (c), we display the relative frequency of safe choice –i.e., $x_A = x_B$ approximately.[26] The share of safe choices increases as the reward for risk-taking decreases (i.e., as the price ratio approaches one from the right). More importantly, across all price ratios, the share of safe choices is much higher under EPU than under EPF. Overall, there are 51% more safe choices in EPU than in EPF.

We also inspect whether there is a different behavior in terms of high risk-taking. In Panel (d), we show the percentage of corner choices –i.e., cases where the participant only purchased one asset, buying only the riskiest / highest return available bundle.[27] As theory predicts, we observe a meager percentage of corner choices for price ratios close to one, which means very few subjects (<5%) are almost perfectly risk neutral. As the return of taking risks increases (the price ratio increases), the incidence of corners choices increases in both treatments. This increase is slightly faster in EPF, although the evidence is mild as the confidence intervals for single values of relative prices overlap. When the relative attractiveness of security B is substantial (p_A/p_B is large), almost one in five people chose the corner regardless of the treatment. In sum, the difference in high-risk behavior is mild, mainly because only a small share of subjects are close to being risk neutral.

Overall, we find strong evidence of divergent behavior between EPF and EPU environments in the direction consistent with Hypothesis 1. These differences between EPU and EPF behavior come mainly from a higher incidence of safe choices in EPU and higher responsiveness to return increases in EPF relative to EPU.

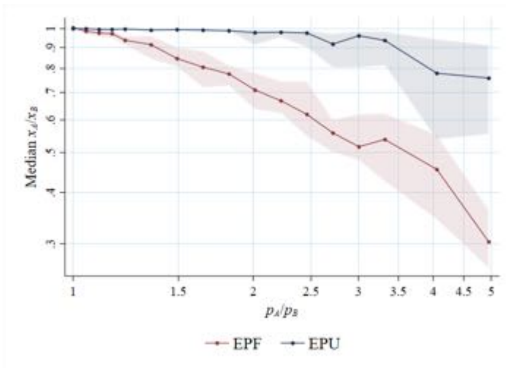
Sharing in Chances (SiC)

Each round of the sharing in chances (SiC) environment presents the DM with two mutually exclusive social outcomes, A and B: (x_A, y_A) and (x_B, y_B). Then, the DM must decide their preferred probabilities for the outcomes so that they add up to one. When people choose probabilities

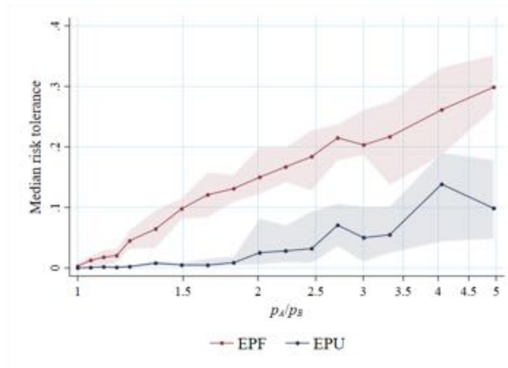
[25] The rational segment of a budget line goes from the safe choice ($x_A = x_B$) to the largest intercept. We call it the "rational segment" because choices outside this segment violate first-order stochastic dominance.

[26] We define safe choice as those $|x_A - x_B| \leq 2$ ECUs.

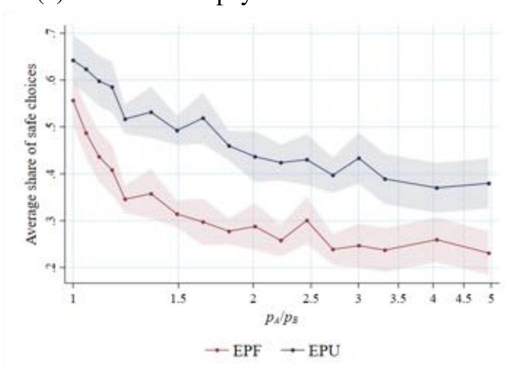
[27] Since the interface is graphical, we need an operational definition of corner choices. We classify a choice as a corner if the distance from the chosen bundle to the nearest intercept is less than 2 ECUs.



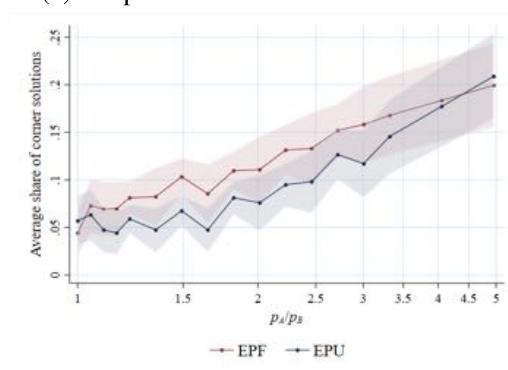
(a) Ratio of own payoffs in states A and B



(b) Nonparametric risk tolerance measure



(c) Share of safe choices



(d) Share of corner solutions

Figure 3: Measures of Risk-Taking under EPF and EPU. The figure plots four different person-choice level measures of risk-taking or avoidance against the price ratio, p_A/p_B (in logarithmic scale in all panels). To take advantage of the symmetric probabilities, we transform the data so that security A always has a lower expected return than security B (i.e., $p_A \geq p_B$). That is, if $p_A > p_B$, we swapped all labels A and B in prices and choices. Panel (a) shows the median ratio of relative security holdings, x_A/x_B (logarithmic scale). We impute $x_A/x_B = 0.001$ when $x_A = 0$ and $x_A/x_B = 1000$ when $x_B = 0$. Panel (b) presents the median of a nonparametric measure of risk tolerance defined as the distance, from the chosen security bundle to the safe choice as a proportion of the rational segment of the budget line (the bundles between the safe choice, $x_A = x_B$, and the largest intercept). Panel (c) displays the relative frequency of safe choices–i.e., those in which $|x_A - x_B| \leq 2$ ECUs. Panel (d) shows the relative frequency of corner choices–i.e., those in which the distance, along the budget line, from the chosen security bundle to the nearest intercept is less than 2 ECUs. All statistics in all panels include the corresponding 95% confidence intervals.

away from zero and one, we say they exhibit ex-ante fairness-seeking behavior or preference for social randomization. The description of the deployed tasks and the theoretical predictions are in sub-sections 4.1 and 3.1, respectively. In these tasks, outcome A gives a material advantage to the DM, and outcome B to the DM's counterpart (i.e., $x_A > y_A$ and $x_B > y_B$). For that reason, we use $\frac{y_B - y_A}{x_A - x_B}$ as a measure of the attractiveness of outcome B. This ratio represents the expected gain for the DM's counterpart relative to the expected DM's loss when we move 1% chance from outcome A to outcome B.

In Figure 4a, we show how the average person assigns a higher probability to outcome B as this outcome becomes more attractive. For example, when the marginal relative gain equals two, the average person assigns a probability of 0.25 to outcome B. Although interesting, this pattern does not necessarily imply people are motivated by ex-ante fairness, as it could still be generated by individuals choosing only A or B with certainty. That is why, in Figures 4b and 4c, we display finer data showing that most people consistently choose interior probabilities –exhibiting ex-ante fairness behavior. In particular, Figure 4b depicts the empirical CDF of $Pr[B]$ for each of the 28 decision rounds. Notably, in most decisions, at least 50% of the people chose non-degenerate probabilities. If we count the number of people with consistent independence (EUT) violations, we realize that it is more than 60% of the people.[28] This figure also shows that people respond to an increased attractiveness of outcome B (as represented by a higher $\frac{y_B - y_A}{x_A - x_B}$) by a higher willingness to choose a $Pr[B]$ higher than 0% and closer to 100%.

Similarly, the bars in Figure 4c show the share of subjects that chose *at least as many inner probabilities* as indicated by the *x*-axis. Lighter bars define inner probabilities as $5\% \leq Pr[B] \leq 95\%$. Darker bars define it as $10\% \leq Pr[B] \leq 90\%$. The figure shows that most subjects consistently depart from EUT in the ex-ante-fairness-seeking sense. Even with a highly conservative assumption that a EUT-behaved agent can have up to four indifferences[29] given the SiC choice tasks we deployed, we still have that almost 70% of all subjects select five or more inner probabilities. Approximately 60% choose inner probabilities in at least half of their decisions. This novel result is one of the main contributions of our paper.

Prior literature, Brock et al. (2013) and Krawczyk and Le Lec (2010), argued that less than 30% of subjects exhibited ex-ante-oriented behavior. However, those studies presented subjects with insufficient choice contexts, and we argue that led to an underestimation of the prevalence of this behavior. For example, mostly selfish individuals only give positive probabilities when the

[28]This slightly depends on how demanding we are in the conditions to label a DM as a "consistent violator" of EUT. For example, someone with a unique non-degenerate probability choice should not be labeled as ex-ante fairness-seeking since that choice could be a product of indifference. Also, how far from 0 and 1 can be considered a non-degenerate probability? We use two conservative thresholds, 5% and 10%.

[29]Note in Figure 9 that we set four sub-sets of choices for the SiC environment. Tasks within each sub-set are constructed so there can be a single indifference among them. That is why having four randomizations is already a conservative number for the maximum number of deviations from EUT.

attractiveness of outcome B is high, a region of choice context that prior studies did not explore.

5.2 Testing Broad Predictions

We formally test the hypotheses presented in Section 4.2. We find evidence favoring the two main hypotheses 1 and 2.

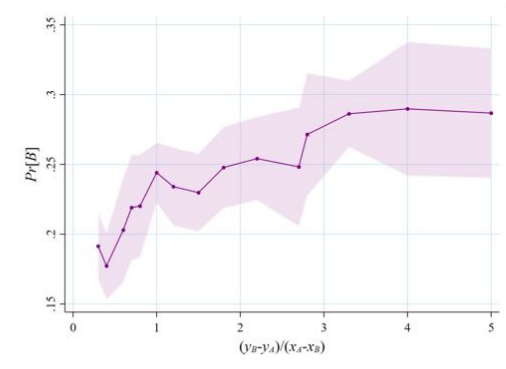
Result 1: People exhibit strong ex-post fairness attitudes. Participants bear more risk in the EPF than in the EPU task. This is mainly driven by more frequent choices of the safe choice in EPU.

First, we test Hypothesis 1, which states that people are more tolerant of risks when payoffs are equal ex-post (state-by-state) than when there is ex-post inequality. We test this by comparing the empirical CDF of the non-parametric risk tolerance measure under EPU and EPF tasks. The analysis shows that subjects bear more risk under EPF than under EPU (See Figure 10 in Appendix E). We test whether the two distributions were equal using a standard Kolmogorov Smirnov (KS) equality-of-distributions test and reject the null hypothesis that the distribution of risk tolerance under EPF has lower values than under EPF (p-value=.0099). Furthermore, we perform a regression at the individual level, where the risk tolerance measure under EPU is the dependent variable, and the same measure under EPF is the independent variable. We then test whether the estimated constant is zero and the estimated coefficient is one. This hypothesis is rejected at the 1% and 5% significance level for 84.18% and 95.57% of participants, respectively. That is, nearly all individuals behave differently in the two environments EPU and EPF. Our experimental evidence provides strong evidence in favor of Hypothesis 1, yielding our first broad result.

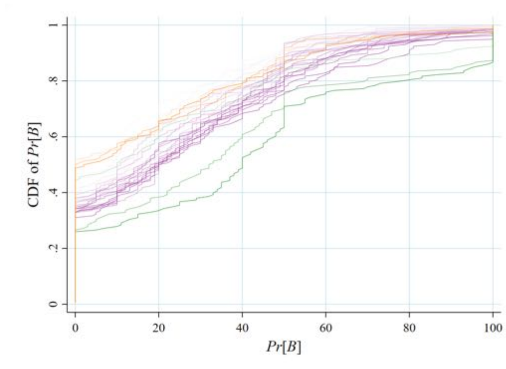
Result 2: Between 50% and 70% of participants consistently exhibit ex-ante fairness-seeking behavior.

Our second hypothesis states that a majority of people exhibit ex-ante fairness-seeking behavior. We argued that the prevalence of this behavior was not correctly measured in the extant literature because subjects were given a limited set of choices where this trait could emerge. As conjectured, we find that when we present people with a richer set of ex-ante dilemmas, most display ex-ante fairness-seeking behavior, even under conservative classifications.

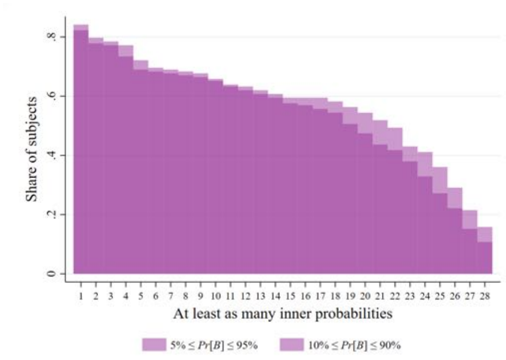
If participants' behavior were to satisfy EUT, their CDF of chosen probabilities in the SiC task would look like a step function. We use that fact to formally compare the distribution of actual choices of each person (the individual empirical CDF) against each possible theoretical (step-function-type) CDF. There are 29 possible theoretical CDFs since we have 28 choices in SiC



(a) Chosen $Pr[B]$



(b) CDF of chosen $Pr[B]$



(c) Share of subjects with at least as many inner $Pr[B]$

Figure 4: Behavior under Sharing in Chances. Panel (a) plots the average $Pr[B]$ at each level of $\frac{y_B - y_A}{x_A - x_B}$, our measure of the relative desirability of outcome B (see text for details). We also include the 95% confidence interval. Panel (b) shows the CDF of the chosen $Pr[B]$ for each of the 28 decision tasks under SiC. We categorize the rounds into three groups, each coded with a different color. The 22 purple functions correspond to rounds where the DM and their partner receive different payoffs in either outcome (see list of outcomes in Appendix C). In the three green functions, both players get the same payoff (25) in outcome B; and in the three orange lines they get the same payoff (25) in outcome A. In all cases, a darker line indicates a context with a higher desirability of outcome B, according to our measure. Each bar in Panel (c) shows the share of subjects that chose at least as many inner probabilities as indicated in the *x*-axis. For each possible number, we plot two overlayed bars: the lighter (darker) one corresponds to a less (more) demanding definition of inner probability, as detailed in the graph's legend.

tasks. We use a demanding standard to classify a subject as exhibiting ex-ante fairness attitudes: we do so when we reject all 29 contrasts with each of the possible theoretical distributions. We find that at least 53% of people consistently violate EUT in the ex-ante fairness-seeking sense.[30]

Result 3: Peoples' ex-ante fairness-seeking behavior is uncorrelated with their degree of risk tolerance.

We now turn to hypothesis 3. As stated above, the GEIA model predicts that more ex-ante-oriented individuals will also be more tolerant of risk. We define non-parametric, individual-level measures of ex-ante fairness concerns using the choices made in the SiC environment. In the first measure, we use the sum of the distances between the chosen probability for outcome B (that is, the outcome that favors the DM's partner) and the nearest extreme probability ($\min\{Pr[B], 1 - Pr[B]\}$). In the second measure, we use the number of choices where $Pr[B] \in [.05, .95]$. For the model-free measure of risk tolerance, we use the non-parametric measure of risk tolerance described in the previous subsection (the distance from the perfectly safe portfolio to the actual choice as a proportion of the rational segment of each budget line). Then, we take the average of those proportions. We use data at the individual level to regress the risk tolerance measure against the measure of ex-ante fairness concern. We find no evidence in favor of the GEIA model's predicted correlation between these measures (coefficient $-.000469$, p-value = 0.381, see Figure 11 in Appendix E).

5.3 Structural Individual Analysis

In this section, the main analysis, we take advantage of the large number of decisions per person (178 choices) to characterize their attitudes towards risks and others in detail and disentangle their motivations. To do so, we fit our model to each person that satisfies a rationality filter. We test whether participants' choices comply with rationality in EPF, EPU, and DetGiv. There is ex-ante fairness in the first two, and in the last one, there is no uncertainty. Therefore, in all those cases, the standard model (with rational, continuous preferences and independence) applies, and preferences can be represented by a continuous (Bernoulli) utility function. Specifically, we measure the extent of violations to the Generalized Axiom of Revealed Preferences (GARP) through Afriat's critical cost efficiency index (CCEI) (Afriat, 1972). The CCEI is bounded between 0 and 1, where 1 reflects perfect satisfaction of GARP. Following Choi et al. (2007); Fisman et al. (2007), we establish a threshold of minimum rationality using a distribution of CCEI based on a sample of 10,000 synthetic agents whose choices are randomly distributed along the budget line (uniformly).

[30]We repeated this test allowing for inner probabilities in the "switching point" to account for the possibility of indifference. We obtain similar results. We also conducted more stringent tests, where we isolated the set of choices where under the null hypothesis of EUT behavior, subjects could have at most one indifference and, therefore, only one inner probability. Again we obtain qualitatively similar results.

We use a threshold of 0.7 since only 10% of random subjects pass this level. Of the 158 subjects, 118 have a CCEI score above the threshold in all three sets of tasks. Figure 12 in Appendix F shows the CCEI distributions for actual participants and those of the synthetic, random agent.

With this rational sub-sample, we jointly estimate the five parameters for each subject using their choices in all four choice environments. That is we use the 178 choices of each participant in all four environments to identify the five parameters of the model. We employ an ad-hoc implementation of minimum-distance estimators that minimizes the individual's weighted sum of absolute prediction errors across rounds of tasks. We use absolute prediction error to reduce the influence of outlier choices and mistakes. The Predictions for each decision round were generated using the first-order conditions presented in subsection 3.1 (in Equations 7, 9 and 13 for the DetGiv, EPF, and EPU, respectively) except for the SiC for which we maximize the utility on the flight.[31,32]

Table 1 and Figure 5 summarize the distributions of estimated parameters. We present our results using the appropriate interpretable transformation of the parameters and provide the percentiles 3, 10, 25, 33, 50, 66, 75, 90, and 97 in each case. Let us first inspect our basic risk attitude parameter, γ. We transform it into $(1-\gamma) \in [0, \infty)$ so it can be interpreted directly as the relative risk aversion coefficient for ex-post fair lotteries. The median subject exhibits an Arrow-Pratt measure $(1-\gamma)$ of 1.7. Also, we find substantial heterogeneity as 50% of subjects are in the $[0.7, 11.06]$ range.

Regarding the selfishness coefficient, a, this can take any value between zero and one, 0.5 if the person is impartial, and one if they are fully selfish. As expected, we only find people with $a \geq 0.5$. The distribution shows two modes (around 0.5 and 1), with the median subject around 0.75.[33]

Parameter ρ is transformed into the *constant elasticity of social substitution*, ESS $= \frac{1}{1-\rho} \in (0, \infty]$. The closer to zero (infinity), the stronger (milder) the convexity of the indifference curves in deterministic giving. Intuitively, EES is a measure of inequality tolerance away from the egalitarian outcome in deterministic contexts. The median ESS equals 1.39, and 50% of subjects are in the $[0.4, 6.87]$ range.

Let us now examine our main parameters. θ measures the degree of overall inequality aversion.

[31]Due to homotheticity, in DetGiv, EPF, and EPU, we can use a ratio (x/y or x_A/x_B) as the endogenous variable. During the experiment, subjects could choose bundles where the numerator or denominator might equal zero. Following Choi et al. (2007), we set the ratios of such extreme bundles as equal to $\omega = 0.001$ when the numerator is equal to zero, and $\omega^{-1} = 1000$ when the denominator is equal to zero; after applying the logarithmic function, these become approximately -6.908 and 6.908, respectively. For consistency, any predicted value in these three tasks that were smaller (larger) than ω (ω^{-1}) was set equal to the corresponding lower or upper limit. Our results are robust to different values of ω.

[32]We estimate the model using data from all four tasks. Then, no obvious relative weight is given to errors from different tasks. For that reason, we let the weights vary across people and set up a procedure that finds the best weight for each person. See Appendix G for details.

[33]If there is no inequality aversion, this median level of selfishness means that the deterministic MRS at the 45-degree line equals 3.

Table 1: Distribution of Elementary Attitudes

Parameter	Percentile								
	3	10	25	33	50	66	75	90	97
Relative risk aversion									
$1-\gamma$	0.007	0.209	0.695	0.989	1.698	8.382	11.064	16.977	21.748
Selfishness									
a	0.500	0.522	0.592	0.636	0.754	0.891	0.928	0.987	1.000
Elasticity of social substitution									
$1/(1-\rho)$	0.116	0.312	0.483	0.691	1.390	2.561	6.873	209.0	2,318
Inequality aversion									
θ	0.000	0.002	0.030	0.052	0.114	0.227	0.265	0.408	0.487
θ/a	0.000	0.002	0.033	0.054	0.156	0.328	0.412	0.760	0.970
Kink in deterministic social preferences at $x = y$									
$\frac{(a+\theta)(\bar{a}+\theta)}{(\bar{a}-\theta)(a-\theta)}$	1.027	1.312	2.725	4.988	21.81	82.51	307.2	5,851	36×10^7
Measures of ex-ante inequality concerns									
δ	0.000	0.003	0.033	0.072	0.209	0.557	0.770	0.970	0.999
δ if $\theta > 0.01$	0.000	0.002	0.028	0.047	0.172	0.527	0.799	0.979	0.999
$\theta\delta$	0.000	0.000	0.001	0.002	0.009	0.031	0.083	0.224	0.297
$\theta\delta/a$	0.000	0.000	0.001	0.003	0.010	0.047	0.147	0.306	0.568
Measures of ex-post inequality concerns									
$1-\delta$	0.001	0.030	0.230	0.428	0.791	0.922	0.967	0.997	1.000
$1-\delta$ if $\theta > 0.01$	0.001	0.021	0.201	0.428	0.828	0.936	0.972	0.998	1.000
$\theta(1-\delta)$	0.000	0.000	0.006	0.013	0.045	0.091	0.150	0.361	0.436
$\theta(1-\delta)/a$	0.000	0.000	0.007	0.015	0.053	0.124	0.253	0.722	0.857
Kink in ex-post unfair lotteries at $x_A = x_B$									
$\left[\frac{1+2\theta(1-\delta)}{1-2\theta(1-\delta)}\right]^2$	1.000	1.001	1.045	1.109	1.439	2.084	3.454	38.38	210.8

Notes: The table reports the results for the 118 rational subjects. All values have been rounded to the third decimal. For θ and δ, we also report the results for the sub-sample of rational subjects with non-negligible inequality aversion ($\theta > 0.01$, $n = 99$).

Since we assume indifference curves for deterministic social preferences are downward sloped (see Section 3), θ can only take values between 0 and 0.5. Approximately 80% of people (99/118) exhibit non-negligible positive values ($\theta > 0.01$) for this coefficient, and the median subject exhibits

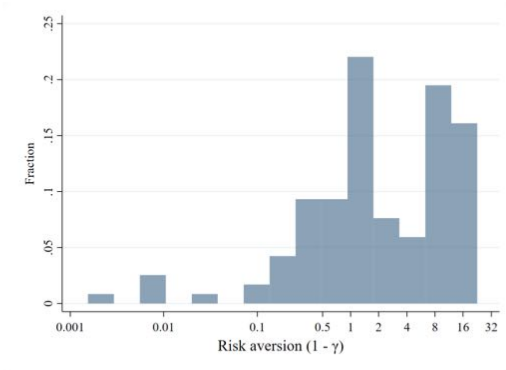
$\theta = 0.11$. That is, the vast majority of subjects do exhibit inequality concerns. Because the scale of deterministic inequality is in the same scale of personal payoff (x^ρ), we can interpret θ/a as the strength of inequality aversion in relation to selfish motives. When we use this transformation, the median subject exhibits a ratio of 0.15. Finally, we can compute the strength of the kink in deterministic choices defined as the MRS above the 45-degree line over the MRS below the 45-degree line. We find that although θ is only 15% of parameter a, for the media subject, still generates a substantial kink – for most people, the indifference curves above $x = y$ are near vertical. The median kink is 21.81, and for more than 3/4 of participants, the kink is above 2.7.

Our model can distinguish whether inequality concerns are ex-ante or ex-post concerns through the parameter δ, which measures the relative weight of ex-ante fairness motives among individuals who exhibit inequality aversion. Over 60% of subjects exhibit clear ex-ante motives $\delta > 0.1$, and one-third give a higher weight to ex-ante inequality than ex-post ($\delta > 0.5$). The median subject exhibit a $\delta = 0.21$. Therefore, according to this measure, ex-post concerns are more important than ex-ante. Since this weight does not matter when $\theta = 0$, we also report in the same table the full size of the ex-ante motives, $\theta\delta$, normalized by the selfishness parameter a. This transformation, $\theta\delta/a$, has a median of 0.01. However, there is important heterogeneity: 25% of participants exhibit a value of 0.14 or higher measure for this measure.

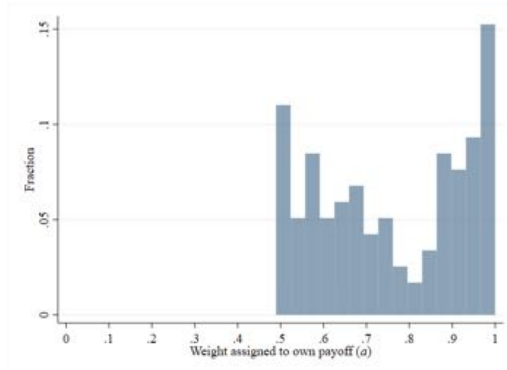
Ex-post inequality concerns are captured by the complementary weight $(1 - \delta)$. Participants exhibit a median ex-post weight of 0.719. Again, the full strength of ex-post motives can be captured by $\theta(1 - \delta)/a$, which displays a median of 0.053. Finally, for ex-post fairness, we have another more interpretable transformation reported in the same table: the kink size in EPU lotteries. This transformation measures how large the range of security prices is where the DM sticks to the safe choice. The median ex-post unfair kink is 1.44, but 50% of participants exhibit kinks ranging from 1.04 to 3.45.

We present two sets of graphs in Figure 6 depicting the actual choices made by two rational subjects (36 and 45) against the predictions made by our model. There are four figures for each of these subjects, one for each task. The actual choices are represented in all of these figures by red dots. In the figures corresponding to the DetGiv, EPF, and EPU, our model's predictions are depicted by blue lines; in the Sharing in Chances figure, our predictions are represented by blue dots, with a dotted gray line highlighting the distance to the actual choice.

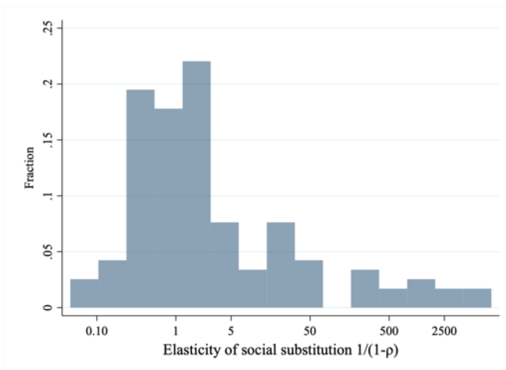
We chose these two subjects because the similarities and differences in their estimated parameter values help to highlight the role these parameters play in generating predictions and because the patterns in their choices exemplify how the data is used to estimate these values. Both subjects assign a mostly impartial weight to their own payoff compared to their partners' (a's are close to 0.5). Moreover, both subjects display high levels of aversion to inequality, as θ is close to 0.5 in both cases. This result is driven by the fact that both subjects chose (relatively) egalitarian bundles in the Deterministic Giving tasks, and their choices are clearly distinct between EPF and EPU.



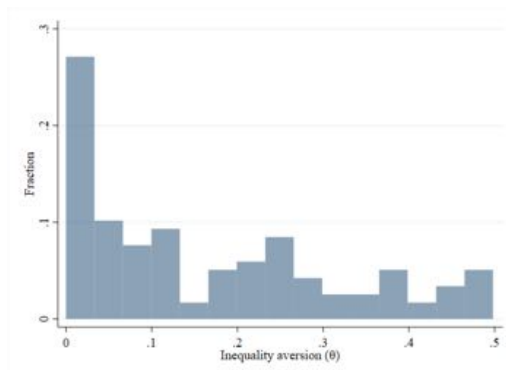
(a) Basic risk aversion



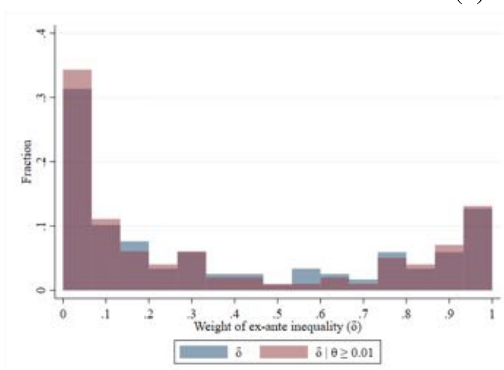
(b) Selfishness



(c) Elasticity of social substitution



(d) Inequality aversion



(e) Ex-ante fairness weight

Figure 5: Distribution of estimated parameters. Calculations include only the 118 subjects in the rational sub-sample. Panel (a) shows the log-scale distribution of $1-\gamma$. Panel (b) the distribution of the degree of selfishness. Panel (c) the elasticity of social substitution in log-scale. Panel (d) presents the distribution of inequality aversion. Panel (e) displays the distribution of δ, the weight given to ex-ante inequality. We plot two distributions: the blue histogram corresponds to the whole rational sub-sample (118 individuals), while the red one corresponds to those among them that have non-negligible inequality aversion ($\theta > 0.01$, 99 individuals).

Despite similar levels of selfishness and inequality aversion, these subjects differ strongly on the weight they assign to ex-ante inequality and to ex-post inequality. Indeed, subject 45 places much more weight on ex-post inequality ($\delta = 0.006$) than subject 36 ($\delta = 0.661$), who cares more strongly about ex-ante inequality. Meanwhile, the latter is more balanced, with a mild bias towards ex-ante inequality.

We also inspect correlations between the parameters and find two relevant correlations between these estimates (See Figure 13 in Appendix H). First, the social curvature parameter (ρ) correlates positively with γ. Second, the degree of inequality aversion shows a strong negative correlation with selfishness. Put differently, more altruistic participants tend to be more inequality-averse.

5.3.1 Model Contrast

We compare our model's predictive power against two versions of the GEIA model: the original EIA with piece-wise linear indifference curves a là Fehr and Schmidt (1999) and the generalized version, where we replace the FS utility with Equation 20 that has the same number of parameters as our model (see Section 3.2). We estimated these two models at the individual level following the same procedure we implemented for our model.

We then conduct out-of-sample predictions and compare the errors of our model against those of GEIAs'. Specifically, for every subject in each iteration, we estimate the models' parameters excluding a randomly chosen 10% of the decision rounds in each task.[34] We use the parameter values for each person-iteration to make predictions in the excluded decision rounds. We store the corresponding choice-level absolute prediction errors. We repeat this process 30 times for each subject, with a new randomly chosen set of decision rounds to be excluded each time.

Our measure of predictive power is the sum of absolute errors of the prediction for each environment. Overall, our model describes subjects' preferences better than both versions of the GEIA model. Results are summarized in Figures, 14, 15 and Table A7 in Appendix I. As expected, the predictive power of the original EIA is rather limited in all four environments, especially in the EPU environment, where, as discussed in 3.2, the GEIA model imposes a structural correlation between risk attitudes and ex-ante fairness concerns. The original EIA model also exhibits higher prediction errors in the other three choice environments. The more flexible GEIA model performs significantly better than the original but underperforms our proposed model. Histograms in Figure 15 compare the proposed model against the GEIA according to the choice-level absolute prediction error distribution for each environment. As expected, the predictive power of deterministic giving is nearly equivalent (both models collapse to the same utility function in deterministic contexts). But our model performs better in the portfolio tasks and in sharing in chances, although the

[34] In the SiC task, 3 out of 28 decision rounds were randomly chosen to be excluded; 5 out of 50 were randomly excluded in the other three task types. We use the same weights for each environment as the estimation that does not exclude any choice.

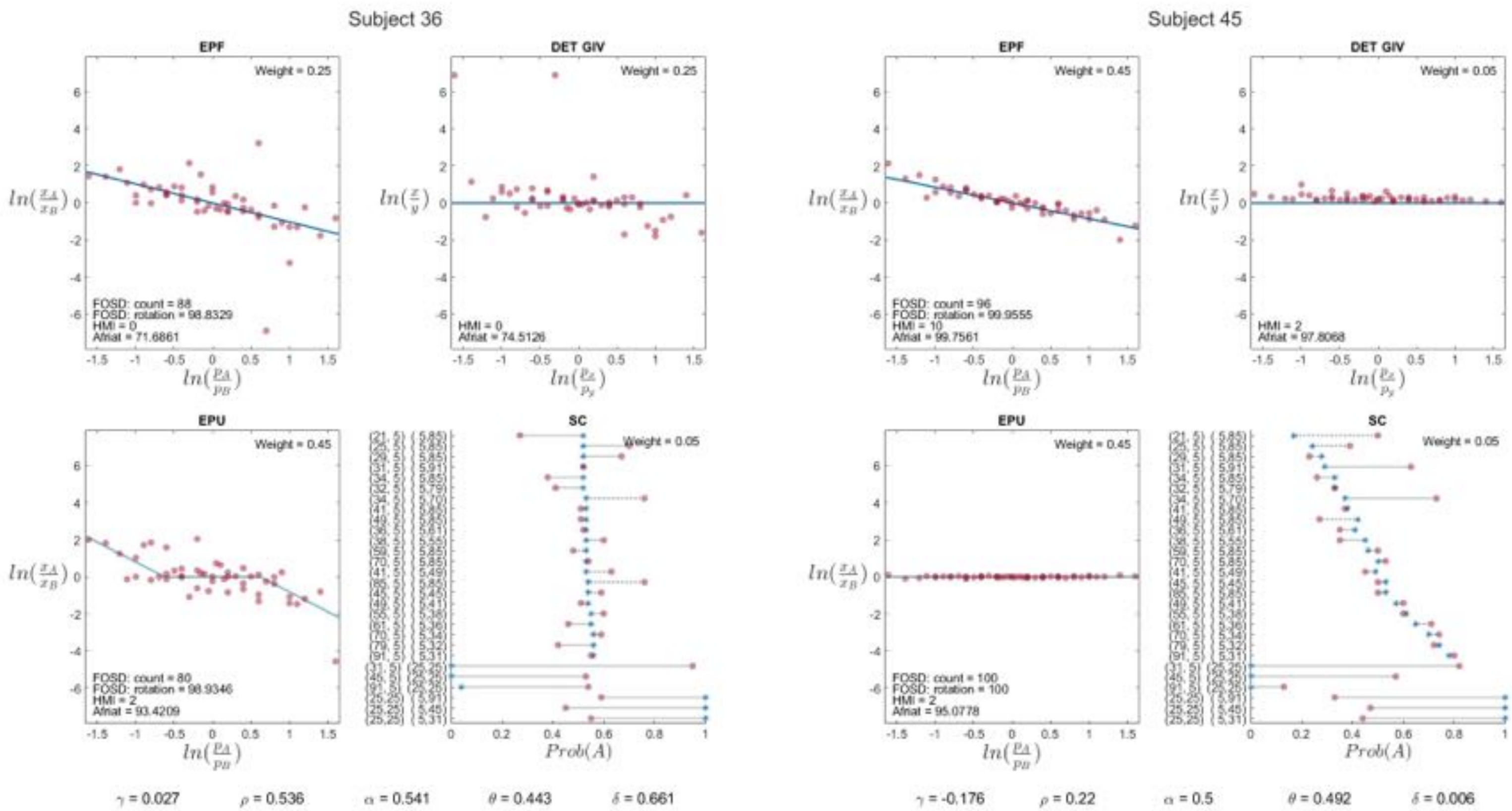


Figure 6: Subjects' Choices: predicted vs. actual decisions. Choices predicted by our model estimation are represented in blue, while actual choices are shown in red dots.

differences are not substantial.

6 Conclusions

We study the interaction and entanglement between risk attitudes and fairness concerns theoretically and experimentally. We proposed a utility model that contemplates risk attitudes and other-regarding preferences in a flexible yet principled manner.

To study the implications of this model, we deploy a laboratory experiment that uses convex choice sets to elicit many choices per participant over completely fair risks, ex-post unfair risks, and deterministic and probabilistic giving. This allows us to test three broad predictions. First, we find that ex-post inequality aversion impacts risk attitudes substantially. People's tendency to choose safely goes up 50% when they face ex-post unfair risks versus ex-post fair ones. Second, we document that most participants (50 to 70%, depending on the measurement approach) consistently exhibit ex-ante fairness-seeking behavior. The prevalence of this behavior is higher than previously suggested, partly because our choice context is richer, so we can better identify the incidence of this attitude. Third, we test one of the main predictions of an alternative model parameterization, the GEIA model. Namely, ex-ante fairness motivations are positively linked with risk tolerance. We do not find evidence that supports this implication of the alternative model.

We perform a structural estimation of our model at the individual level. To our knowledge, we are the first to precisely measure the relative weight of ex-ante and ex-post fairness-seeking motives; and the first to characterize each individual in five fundamental dimensions of preferences: basic risk attitudes, selfishness/altruism, non-aversive fairness concerns (the elasticity of social substitution), overall inequality aversion, and the relative importance of ex-post and ex-ante fairness motives. This empirical exercise also corroborates our first two main findings: most human participants exhibit positive levels of the models' parameters that generate both ex-post and ex-ante fairness-seeking behavior. Finally, we compare our model to the alternative model, the EIA model, and its generalization (GEIA) in terms of predictive power by conducting out-of-sample prediction exercises. We find that our model outperforms both models.

References

Adam, M. T., Kroll, E. B., and Teubner, T. (2014). A note on coupled lotteries. *Economics Letters*, 124(1):96–99.

Afriat, S. N. (1972). Efficiency estimation of production functions. *International economic review*, pages 568–598.

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

Andreoni, J. and Sprenger, C. (2012a). Estimating time preferences from convex budgets. *American Economic Review*, 102(7):3333–56.

Andreoni, J. and Sprenger, C. (2012b). Risk preferences are not time preferences. *American Economic Review*, 102(7):3357–76.

Bolton, G., Brandts, J., and Ockenfels, A. (2005). Fair procedures: Evidence from games involving lotteries. *The Economic Journal*, 115(506):1054–1076.

Bolton, G. and Ockenfels, A. (2000). ERC: A theory of Equity, Reciprocity, and Competition. *The American Economic Review*, 90(1):166–193.

Bolton, G. and Ockenfels, A. (2010). Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states: Comment. *The American Economic Review*, 100(1):628–633.

Borah, A. (2013). Other-regarding preferences, concerns for others' opportunities and violations of stochastic dominance: A choice-theoretic analysis. Working paper. Retrieved from https://ssrn.com/abstract=2608933.

Brennan, G., González, L. G., Güth, W., and Levati, M. V. (2008). Attitudes toward private and collective risk in individual and strategic choice situations. *Journal of Economic Behavior & Organization*, 67(1):253–262.

Brock, J. M., Lange, A., and Ozbay, E. Y. (2013). Dictating the risk: Experimental evidence on giving in risky environments. *The American Economic Review*, 103(1):415–437.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Cappelen, A. W., Konow, J., Sørensen, E. Ø., and Tungodden, B. (2013). Just luck: An experimental study of risk-taking and fairness. *The American Economic Review*, 103(4):1398–1413.

Chakravarty, S., Harrison, G. W., Haruvy, E. E., and Rutström, E. E. (2011). Are you risk averse over other people's money? *Southern Economic Journal*, 77(4):901–913.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *The quarterly journal of economics*, 117(3):817–869.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree: An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, 97(5):1921–1938.

Cox, J., Friedman, D., and Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1):17–45.

Dillenberger, D., Gottlieb, D., and Ortoleva, P. (2020). Stochastic impatience and the separation of time and risk preferences.

Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.

Fafchamps, M., Kebede, B., and Zizzo, D. (2015). Keep up with the winners: Experimental evidence on risk taking, asset integration, and peer effects. *European Economic Review*, 79(C):59–79.

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

Fehr, E. and Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism – experimental evidence and new theories. In Kolm, S.-C. and Ythier, J. M., editors, *Handbook of the Economics of Giving, Altruism and Reciprocity*, volume 1, chapter 8, pages 615 – 691. Elsevier.

Fisman, R., Kariv, S., and Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, 97(5):1858–1876.

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

Gamba, A., Manzoni, E., and Stanca, L. (2017). Social comparison and risk taking behavior. *Theory and Decision*, 82(2):221–248.

Gaudeul, A. (2013). Social preferences under uncertainty. Jena Economic Research Papers 2013-024, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Gaudeul, A. (2015). Social Preferences under Risk: Ex-Post Fairness vs. Efficiency. Jena Economic Research Papers 2015-010, Friedrich-Schiller-University Jena.

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

Harrison, G. W., Lau, M. I., Rutström, E. E., and Tarazona-Gómez, M. (2012). Preferences over social risk. *Oxford Economic Papers*, 65(1):25–46.

Jaramillo, M. and López Vargas, K. (2019). Interpersonal comparisons and risk attitudes: An artefactual field experiment. *Economics Letters*, 179:16 – 18.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291.

Karni, E. and Safra, Z. (2002). Individual sense of justice: a utility representation. *Econometrica*, 70(1):263–284.

Karni, E., Salmon, T., and Sopher, B. (2008). Individual sense of fairness: an experimental study. *Experimental Economics*, 11(2):174–189.

Krawczyk, M. and Le Lec, F. (2010). 'Give me a chance!' An experiment in social decision under risk. *Experimental economics*, 13(4):500–511.

Kreps, D. M. and Porteus, E. L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica: journal of the Econometric Society*, pages 185–200.

López Vargas, K. (2015). Risk attitudes and fairness: Theory and experiment. In López-Vargas, K., editor, *Essays in Behavioral and Experimental Economics* (Phd Dissertation), chapter 1, pages 1–37. University of Maryland. Retrieved from https://drum.lib.umd.edu/handle/1903/16902.

Meier, S. (2007). A survey of economic theories and field evidence on pro-social behavior. In Frey, B. S. and Stutzer, A., editors, *Economics and Psychology: A Promising New Cross-Disciplinary Field*, pages 51–88. MIT Press, Cambridge.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of economic behavior & organization*, 3(4):323–343.

Rohde, I. M. and Rohde, K. I. (2011). Risk attitudes in a social context. *Journal of Risk and Uncertainty*, 43(3):205–225.

Saito, K. (2013). Social preferences under risk: Equality of opportunity versus equality of outcome. *The American Economic Review*, 103(7):3084–3101.

Schwerter, F. (2013). Social reference points and risk taking. Bonn Econ Discussion Papers 11/2013, Graduate School of Economics, Bonn. Retrieved from https://www.econstor.eu/handle/10419/92981.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The american economic review*, 67(2):76–90.

Trautmann, S. T. (2009). A tractable model of process fairness under risk. *Journal of Economic Psychology*, 30(5):803–813.

Trautmann, S. T. and Vieider, F. M. (2012). Social influences on risk attitudes: Applications in economics. In *Handbook of risk theory*, pages 575–600. Springer.

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

Van Koten, S., Ortmann, A., and Babicky, V. (2013). Fairness in risky environments: Theory and evidence. *Games*, 4(2):208–242.

Zame, W. R., Tungodden, B., Sørensen, E. Ø., Kariv, S., and Cappelen, A. W. (2020). Linking social and personal preferences: Theory and experiment. *Working Paper*.

Appendix
For Online Publication

A Proofs

A.1 Proof to Example 3.2

Proof. From equation 6, first we set $\rho = 1$:

$$u(x, y) = (a - s_\Delta \theta)x + (\bar{a} + s_\Delta \theta)y$$

Next we set $a = 1 + \frac{\alpha-\beta}{2}$ and $\bar{a} = 1 - \left(1 + \frac{\alpha-\beta}{2}\right) = \frac{\beta-\alpha}{2}$:

$$u(x, y) = \left(1 + \frac{\alpha - \beta}{2} - s_\Delta \theta\right) x + \left(\frac{\beta - \alpha}{2} + s_\Delta \theta\right) y$$
$$u(x, y) = x + \left(\frac{\alpha - \beta}{2} - s_\Delta \theta\right) x + \left(\frac{\beta - \alpha}{2} + s_\Delta \theta\right) y$$
$$u(x, y) = x + \frac{\alpha - \beta}{2}(x - y) + s_\Delta \theta(y - x)$$
$$u(x, y) = x + \left(\frac{\alpha - \beta}{2} - s_\Delta \theta\right)(x - y)$$

Next we set $\theta = \frac{\alpha+\beta}{2}$:

$$u(x, y) = x + \left(\frac{\alpha - \beta}{2} - s_\Delta \frac{\alpha + \beta}{2}\right)(x - y)$$

Now let's divide the problem by cases, first let's suppose $x \geq y$, which means $s_\Delta = 1$:

$$u(x, y) = x + \left(\frac{\alpha - \beta}{2} - \frac{\alpha + \beta}{2}\right)(x - y)$$
$$u(x, y) = x - \beta(x - y)$$

Finally, we suppose $x < y$, which means $s_\Delta = -1$:

$$u(x, y) = x + \left(\frac{\alpha - \beta}{2} + \frac{\alpha + \beta}{2}\right)(x - y)$$
$$u(x, y) = x - \alpha(y - x)$$

Therefore, if $\rho = 1$, $a = 1 + \frac{\alpha-\beta}{2}$ and $\theta = \frac{\alpha+\beta}{2}$, we get that our model collapses to:

$$u(x, y) = \begin{cases} x - \beta(x - y) & \text{if } x \geq y \\ x - \alpha(y - x) & \text{if } x < y \end{cases} \tag{21}$$

Which is equivalent to:

$$u(x,y) = x - \beta(x-y)^+ - \alpha(y-x)^+$$

And this expression is exactly Equation 5, the FS99 model. □

A.2 Proof to Proposition 3.1

Proof. To prove the first numeral, suppose $x > y$ and $\rho > 0$. Then, $F_{X^\rho}(t) - F_{Y^\rho}(t) = 0$ for $t < y^\rho$ or $t > x^\rho$, and $F_{X^\rho}(t) - F_{Y^\rho}(t) = 1$ for $t \in [y^\rho, x^\rho]$. Thus, $\int |F_{X^\rho} - F_{Y^\rho}|^2 dt = x^\rho - y^\rho$. This also holds for the cases where $x \leq y$ and $\rho \leq 0$. Similarly, when in the other two cases ($x < y$, $\rho > 0$) and ($x \geq y$, $\rho \leq 0$), we have $\int |F_{X^\rho} - F_{Y^\rho}|^2 dt = y^\rho - x^\rho$. Together, these imply $\int |F_{X^\rho} - F_{Y^\rho}|^2 dt = |x^\rho - y^\rho|$. This means Equation 2 can be written as:

$$U(F) = \frac{1}{\gamma}\mathbb{E}\left[ax^\rho + \bar{a}y^\rho - s_\rho\theta\left|x^\rho - y^\rho\right|\right]^{\frac{\gamma}{\rho}} \tag{22}$$

Dropping the expectation (there is no randomness), rearranging the inside of the brackets, and using the appropriate increasing transformation, these deterministic preferences can be written as:

$$u(x,y) = \left[(a - s_\Delta\theta)x^\rho + (\bar{a} + s_\Delta\theta)y^\rho\right]^{\frac{1}{\rho}} \tag{23}$$

where $s_\Delta = \text{sign}(x-y)$. This is a utility that is differentiable a.e. It is only non-differentiable along the 45-degree line, where the optimal choice is a kink type. Everywhere else, optimal behavior can be characterize with standard first order conditions. It is useful to calculate the |MRS|:

$$|\text{MRS}| = \left(\frac{a - s_\Delta\theta}{\bar{a} + s_\Delta\theta}\right)\left(\frac{x}{y}\right)^{\rho-1} \tag{24}$$

The limit of the |MRS| when x/y converges to 1 from above is $\frac{a-\theta}{\bar{a}+\theta}$ ($s_\Delta = 1$). Similarly, the limit of the |MRS| when x/y converges to 1 from below is $\frac{a+\theta}{\bar{a}-\theta}$ ($s_\Delta = -1$).

That is, there is a kink along $x = y$ if and only if $\theta > 0$. If we assume for simplicity that indifference curves are downwards sloping ($\theta < \min\{a, \bar{a}\}$), the form of the |MRS| means that indifference curves are flatter below the 45-degree line ($x > y$) than above it ($x < y$), as $\frac{a+\theta}{\bar{a}-\theta} > \frac{a-\theta}{\bar{a}+\theta}$. Since $\rho - 1 \leq 0$, the |MRS| is always weakly decreasing with respect to $\frac{x}{y}$.

All together, this is sufficient to characterize the DM's optimal solution,

$$\frac{x}{y} = \begin{cases} \left(\frac{p_x}{p_y}\frac{\bar{a}+\theta}{a-\theta}\right)^{\frac{1}{\rho-1}} & \text{if } \frac{p_x}{p_y} < \frac{a-\theta}{\bar{a}+\theta} \\ 1 & \text{if } \frac{p_x}{p_y} \in \left[\frac{a-\theta}{\bar{a}+\theta}, \frac{a+\theta}{\bar{a}-\theta}\right] \\ \left(\frac{p_x}{p_y}\frac{\bar{a}-\theta}{a+\theta}\right)^{\frac{1}{\rho-1}} & \text{if } \frac{p_x}{p_y} > \frac{a+\theta}{\bar{a}-\theta} \end{cases} \tag{25}$$

which can be written as:

$$\frac{x}{y} = \left[\min\left\{\max\left\{1, \frac{p_x}{p_y}\frac{\bar{a}-\theta}{a+\theta}\right\}, \frac{p_x}{p_y}\frac{\bar{a}+\theta}{a-\theta}\right\}\right]^{\frac{1}{\rho-1}} \tag{26}$$

Finally, for the data fitting, it is convenient to use logarithms and express the optimal solution $z \equiv \ln(x/y)$ as a symmetric, piece-wise linear function of $p \equiv \ln(p_x/p_y)$.

$$z = \frac{1}{\rho - 1} \min\left\{\max\left\{0, p + \ln(\bar{a} - \theta) - \ln(a + \theta)\right\}, p + \ln(\bar{a} + \theta) - \ln(a - \theta)\right\} \tag{27}$$

□

A.3 Proof to Proposition 3.3

Proof. Since $D = 0$ in this task, we can write the utility function as:

$$U = \frac{1}{\gamma}\frac{1}{2}\left[a x_A^\rho + \bar{a} x_B^\rho - s_\rho \theta \bar{\delta} |x_A^\rho - x_B^\rho|\right]^{\frac{\gamma}{\rho}} + \frac{1}{\gamma}\frac{1}{2}\left[a x_B^\rho + \bar{a} x_A^\rho - s_\rho \theta \bar{\delta} |x_B^\rho - x_A^\rho|\right]^{\frac{\gamma}{\rho}} \tag{28}$$

Therefore, we can group x's inside the brackets and rewrite it as:

$$U = \frac{1}{2\gamma}\left(\left[(a - s_\Delta \theta\bar{\delta}) x_A^\rho + (\bar{a} + s_\Delta \theta\bar{\delta}) x_B^\rho\right]^{\frac{\gamma}{\rho}} + \left[(a + s_\Delta \theta\bar{\delta}) x_B^\rho + (\bar{a} - s_\Delta \theta\bar{\delta}) x_A^\rho\right]^{\frac{\gamma}{\rho}}\right) \tag{29}$$

where $s_\Delta = \text{sign}(x - y)$. We know that, outside the 45-degree line, preferences are differentiable and optimality is achieved with tangency. The |MRS| thus takes the following form everywhere, except when $x_A = x_B$:

$$\frac{U_{x_A}}{U_{x_B}} = \left(\frac{x_A}{x_B}\right)^{\rho-1} \frac{V(a - s_\Delta\theta\bar{\delta}) + (\bar{a} - s_\Delta\theta\bar{\delta})}{V(\bar{a} + s_\Delta\theta\bar{\delta}) + (a + s_\Delta\theta\bar{\delta})} \tag{30}$$

where

$$V = \left(\frac{\frac{x_A^\rho}{x_B^\rho}(a - s_\Delta\theta\bar{\delta}) + (\bar{a} + s_\Delta\theta\bar{\delta})}{\frac{x_A^\rho}{x_B^\rho}(\bar{a} - s_\Delta\theta\bar{\delta}) + (a + s_\Delta\theta\bar{\delta})}\right)^{\frac{\gamma}{\rho}-1} \tag{31}$$

The limit of |MRS| when $\frac{x_A}{x_B}$ converges to 1 from above ($s_\Delta = 1$) is $\frac{(1-2\theta\bar{\delta})}{(1+2\theta\bar{\delta})} < 1$; similarly, its limit when $\frac{x_A}{x_B}$ converges to 1 from below ($s_\Delta = -1$), is $\frac{(1+2\theta\bar{\delta})}{(1-2\theta\bar{\delta})} > 1$. Note that the limit of V in either case is 1. With the assumption that weak risk aversion holds (indifference curves in the state contingent space are convex), these limits imply a kink in the indifference curves at the 45-degree line. That is, there is a range of relative prices for which the consumer maximizes at the kink located at the 45-degree loci (where $x_A = x_B$).

The previous analysis implies the following optimal choice:

$$\frac{x_A}{x_B} = \begin{cases} \left(\frac{p_A}{p_B}\frac{V_+(\bar{a}+\theta\bar{\delta})+(a+\theta\bar{\delta})}{V_+(a-\theta\bar{\delta})+(\bar{a}-\theta\bar{\delta})}\right)^{\frac{1}{\rho-1}} > 1 & \text{if } \frac{p_A}{p_B} < \frac{(1-2\theta\bar{\delta})}{(1+2\theta\bar{\delta})} \\ 1 & \text{if } \frac{p_A}{p_B} \in \left[\frac{(1-2\theta\bar{\delta})}{(1+2\theta\bar{\delta})}, \frac{(1+2\theta\bar{\delta})}{(1-2\theta\bar{\delta})}\right] \\ \left(\frac{p_A}{p_B}\frac{V_-(\bar{a}-\theta\bar{\delta})+(a-\theta\bar{\delta})}{V_-(a+\theta\bar{\delta})+(\bar{a}+\theta\bar{\delta})}\right)^{\frac{1}{\rho-1}} < 1 & \text{if } \frac{p_A}{p_B} > \frac{(1+2\theta\bar{\delta})}{(1-2\theta\bar{\delta})} \end{cases} \tag{32}$$

where

$$V_+ = \left(\frac{\frac{x_A^\rho}{x_B^\rho}(a - \theta\bar{\delta}) + (\bar{a} + \theta\bar{\delta})}{\frac{x_A^\rho}{x_B^\rho}(\bar{a} - \theta\bar{\delta}) + (a + \theta\bar{\delta})} \right)^{\frac{\gamma}{\rho}-1} \qquad V_- = \left(\frac{\frac{x_A^\rho}{x_B^\rho}(a + \theta\bar{\delta}) + (\bar{a} - \theta\bar{\delta})}{\frac{x_A^\rho}{x_B^\rho}(\bar{a} + \theta\bar{\delta}) + (a - \theta\bar{\delta})} \right)^{\frac{\gamma}{\rho}-1} \tag{33}$$

The previous statement can be summarized as:

$$\frac{x_A}{x_B} = \left[\max\left(\min\left(1, \frac{p_A}{p_B} \frac{V_+(\bar{a} + \theta\bar{\delta}) + (a + \theta\bar{\delta})}{V_+(a - \theta\bar{\delta}) + (\bar{a} - \theta\bar{\delta})} \right), \frac{p_A}{p_B} \frac{V_-(\bar{a} - \theta\bar{\delta}) + (a - \theta\bar{\delta})}{V_-(a + \theta\bar{\delta}) + (\bar{a} + \theta\bar{\delta})} \right) \right]^{\frac{1}{\rho-1}} \tag{34}$$

□

A.4 Proof to Proposition 3.4

Proof. (Assuming differentiability) Utility takes the following form:

$$U(q) = qw(g_A(q)) + (1 - q)w(g_B(q)) \tag{35}$$

where

$$w(z) = (1/\gamma)z^\gamma \tag{36}$$

$$g_s(q) = (ax_s^\rho + \bar{a}y_s^\rho - s_\rho\theta\bar{\delta}|x_s^\rho - y_s^\rho| - s_\rho\theta\delta D(q))^{\frac{1}{\rho}} \text{ for } s \in \{A, B\} \tag{37}$$

The DM chooses q, the probability of state A (with $1-q$ as the probability of state B). $w(g_s(q))$ for $s \in \{A, B\}$ is:

$$w(g_s(q)) = (1/\gamma)(ax_s^\rho + (1 - a)y_s^\rho - s_\rho\theta(1 - \delta)|x_s^\rho - y_s^\rho| - s_\rho\theta\delta D(q))^{\frac{\gamma}{\rho}} \text{ for } s \in \{A, B\} \tag{38}$$

Recall that $D(F_X, F_Y) = \int_{-\infty}^{\infty}(F_{X^\rho}(t) - F_{Y^\rho}(t))^2 dt$. F_{X^ρ} and F_{Y^ρ}, the marginal distributions of x^ρ and y^ρ, depend on the probability of each state; in other words, they depend on q. $D(q)$ is thus the only element inside $w(g_A(q))$ and $w(g_B(q))$ that depends on q.

To see exactly how $D(q)$ depends on q, we must first state the functional form of $D(q)$. Because $x_B \leq x_A$ and $y_A \leq y_B$ in all choice problems presented to the DM, the ordering of the four payoffs will always be one of the following:

$$\begin{aligned} &x_B \leq x_A \leq y_A \leq y_B && y_A \leq y_B \leq x_B \leq x_A \\ &y_A \leq x_B \leq x_A \leq y_B && x_B \leq y_A \leq y_B \leq x_A \\ &x_B \leq y_A \leq x_A \leq y_B && y_A \leq x_B \leq y_B \leq x_A \end{aligned}$$

By taking the integral with respect to q^2 over the difference of marginal distributions, D can be expressed the following way:

$$\begin{aligned} D(q) =& s_\rho((\bar{y_A}^\rho - y_A^\rho + x_A^\rho - \bar{x_A}^\rho)q^2 + (\bar{x_B}^\rho - x_B^\rho + y_B^\rho - \bar{y_B}^\rho)(1 - q)^2 \\ &+ (\bar{x_A}^\rho - \bar{x_B}^\rho)(1 - 2q)^2 + s_\Delta(\bar{y_B}^\rho - \bar{x_A}^\rho)) \end{aligned} \tag{39}$$

where

$$\begin{aligned}
\bar{y}_B &= min(y_B, max(x_A, x_B, y_A)) \\
\bar{x}_B &= max(x_B, min(x_A, y_A, y_B)) \\
\bar{y}_A &= max(y_A, min(x_A, x_B, y_B)) \\
\bar{x}_A &= min(x_A, max(x_B, y_A, y_B)) \\
s_\Delta &= sign(y_A - x_A)
\end{aligned}$$

Its first derivative with respect to q is the following:

$$\begin{aligned}
D'(q) = & s_\rho(2(\bar{y_A}^\rho - y_A^\rho)q - 2(\bar{x_B}^\rho - x_B^\rho)(1-q) + 2(x_A^\rho - \bar{x_A}^\rho)q - 2(y_B^\rho - \bar{y_B}^\rho)(1-q) \\
& - 4(\bar{x_A}^\rho - \bar{x_B}^\rho)(1-2q))
\end{aligned} \tag{40}$$

The DM thus faces the following FOC:

$$q = \frac{w(g_B(q)) - w(g_A(q)) - w'(g_B(q))g'_B(q)}{w'(g_A(q))g'_A(q) - w'(g_B(q))g'_B(q)} \tag{41}$$

where

$$w'(g_s(q)) = (g_s(q))^{\gamma-1} \text{ for } s \in \{A, B\}$$

$$g'_s(q)) = \frac{1}{\rho}(ax_s^\rho + (1-a)y_s^\rho - s_\rho\theta(1-\delta)|x_s^\rho - y_s^\rho| - s_\rho\theta\delta D(q))^{\frac{1}{\rho}-1}(-s_\rho)\theta\delta D'(q) \text{ for } s \in \{A, B\}$$

The utility function is quasi-concave with respect to q, which can only take values in the range of $[0, 1]$.[35] This implies that either utility is maximized with $0 < q < 1$ (with $U_q(0) > 0$ and $U_q(1) < 0$), or it is maximized at either $q = 0$ (with $U_q(0) \leq 0$) or $q = 1$ (with $U_q(1) \geq 0$). Thus, the DM's optimal choice can be summarized in the following expression:

$$q = \begin{cases} 0 & \text{if } U_q(0) < 0 \\ 1 & \text{if } U_q(1) > 0 \\ \frac{w(g_B(q)) - w(g_A(q)) - w'(g_B(q))g'_B(q)}{w'(g_A(q))g'_A(q) - w'(g_B(q))g'_B(q)} \in (0,1) & \text{otherwise} \end{cases} \tag{42}$$

□

Proof. (Not assuming differentiability) Fix two deterministic events S = A, B. Assume A≥ B and notice

$$U(S) = \frac{1}{\gamma}\left[ax_s^\rho + \bar{a}y_s^\rho - s_\rho\theta|x_s^\rho - y_s^\rho|\right]^{\frac{\gamma}{\rho}}$$

[35]Strictly speaking, there are small regions of the parameters and choice space where the utility is not quasiconcave. We ran simulations and checked that quasiconcavity holds for all relevant purposes. That is, the relationship $U(\hat{q}) \geq \min[U(q_1), U(q_2)]$, where $\hat{q} = \lambda q_1 + (1-\lambda)q_2$, holds for all possible values of $q_1, q_2 \in [0, 1]$ in steps of 0.1 and all possible values of $\lambda \in [0.1, 0.9]$ in steps of 0.1, and for all of the (x_A, y_A, x_B, y_B) bundles used in the 28 SiC rounds of our experiment. We did this for 243 different sets of parameters, comprised of all combinations of the 25th, 50th, and 75th percentiles of each of the five parameters as estimated for our model for the rational sub-sample of subjects.

for S=A,B. Define,

$$\tilde{U}(S) = ax_s^{\rho} + \bar{a}y_s^{\rho} - s_{\rho}\theta|x_s^{\rho} - y_s^{\rho}|.$$

Clearly, $U(S) = \frac{1}{\gamma}[\tilde{U}(S)]^{\frac{\gamma}{\rho}}$ and $U(S) \geq U(S')$ iff $\tilde{U}(S) \geq \tilde{U}(S')$. Let $d_s = |x_s^{\rho} - y_s^{\rho}|$, then adding and subtracting from inside $s_{\rho}\theta(\cdot)$, we can rewrite $U(q) = U(qA + (1-q)B)$ as

$$U(q) = \frac{1}{\gamma}\left(q\left[\tilde{U}(A) - s_{\rho}\theta\delta\left(D(q) - d_A\right)\right]^{\frac{\gamma}{\rho}} + (1-q)\left[\tilde{U}(B) - s_{\rho}\theta\delta\left(D(q) - d_B\right)\right]^{\frac{\gamma}{\rho}}\right).$$

Now, let $k_S(q) = s_{\rho}\theta\delta\left(D(q) - d_B\right)$ and using $U(B) = U(qB + (1-q)B)$, the critical inequality that we need to evaluate is whether $U(q) > U(B)$ which is equivalent to

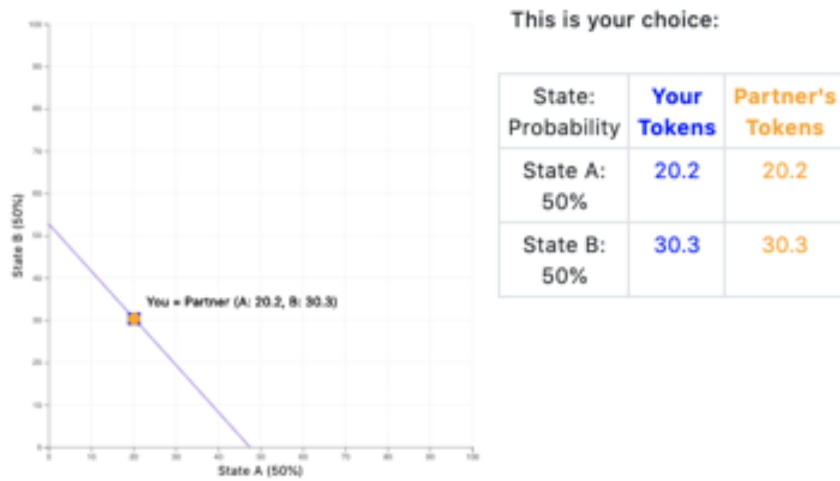
$$\frac{1}{\gamma}\left(q\left[\tilde{U}(A) - k_A(q)\right]^{\frac{\gamma}{\rho}} + (1-q)\left[\tilde{U}(B) - k_B(q)\right]^{\frac{\gamma}{\rho}}\right) > \frac{1}{\gamma}\left(q\left[\tilde{U}(B)\right]^{\frac{\gamma}{\rho}} + (1-q)\left[\tilde{U}(B)\right]^{\frac{\gamma}{\rho}}\right).$$

This inequality can be rewritten as

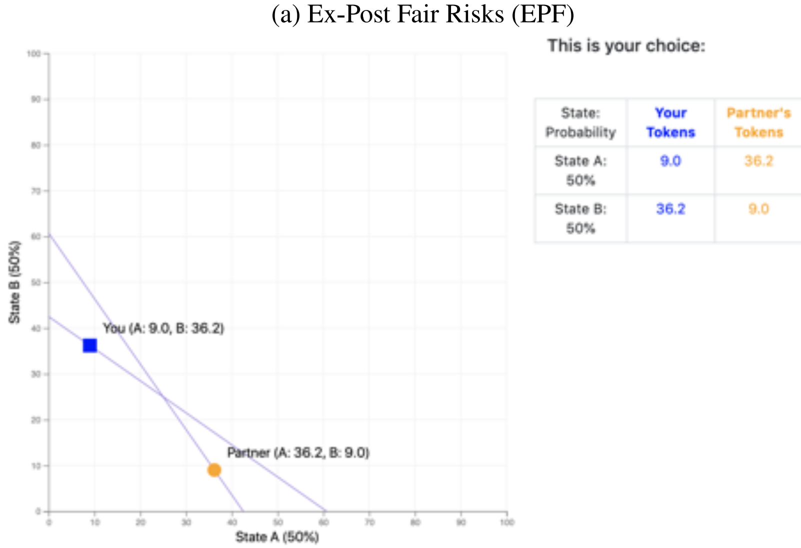
$$q\left(\left[\tilde{U}(A) - k_A(q)\right]^{\frac{\gamma}{\rho}} - \left[\tilde{U}(B)\right]^{\frac{\gamma}{\rho}}\right) > (1-q)\left(\left[\tilde{U}(B)\right]^{\frac{\gamma}{\rho}} - \left[\tilde{U}(B) - k_B(q)\right]^{\frac{\gamma}{\rho}}\right).$$

This inequality describes all instances over which our model predicts a preference for ex-ante preferences (quasi-concavity of social preferences). For our second item it is sufficient to notice that this inequality holds trivially when $k_B(q), k_A(q) \leq 0$. It is also important to note that our sharing in chances environment implicitly assumes $d_s > 0$ for some s. □

B Experimental Design: Graphical Interface

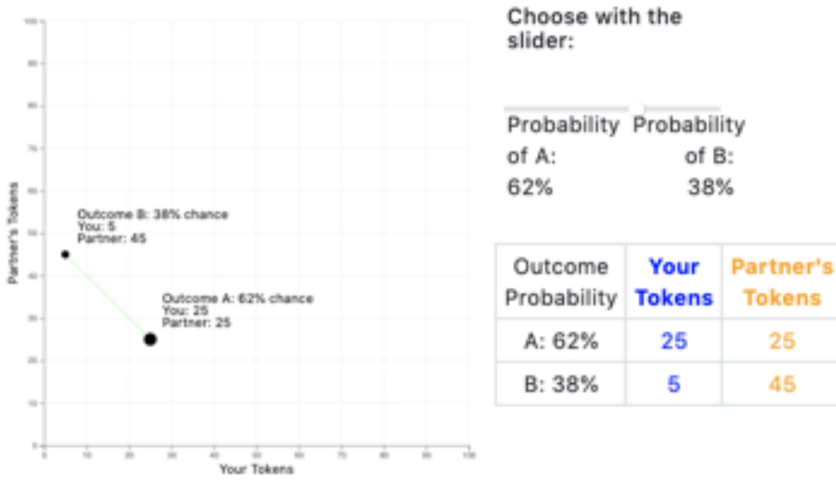


(a) Ex-Post Fair Risks (EPF)



(b) Ex-Post Unfair Risks (EPU)

Figure 7: Experiment Interface for Tasks EPF and EPU. In Panel (a) shows the interface for EPF where participants make a decision by dragging the blue circle (DM portfolio) or the orange square (the partner's portfolio) along the budget line on the plane of State A payoffs - State B payoffs. These two shapes always move together. In Panel (b), participants make a decision in the EPU task by dragging the blue circle (DM portfolio) or the orange square (the partner's portfolio) along the budget line. These two shapes move along their corresponding budget, always mirroring each other so that $y_A = x_B$ and $y_B = x_A$. For the sake of clarity, besides the graphical representation, a table at the top right corner of the decision screen also shows the current choice in text form.



(a) Sharing in Chances (SiC)

(b) Deterministic Giving (DetGiv)

Figure 8: Experiment Interface for Sharing in Chances and Deterministic Giving. In Panel (a), we show the interface for the SiC task. The two fixed outcomes A and B are represented "(me, other)" plane. Participants make a decision about the probability of each state using the slider at the top right corner of the decision box. In Panel (b), participants make a decision in the DetGiv task by dragging the red square along the budget line. The axes represent DM's payoff (x axis) and the DM's partners (y axis). For the sake of clarity, besides the graphical representation, a table at the top right corner of the decision screen also shows the current choice in text form.

C Experimental Design: Choice contexts

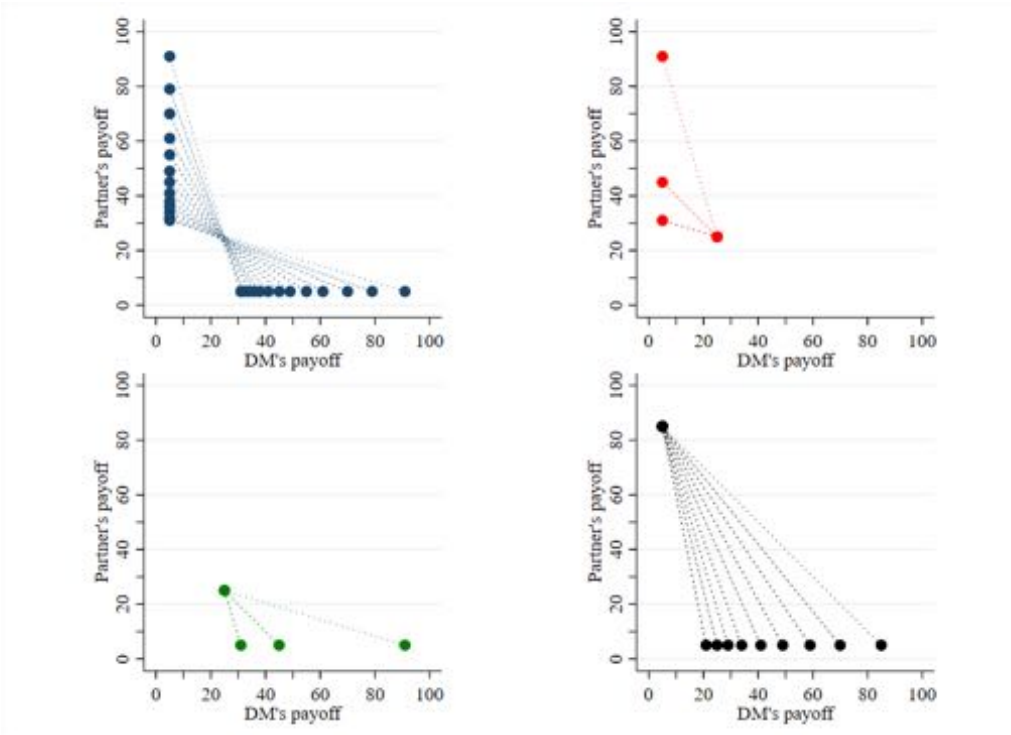


Figure 9: Outcomes defining each choice in SiC environment. Each decision task in SiC is defined over two fixed outcomes $(x_A, y_A); (x_B, y_B)$. In the figures above, each choice problem is represented by two circles connected by a dotted line. In the top left panel, we show the 13 choices that share the following feature. If the DM chooses $q = 50\%$, then the expected value of the outcome equals $(25, 25)$. The top right panel shows three choices where the materially convenient outcome (A) is always perfectly fair $x_A = y_A = 25$. The bottom left panel shows three choices where the materially least convenient outcome (B) is always perfectly fair $x_B = y_B = 25$. Finally, the bottom right panel shows nine choice problems where outcome B remains constant $(x_B, y_B) = (5, 85)$; so does $y_A = 5$, and we only vary x_A.

D Descriptive Analysis

Table A2: Deterministic Giving: summary statistics by round

p_x/p_y	m	Median($\frac{x}{y}$)	Egalitarian choices (%)	Selfish choices (%)	p_x/p_y	m	Median($\frac{x}{y}$)	Egalitarian choices (%)	Selfish choices (%)
0.20	20.19	11.95	6.33	34.81	1.00	70.00	2.12	22.78	24.05
0.25	24.66	9.97	8.86	27.85	1.00	51.28	1.86	18.35	24.68
0.30	30.12	8.48	8.23	30.38	1.11	52.63	1.79	17.09	20.89
0.33	33.32	6.78	6.96	32.28	1.11	54.05	1.84	19.62	22.15
0.37	34.20	5.63	6.33	28.48	1.22	55.54	1.67	18.99	21.52
0.37	36.79	5.48	7.59	31.65	1.22	77.75	2.15	10.76	22.78
0.41	35.16	5.93	11.39	29.75	1.22	100.00	2.08	12.66	22.15
0.45	36.23	4.84	6.33	27.22	1.35	58.75	1.60	17.09	23.42
0.45	44.93	5.67	8.23	31.01	1.49	62.30	1.42	20.89	23.42
0.50	37.41	4.21	11.39	27.85	1.49	87.21	1.56	12.03	17.72
0.55	38.72	4.00	10.13	24.05	1.49	100.00	1.69	15.19	19.62
0.55	54.21	4.01	8.86	29.75	1.65	66.22	1.58	13.92	22.78
0.55	54.88	3.50	8.23	27.22	1.82	70.55	1.47	17.72	24.68
0.61	40.16	2.66	13.92	24.05	1.82	98.77	1.47	10.76	20.89
0.67	41.76	3.47	10.13	24.68	1.82	100.00	1.54	7.59	23.42
0.67	58.46	3.45	8.23	25.95	2.01	75.34	1.35	18.99	22.78
0.67	67.03	3.63	7.59	32.28	2.23	80.64	1.27	20.89	25.32
0.74	43.52	2.63	9.49	24.68	2.23	100.00	1.39	11.39	23.42
0.82	45.47	2.33	17.09	23.42	2.46	86.49	1.16	15.82	21.52
0.82	63.66	2.72	11.39	28.48	2.72	92.96	1.24	16.46	27.22
0.82	81.87	4.50	10.76	29.75	2.72	100.00	1.33	12.03	29.75
0.82	46.52	2.36	13.29	24.05	3.00	100.10	1.19	16.46	24.05
0.90	47.62	1.94	15.82	21.52	3.32	100.00	1.33	8.86	29.11
0.90	48.78	2.08	22.15	23.42	4.06	100.00	1.23	14.56	29.11
1.00	50.00	2.09	22.78	23.42	4.95	100.00	1.16	13.92	41.77

Notes: Columns (4) and (9) show the relative frequency of egalitarian choices per round. A choice is considered egalitarian when the difference between the DM's payoff and their partner's is 2 or less. Columns (5) and (10) show the relative frequency of selfish choices per round, that is, choices where the partner's payoff is 2 or less.

Table A3: Ex-post fair and ex-post unfair risks: summary statistics by round (Part 1)

p_A/p_B	m	Median (x_A/x_B) EPF	Median (x_A/x_B) EPU	Median (Risk tol.) EPF	Median (Risk tol.) EPU	p_A/p_B	m	Median (x_A/x_B) EPF	Median (x_A/x_B) EPU	Median (Risk tol.) EPF	Median (Risk tol.) EPU
0.20	20.19	4.01	1.49	0.334	0.076	1.00	70.00	1.01	1.00	0.003	0.000
0.25	24.66	2.58	1.63	0.240	0.113	1.05	51.28	1.00	1.00	0.001	0.001
0.30	30.12	2.54	1.10	0.262	0.025	1.11	52.63	1.00	1.00	0.000	0.002
0.33	33.32	2.23	1.09	0.234	0.024	1.16	54.05	1.00	1.00	0.001	0.001
0.37	34.20	1.77	1.20	0.173	0.051	1.22	55.54	0.98	1.00	0.007	0.001
0.37	36.79	2.17	1.10	0.240	0.026	1.22	77.75	0.97	1.00	0.012	0.001
0.41	35.16	1.93	1.04	0.214	0.013	1.22	100	0.97	1.00	0.012	0.001
0.45	36.23	1.59	1.02	0.156	0.007	1.35	58.75	0.98	0.99	0.008	0.004
0.45	44.93	1.73	1.17	0.185	0.051	1.49	62.30	0.95	1.00	0.021	0.001
0.50	37.41	1.69	1.03	0.188	0.010	1.49	87.21	0.96	1.00	0.016	0.001
0.55	38.72	1.45	1.01	0.137	0.003	1.49	100	0.98	1.00	0.008	0.002
0.55	54.21	1.54	1.02	0.159	0.008	1.65	66.22	0.88	1.00	0.050	0.000
0.55	54.88	1.48	1.03	0.146	0.011	1.82	70.55	0.88	1.00	0.047	0.001
0.61	40.16	1.42	1.02	0.136	0.006	1.82	98.77	0.91	1.00	0.032	0.001
0.67	41.76	1.33	1.01	0.118	0.004	1.82	100	0.84	0.99	0.065	0.002
0.67	58.46	1.33	1.02	0.117	0.006	2.01	75.34	0.83	0.98	0.065	0.005
0.67	67.03	1.30	1.01	0.108	0.003	2.23	80.64	0.72	0.99	0.109	0.003
0.74	43.52	1.24	1.01	0.092	0.004	2.23	100	0.77	0.99	0.086	0.005
0.82	45.47	1.11	1.00	0.049	0.000	2.46	86.49	0.74	0.99	0.094	0.004
0.82	63.66	1.14	1.01	0.058	0.002	2.72	92.96	0.65	0.87	0.124	0.038
0.82	81.87	1.13	1.00	0.054	0.002	2.72	100	0.60	0.97	0.152	0.009
0.86	46.52	1.07	1.01	0.032	0.004	3.00	100.10	0.62	0.98	0.131	0.005
0.90	47.62	1.08	1.00	0.036	0.003	3.32	100	0.68	0.95	0.101	0.011
0.95	48.78	1.06	1.01	0.027	0.005	4.06	100	0.51	0.90	0.160	0.023
1.00	50.00	1.01	1.00	0.004	0.000	4.95	100	0.38	0.81	0.215	0.039

Notes: Columns (3),(4),(9) and (10) show the median ratio of relative security holdings x_A/x_B. Columns (5),(6), (11) and (12) present the median of a nonparametric measure of risk tolerance defined as the distance, along the budget line, from the chosen security bundle to the safe choice as a proportion of the rational segment of the budget line (the bundles between the safe choice, $x_A/x_B = 1$, and the largest intercept).

Table A4: Ex-post fair and ex-post unfair risks: summary statistics by round (Part 2)

p_A/p_B	m	Safe Choices (%)		Corner Choices (%)		p_A/p_B	m	Safe Choices (%)		Corner Choices (%)	
		EPF	EPU	EPF	EPU			EPF	EPU	EPF	EPU
0.20	20.19	21.52	36.71	17.72	20.25	1.00	70.00	50.63	61.39	5.70	5.06
0.25	24.66	27.85	37.34	15.19	16.46	1.05	51.28	49.37	63.29	8.23	5.70
0.30	30.12	22.15	39.87	17.72	14.56	1.11	52.63	47.47	60.76	6.96	5.06
0.33	33.32	25.32	41.77	14.56	10.76	1.16	54.05	43.67	59.49	8.23	4.43
0.37	34.20	28.48	41.14	12.66	10.13	1.22	55.54	34.18	56.33	7.59	3.80
0.37	36.79	19.62	38.61	16.46	14.56	1.22	77.75	37.97	50.63	7.59	6.33
0.41	35.16	28.48	43.67	10.76	10.13	1.22	100.00	33.54	49.37	8.86	8.86
0.45	36.23	26.58	42.41	10.13	8.23	1.35	58.75	37.34	54.43	8.86	4.43
0.45	44.93	24.05	37.34	11.39	12.03	1.49	62.30	35.44	53.80	9.49	4.43
0.50	37.41	29.11	47.47	8.86	7.59	1.49	87.21	31.01	45.57	8.86	6.33
0.55	38.72	33.54	49.37	8.86	5.70	1.49	100.00	32.28	46.20	12.66	8.23
0.55	54.21	23.42	45.57	13.29	8.23	1.65	66.22	32.91	51.90	6.96	4.43
0.55	54.88	24.05	44.94	9.49	8.23	1.82	70.55	32.91	48.73	8.86	8.23
0.61	40.16	26.58	51.90	10.13	5.06	1.82	98.77	26.58	42.41	13.92	9.49
0.67	41.76	32.91	53.80	8.86	6.33	1.82	100.00	25.95	44.94	11.39	8.86
0.67	58.46	27.85	46.84	8.86	6.96	2.01	75.34	28.48	39.87	13.29	7.59
0.67	67.03	29.11	49.37	13.29	8.23	2.23	80.64	28.48	46.20	14.56	8.23
0.74	43.52	34.18	51.90	7.59	5.06	2.23	100.00	24.05	43.67	16.46	9.49
0.82	45.47	37.97	55.70	7.59	5.70	2.46	86.49	31.65	42.41	15.82	9.49
0.82	63.66	33.54	48.73	8.23	4.43	2.72	92.96	23.42	38.61	17.72	11.39
0.82	81.87	30.38	49.37	8.86	6.33	2.72	100.00	24.05	40.51	13.92	14.56
0.86	46.52	37.97	57.59	5.70	4.43	3.00	100.10	24.05	44.94	17.09	12.66
0.90	47.62	39.87	58.86	6.96	4.43	3.32	100.00	25.32	37.97	15.82	14.56
0.95	48.78	48.10	61.39	6.33	6.96	4.06	100.00	24.05	36.71	21.52	18.99
1.00	50.00	60.76	67.09	3.16	6.33	4.95	100.00	24.68	39.24	22.15	21.52

Notes: Columns (3),(4),(9) and (10) present the relative frequency of safe choices–i.e., those in which $|x_A - x_B| \leq 2$ ECUs. Columns (5),(6),(11) and (12) show the relative frequency of the riskiest choices–i.e., those in which the distance, along the budget line, from the chosen security bundle to the nearest intercept is less than 2 ECUs.

Table A5: Sharing in Chances: summary statistics by round

(x_A, y_A)	(x_B, y_B)	$\frac{y_B - y_A}{x_A - x_B}$	Avg. $Pr[B]$ (%)	Med. $Pr[B]$ (%)	$Pr[B] \in [5, 95]$ (%)	$Pr[B] \in [10, 90]$ (%)
(31, 5)	(5, 91)	3.30	27.13	20.00	60.13	55.06
(32, 5)	(5, 79)	2.70	24.82	17.00	56.96	52.53
(34, 5)	(5, 70)	2.20	25.67	19.00	59.49	54.43
(36, 5)	(5, 61)	1.80	26.27	20.00	62.66	59.49
(38, 5)	(5, 55)	1.50	24.14	19.50	59.49	55.70
(41, 5)	(5, 49)	1.20	24.79	19.00	64.56	57.59
(45, 5)	(5, 45)	1.00	23.17	16.50	64.56	57.59
(49, 5)	(5, 41)	0.80	22.01	16.00	60.76	53.16
(55, 5)	(5, 38)	0.70	21.91	18.00	60.13	51.90
(61, 5)	(5, 36)	0.60	20.30	13.50	58.86	51.90
(70, 5)	(5, 34)	0.40	18.02	10.50	55.70	49.37
(79, 5)	(5, 32)	0.40	17.44	9.00	51.27	44.30
(91, 5)	(5, 31)	0.30	14.85	0.50	45.57	39.24
(25, 25)	(5, 91)	3.30	19.20	3.00	46.20	41.14
(25, 25)	(5, 45)	1.00	18.20	0.50	46.20	40.51
(25, 25)	(5, 31)	0.30	18.13	0.00	44.94	41.77
(31, 5)	(25, 25)	3.30	41.38	40.00	57.59	54.43
(45, 5)	(25, 25)	1.00	36.82	32.50	55.06	51.27
(91, 5)	(25, 25)	0.30	24.44	10.00	45.57	39.24
(21, 5)	(5, 85)	5.00	28.68	20.50	60.76	55.70
(25, 5)	(5, 85)	4.00	28.98	20.00	61.39	53.80
(29, 5)	(5, 85)	3.30	26.80	21.00	62.031	56.33
(34, 5)	(5, 85)	2.80	27.15	20.00	62.66	57.59
(41, 5)	(5, 85)	2.20	25.16	19.50	58.86	55.70
(49, 5)	(5, 85)	1.80	23.28	16.50	56.96	52.53
(59, 5)	(5, 85)	1.50	21.82	17.00	57.59	53.80
(70, 5)	(5, 85)	1.20	22.02	13.50	56.33	48.73
(85, 5)	(5, 85)	1.00	19.42	11.00	55.70	49.37

Notes: Columns (1) to (2) show the fixed outcomes A and B that define each task in this environment. Columns (4) and (5) reflect the average and the median probability given to state B respectively, while columns (6)-(7) show the share of individuals that choose strict randomization with $Pr[B]$ within the specified ranges. The color coding of the first two columns corresponds to the same color coding in Figure 9. In blue, the 13 tasks with $x_B = y_A = 5$, so that if the DM chooses $Pr[A] = 50\%$, the expected outcome is $(25, 25)$. In red, the three choices where outcome (A) is perfectly fair $x_A = y_A = 25$. In green, the three choices where outcome (B) is always perfectly fair $x_B = y_B = 25$. Finally, in black, the nine tasks where outcome B remains constant $(x_B, y_B) = (5, 85)$; so does $y_A = 5$, and we only vary x_A.

E Testing Broad Predictions

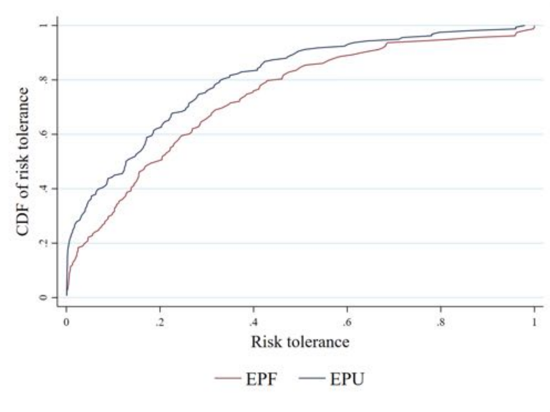


Figure 10: CDF of nonparametric measure of risk tolerance.We perform a KS equality-of-distributions test and reject the null hypothesis that the distribution of risk tolerance under ex-post unfairness has lower values than under ex-post fairness (p-value=.0099).

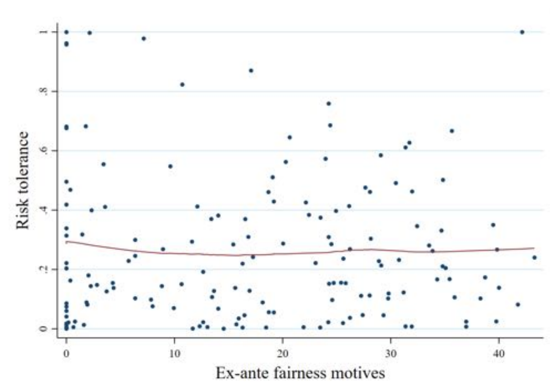
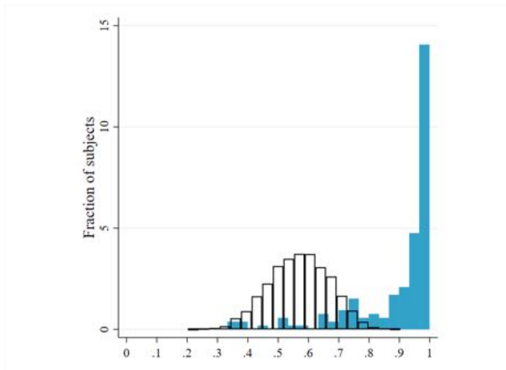
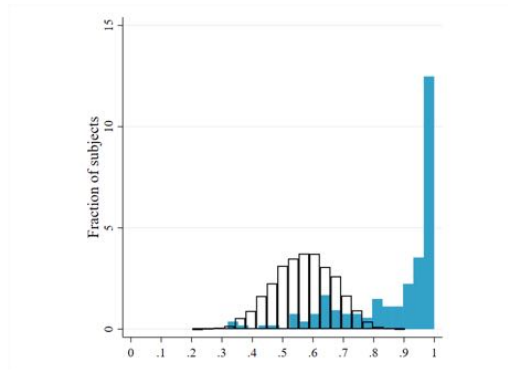


Figure 11: Non parametric regression of risk tolerance on ex-ante fairness motives, at the participant level. Ex-ante fairness concerns is defined as the sum of nearest distances between $Pr[B]$ and either 0 or 1, for each subject. Risk tolerance is the average of risk tolerance demonstrated in EPF tasks.

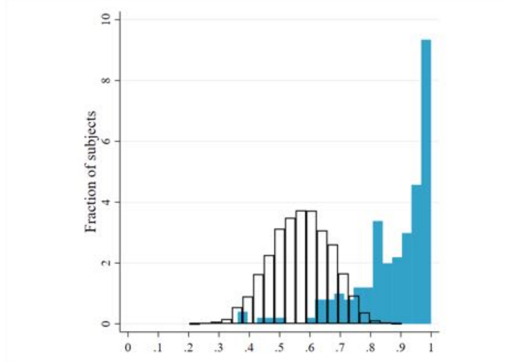
F Rationality Testing



(a) Ex-post fair risks



(b) Ex-post unfair risks



(c) Deterministic giving

Figure 12: Distributions of Afriat's (1972) Critical Cost Efficiency Index (CCEI). The CCEI's distribution of participants is represented in blue and the CCEIS's distribution of 10000 synthetic subjects is represented in white.

G Structural Individual Analysis: Technical Detail

The predictions for each task type were calculated in the following manner:

1. For the **Deterministic Giving** task type, predictions of the decision variable (payoff ratio) were given by the closed-form solution presented in Equation 7. We applied logarithms to both sides of the expression.

2. For the **Ex-Post-Fair Risks** environment, predictions of the decision variable (securities ratio) were given by the closed-form solution presented in Equation 9. We applied the logarithmic function to both sides of the expression.

3. For the **Ex-Post-Unfair, but Ex-Ante-Fair Risks** environment, predictions were given by the implicit function presented in Equation 13. As the decision variable x (securities ratio) cannot be isolated from either side of the equation, we applied the logarithmic function to both sides,

re-write the expression to have the form $\log(x) - \log(f(exp(log(x)), p, \gamma, \rho, a, \theta, \delta)) = 0$, and finally used an optimizer to find the value of x that minimizes the absolute of $\log(x) - \log(f(exp(log(x)), p, \gamma, \rho, a, \theta, \delta))$ given the price ratio p and the parameters.

4. For the **Sharing in Chances** task type, we obtained predictions of the decision variable q ($Pr[A]$) by computing the utility resulting from each possible value of q between 0 and 1 in steps of 0.01 (as given by Equation 2) and choosing the one with the highest utility given the set of outcome bundles and the parameters.

Thus, given a set of values for the parameters, a prediction can be generated for each decision task the subjects face during the experiment, as well as a corresponding absolute prediction error (the absolute value of the difference between the actual choice made and the prediction). We use a nonlinear solver in Matlab to find the parameter values that minimize each subject's weighted sum of prediction errors, with different weights being given to different ask types.[36]

We use 35 different sets of weights; these comprise all combinations that arise when each task type has at least a weight of 5% (meaning that, at most, a task type can weight 85%) and weight increases in intervals of 20 percentage points.[37] As such, we estimate 35 sets of parameter values for every subject. In addition, we compute the sum of prediction errors in each type task that arises from each of these sets of parameter values, thus obtaining 4 error vectors with 35 elements in each. Finally, we choose one set of parameter values among the 35 for every subject to be used in all subsequent analyses. More specifically, we choose the set that minimizes the CES function that takes as inputs the subject's normalized error vectors.

The Matlab solver that we used allows for upper and lower boundaries in each of the parameters being optimized. We chose it to ensure that the estimated values for each of the parameters were contained in the permissible ranges, as described in Section 3. We also defined the objective function in such a way that the solver would avoid parameter value sets where the expression $ax^\rho + \bar{a}y^\rho - s_\rho\theta\left(\bar{\delta}\, d + \delta\, D\right)$ is negative for any $Prob[A]$ under any of the bundles used in the Sharing in Chances task type choices.

[36]The prediction errors of the Sharing in Chances task type were also re-scaled to maintain comparable scale with respect to the other tasks. In particular, this task type only had 28 decision rounds while the others had 50, so each prediction error in this task type was multiplied by a factor of $\frac{50}{28}$. Furthermore, the maximum value of each prediction error in the task type is 1 (as the decision variable q can only take values between 0 and 1), while the maximum value in the other three tasks is approximately 14 (as the logarithm of the decision variable can take values between -6.908 and 6.908); as such, each Sharing in Chances prediction error was also multiplied by a factor of 14.

[37]For example, [DetGiv, EPF, EPU, SiC] = [5%, 25%, 25%, 45%] is one of these sets of weights.

H Parameter Correlation

Table A6: Correlation among estimated parameters

	$\ln(1-\gamma)$	$\ln(1-\rho)$	a	θ	δ
$\ln(1-\gamma)$	1				
$\ln(1-\rho)$	0.478	1			
	(0.0000)				
a	0.125	-0.024	1		
	(0.1761)	(0.7969)			
θ	0.095	0.162	-0.786	1	
	(0.3084)	(0.0788)	(0.0000)		
δ	-0.049	0.22	-0.141	-0.037	1
	(0.5958)	(0.0194)	(0.1286)	(0.6891)	

Notes: The table presents the pairwise correlation coefficients among the five parameters, as calculated according to our model for the 118 rational subjects. All correlation coefficients were rounded to the third decimal. We also include the corresponding p-value below each correlation coefficient in parentheses.

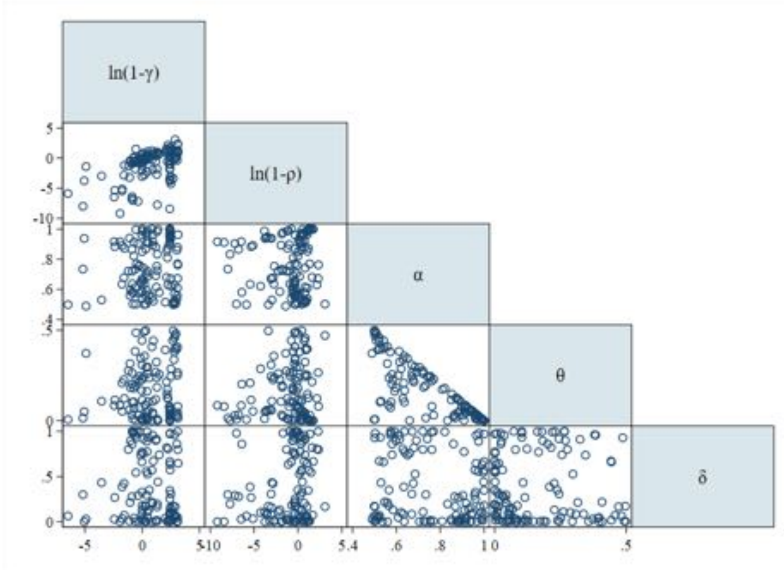


Figure 13: Scatter plot matrix of estimated parameters. The x-axis of each scatter plot is indicated in the right-hand side parameter box and the y-axis is indicated in the parameter box above it. We plot the estimated parameters of the 118 rational subjects.

I The Alternative Parameterization: GEIA Model

The Expected Inequality Aversion model EIA was introduced by Fudenberg and Levine (2012) and Fehr and Schmidt (1999) and axiomatized by Saito (2013), who gave it its name. The DM's decision utility balances: (i) the expected utility of the prospects she faces and (ii) the utility of the expected outcomes. Following the discussion from López Vargas (2015), we present a more general version of the EIA model and, henceforth, use the term *generalized EIA* or GEIA. Specifically, we replace the Fehr-Schmidt utility (FS99) with a more flexible form that nests the FS99 model. Specifically, the DM's decision utility takes the following form:

$$U(L) = \delta u(\mathbb{E}x, \mathbb{E}y) + \bar{\delta}\mathbb{E}u(x, y) \tag{43}$$

where:

$$u(x, y) = \gamma^{-1} \left[ax^{\rho} + \bar{a}y^{\rho} - s_{\rho}\theta|x^{\rho} - y^{\rho}|\right]^{\frac{\gamma}{\rho}} . \tag{44}$$

The parameter δ has the same interpretation as our model: it represents the relative importance the individual gives to ex-ante fairness relative to overall fairness considerations. This model nests the model formalized in Saito (2013) when Equation 44 takes the form of the FS99 model. We use the same Greek letters in this model as in our model since their interpretation is roughly parallel. These Greek letters, however, are associated with our model everywhere else in the paper. The GEIA model and ours have the same number of parameters. In that sense, they have equal footing for contrasting predictive power. Both models give identical predictions in deterministic settings as they collapse to the same model when there is no risk.

The way the GEIA model introduces ex-ante fairness considerations has an implication. The model imposes and predicts a strict positive correlation between ex-ante fairness motives and risk tolerance. In the extreme, a fully ex-ante-motivated decision-maker should be risk neutral. As discussed in the body of the paper, this is an intrinsic property of the model that prevents the disentangling of social and risk attitudes. Furthermore, in this model, the DM's behavior does not obey the independence axiom even when outcomes are ex-post and ex-ante fair with certainty.

I.1 Theoretical Implications

To formalize the implications of the GEIA model and contrast it with ours, extending standard notions of risk aversion to the two-dimensional case is helpful.

Definition I.1. *The* ***certainty equivalent set*** *of a lottery F, denoted as $c(F)$, is the set of bundles (x, y) in $\mathbb{R}^2_+$ that the DM finds equally preferred as the lottery F.*

Definition I.2. *The* ***risk premium*** *of a lottery F is the L^1-distance from $(\mathbb{E}x, \mathbb{E}y)$ to $c(F)$.*

This metric adds up the money the DM is willing to give up from his expected payoff and from the other agent's expected payoff to exchange the lottery for the pair $(\mathbb{E}x, \mathbb{E}y)$. Defining a single-value fair equivalent of a (me, other) bundle is also helpful.

Definition I.3. *The* ***fair equivalent*** *of a pair (x, y), denoted as $\hat{x}$, is the amount that, if given to both agents, makes the DM indifferent between $(\hat{x}, \hat{x})$ and (x, y).*

We say an amount w is the **fair-certainty equivalent** of a lottery F if w is the fair equivalent of an element of the certainty equivalent set of F. The following proposition I.1 summarizes the main implications of the GEIA model.

Proposition I.1. *In the GEIA model specified in Equations 43 and 44:*

(a) The fair-certainty equivalent of a lottery F is increasing in δ.

(b) If $\delta > 0$, the independence axiom does not hold; even in the absence of any ex-ante expected inequality.

The proof is at the end of this Appendix. This proposition formalizes the notion that the GEIA model predicts that stronger ex-ante fairness motivations lead to increased risk tolerance. Our model, in contrast, makes no predictions regarding the relationship between ex-ante fairness concerns and risk tolerance. Whether this prediction holds true is one of the main questions tackled in this paper (see Hypothesis 3 in Section 4.2).

We also characterize the optimal behavior of the GEIA in all four types of choice problems we studied in Section 3.1.

Proposition I.2. *Consider a DM whose preferences are given by the GEIA model specified in Equations 43 and 44:*

1. *If there is no risk and the DM faces the choice problem described in definition 3.1 (DetGiv), the optimal behavior is as described in Proposition 3.1.*

2. *If risks are completely fair (ex-post and ex-ante), and the choice problem is as given in definition 3.2 (EPF), the optimal behavior obeys:*

$$p = \frac{\delta(0.5x + 0.5)^{\gamma-1} + \bar{\delta}x^{\gamma-1}}{\delta(0.5x + 0.5)^{\gamma-1} + \bar{\delta}} \tag{45}$$

where x denotes x_A/x_B, and p represents p_A/p_B.

3. *If risks are ex-post unfair and ex-ante fair in the sense of the choice problem given in definition 3.3 (EPU), the optimal behavior obeys:*

$$\frac{p_A}{p_B} = \frac{\delta f(x_A, x_B) + \bar{\delta}x_A^{\rho-1}\left[g(x_A, x_B)(a - s_\Delta\theta) + h(x_A, x_B)(\bar{a} - s_\Delta\theta)\right]}{\delta f(x_A, x_B) + \bar{\delta}x_B^{\rho-1}\left[g(x_A, x_B)(\bar{a} + s_\Delta\theta) + h(x_A, x_B)(a + s_\Delta\theta)\right]} \tag{46}$$

where

$$f(x_A, x_B) = [0.5x_A + 0.5x_B]^{\gamma-1} \tag{47}$$

$$g(x_A, x_B) = \left[(a - s_\Delta\theta)x_A^\rho + (\bar{a} + s_\Delta\theta)x_B^\rho\right]^{\frac{\gamma}{\rho}-1} \tag{48}$$

$$h(x_A, x_B) = \left[(a + s_\Delta\theta)x_B^\rho + (\bar{a} - s_\Delta\theta)x_A^\rho\right]^{\frac{\gamma}{\rho}-1} \tag{49}$$

4. *If $\delta > 0$, there exist choice problems with ex-ante fairness dilemmas as described in definition 3.4, for which the DM strictly prefers randomization exhibiting ex-ante fairness-seeking behavior.*

The proof is at the end of this Appendix. Note the GEIA model collapses to homothetic preferences in DetGiv and EPF contexts but not under EPU.

I.2 Empirical Implications

The propositions above characterize the influence of the ex-ante parameter even in perfectly egalitarian choice problems. In Hypothesis 3, we state the empirical implication of this feature of the GEIA model. Individuals with higher ex-ante fairness motivations (measured by frequent interior choices in the probabilistic giving task - SiC) will tolerate more risk – i.e., behave closer to risk-neutrality. As described in the results section, we find no evidence in favor of this implication.

We also compare the predictive power of our model against the GEIA's. To do so, we estimate the original EIA and the GEIA models following the same procedure as our model. Since the GEIA model predicts that participants with stronger ex-ante fairness will be more tolerant of risks, in the estimation, people who exhibit both strong ex-ante motives and risk aversion return a positive δ but very negative values of γ. That is, the GEIA model requires people to compensate positive δ values with unreasonably very negative values of γ. The interpretation of γ is, therefore, difficult and entwined with δ.

We then conduct an out-of-sample prediction contrast and compare the errors of our model against those of the EIA and GEIA. Specifically, we estimate the models' parameters for every subject over many iterations. In each iteration, we exclude a randomly chosen 10% of the decision rounds in each task. In the SiC task, 3 out of 28 decision rounds were randomly chosen to be excluded; 5 out of 50 were randomly excluded in the other three task types. We use the parameter values for this person iteration to compute the corresponding predictions of the excluded decision rounds. We store the corresponding choice-level absolute prediction errors. We repeat this process 30 times for each subject, with a new randomly chosen set of decision rounds to be excluded each time.

Our measure of predictive power is the absolute errors in each task. Overall our model describes subjects' preferences better than the EIA and GEIA models, as we often observe lower values of the absolute errors in our model relative to GEIA's. This can be observed in Table A7 and Figures 14 and 15 that characterize the distribution of the choice-level absolute prediction errors by task type. As expected, our model performs the EIA and GEIA models in choice environments that involve risks. The data and Matlab code for this exercise is available from the authors upon request.

Table A7: Descriptive table of MSE under our model and the GEIA model

		$p(10)$	$p(25)$	$p(50)$	$p(75)$	$p(90)$
DetGiv	Our model	0.013	1.311	3.466	6.657	11.161
	GEIA	0	1.544	3.256	6.855	10.175
EPF	Our model	0.505	1.023	1.999	4.943	8.583
	GEIA	0.643	1.315	2.388	4.477	8.936
EPU	Our model	0.261	0.612	1.833	3.599	7.636
	GEIA	0.185	0.905	1.919	3.663	7.661
SiC	Our model	0	1.95	9	15.95	22.705
	GEIA	0	3.65	11.05	16.85	24.97

Notes: The table summarizes the distribution of the out-of-sample prediction error generated by our and the EIA/GEIA models. For each of the 118 rational subjects, we estimated their parameter values excluding 10% of rounds from each of the four tasks; generated predictions for said rounds; and calculated the corresponding prediction error as the absolute value of the difference between the prediction and the actual choice made. We repeated this process 30 times, for a total of $(5+5+5+3)*30 = 540$ choice-level prediction errors for each subject. The table reports five percentiles of this variable in logarithms, divided by task. Prediction errors lower than 0.00001 were capped at this value before applying the logarithmic transformation ($\log(0.00001) \approx -11.51$). The SiC task errors were multiplied by a factor of $14 * \frac{5}{3}$ to maintain the scale consistent across task types.

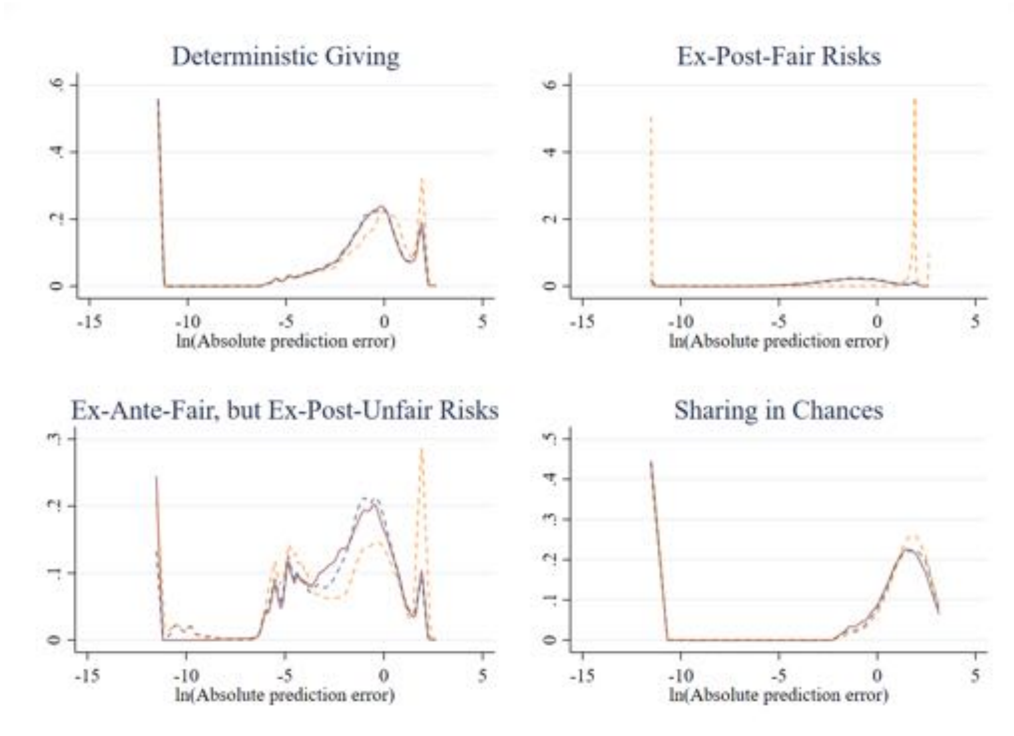


Figure 14: Predictive power comparison (kernel density). The figure plots the choice-level kernel density of the absolute prediction error in logarithms for each of the four task types, with (solid) red lines for our model, (dashed) blue lines for the GEIA model and (dashed) orange lines for the EIA model (equivalent to setting $\gamma = \rho = 1$ in the GEIA model), across all 30 randomizations for each of the 118 subjects in the rational sub-sample. Prediction errors lower than 0.00001 were capped at this value before applying the logarithmic transformation ($\log(0.00001) \approx -11.51$). The SiC task errors were multiplied by a factor of $14 * \frac{5}{3}$ to maintain the scale consistent across task types.

I.3 Proofs to Propositions of the GEIA Model

Proof to Proposition I.1

Proof. Let us define

$$v(\hat{x}) = u(\hat{x}, \hat{x}) \tag{50}$$

By the definition of the fair-certain equivalent and for any lottery F:

$$v(\hat{x}) = U(F) = \delta u(\mathbb{E}x, \mathbb{E}y) + \bar{\delta}\mathbb{E}u(x, y) \tag{51}$$

So we have:

$$\hat{x} = v^{-1}\left[\delta u(\mathbb{E}x, \mathbb{E}y) + \bar{\delta}\mathbb{E}u(x, y)\right] \tag{52}$$

$$\hat{x} = v^{-1}\left[\delta\left(u(\mathbb{E}x, \mathbb{E}y) - \mathbb{E}u(x, y)\right) + \mathbb{E}u(x, y)\right] \tag{53}$$

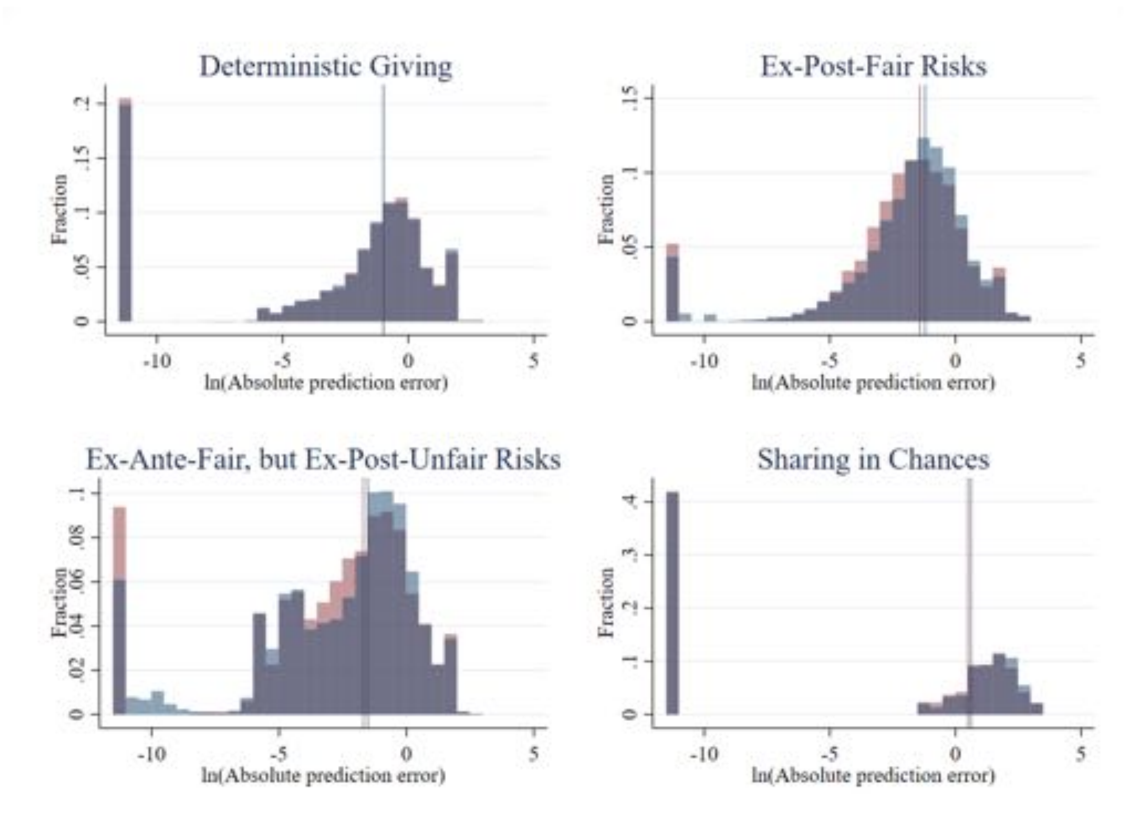


Figure 15: Predictive power comparison. The figure plots the choice-level histogram of the absolute prediction error in logarithms for each of the four task types, with red bars for our model and blue bars for the GEIA model, across all 30 randomizations for each of the 118 subjects in the rational sub-sample. We include vertical lines of the corresponding colors at the median. Prediction errors lower than 0.00001 were capped at this value before applying the logarithmic transformation ($\log(0.00001) \approx -11.51$). The SiC task errors were multiplied by a factor of $14 * \frac{5}{3}$ to maintain the scale consistent across task types.

Since $v(\cdot)$ is an increasing function and by Jensen's inequality:

$$\frac{\partial \hat{x}}{\partial \delta} > 0 \tag{54}$$

□

Proposition I.2

1. DetGiv

It follows immediately from the definition of the utility function:

$$U(L) = \delta u(\mathbb{E}x, \mathbb{E}y) + \bar{\delta}\mathbb{E}u(x, y) \tag{55}$$

where:

$$u(x, y) = \gamma^{-1}\left[ax^{\rho} + \bar{a}y^{\rho} - s_{\rho}\theta|x^{\rho} - y^{\rho}|\right]^{\frac{\gamma}{\rho}} \tag{56}$$

2. Ex-Post Fair Risks

When risks are ex-post fair (that is, $x = y$), we obtain the following expression:

$$u(x, y) = u(x) = \gamma^{-1}(x^{\rho})^{\frac{\gamma}{\rho}} = \gamma^{-1}x^{\gamma} \tag{57}$$

As such, expected utility and the utility of expectations collapse, respectively, into the following:

$$\mathbb{E}u(x, y) = \mathbb{E}u(x) = 0.5\gamma^{-1}x_A^{\gamma} + 0.5\gamma^{-1}x_B^{\gamma} = 0.5\gamma^{-1}\left[x_A^{\gamma} + x_B^{\gamma}\right] \tag{58}$$

$$u(\mathbb{E}x, \mathbb{E}y) = u(\mathbb{E}x) = \gamma^{-1}\left[0.5x_A + 0.5x_B\right]^{\gamma} \tag{59}$$

Utility is thus equal to the following expression:

$$U(L) = \delta\gamma^{-1}\left[0.5x_A + 0.5x_B\right]^{\gamma} + \bar{\delta}0.5\gamma^{-1}\left[x_A^{\gamma} + x_B^{\gamma}\right] \tag{60}$$

Maximizing over x_A and x_B, the FOC is given by:

$$\frac{p_A}{p_B} = \frac{\delta(0.5x_A + 0.5x_B)^{\gamma-1} + \bar{\delta}x_A^{\gamma-1}}{\delta(0.5x_A + 0.5x_B)^{\gamma-1} + \bar{\delta}x_B^{\gamma-1}} \tag{61}$$

which can be further simplified as

$$p = \frac{\delta(0.5x + 0.5)^{\gamma-1} + \bar{\delta}x^{\gamma-1}}{\delta(0.5x + 0.5)^{\gamma-1} + \bar{\delta}} \tag{62}$$

where x denotes x_A/x_B, and p represents p_A/p_B.

3. Ex-Post Unfair (but Ex-Ante Fair) Risks

When risks ex-ante fair (meaning $\mathbb{E}x = \mathbb{E}y$) but ex-post unfair (implying $x_A = y_B$ and $x_B = y_A$), utility takes the following form:

$$U(L) = \delta\gamma^{-1}\left(0.5x_A + 0.5x_B\right) + \bar{\delta}0.5\gamma^{-1}(ax_A^\rho + \bar{a}x_B^\rho - s_\Delta\theta(x_A^\rho - x_B^\rho))^{\frac{\gamma}{\rho}} \\ + \bar{\delta}0.5\gamma^{-1}(ax_B^\rho + \bar{a}x_A^\rho - s_\Delta\theta(x_A^\rho - x_B^\rho))^{\frac{\gamma}{\rho}} \tag{63}$$

Note that we can set the functions

$$f(x_A, x_B) = [0.5x_A + 0.5x_B]^{\gamma-1} \tag{64}$$

$$g(x_A, x_B) = \left[(a - s_\Delta\theta)x_A^\rho + (\bar{a} + s_\Delta\theta)x_B^\rho\right]^{\frac{\gamma}{\rho}-1} \tag{65}$$

$$h(x_A, x_B) = \left[(a + s_\Delta\theta)x_B^\rho + (\bar{a} - s_\Delta\theta)x_A^\rho\right]^{\frac{\gamma}{\rho}-1} \tag{66}$$

Such that the FOC can be expressed in the following manner:

$$\frac{P_A}{P_B} = \frac{\delta f(x_A, x_B) + \bar{\delta}x_A^{\rho-1}\left[g(x_A, x_B)(a - s_\Delta\theta) + h(x_A, x_B)(\bar{a} - s_\Delta\theta)\right]}{\delta f(x_A, x_B) + \bar{\delta}x_B^{\rho-1}\left[g(x_A, x_B)(\bar{a} + s_\Delta\theta) + h(x_A, x_B)(a + s_\Delta\theta)\right]} \tag{67}$$

This condition applies everywhere outside the 45-degree line. When nearing the kink, however, it becomes necessary to evaluate the |MRS| when $x_A = x_B = x$ and therefore $g(x_A, x_B) = h(x_A, x_B) = x^{\gamma-\rho}$.

$$|MRS| = \frac{\delta + \bar{\delta}[1 - 2s_\Delta\theta]}{\delta + \bar{\delta}[1 + 2s_\Delta\theta]} \tag{68}$$

When approaching the 45-degree line from the left ($x_A < x_B$), we have $s_\Delta = -1$ and thus

$$|MRS| = \frac{\delta + \bar{\delta}[1 + 2\theta]}{\delta + \bar{\delta}[1 - 2\theta]} > 1 \tag{69}$$

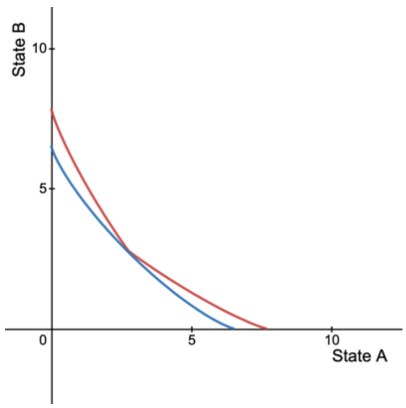
On the other hand, when approaching the 45-degree line from the right ($x_A > x_B$), we have $s_\Delta = 1$ and thus:

$$|MRS| = \frac{\delta + \bar{\delta}[1 - 2\theta]}{\delta + \bar{\delta}[1 + 2\theta]} < 1 \tag{70}$$

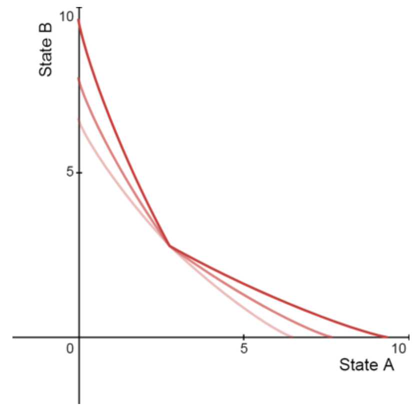
4. Sharing in Chances

The full proof is excluded for the sake of space, but is available from the authors. Simple examples are immediate.

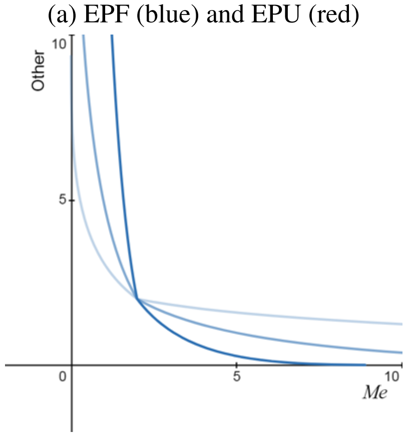
J The Model: Indifference Curves



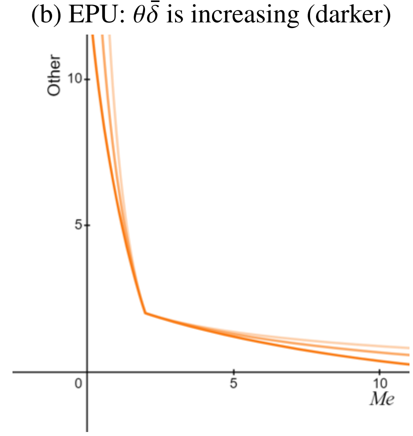
(a) EPF (blue) and EPU (red)



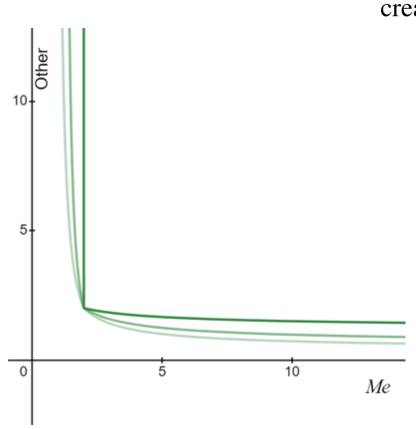
(b) EPU: $\theta\bar{\delta}$ is increasing (darker)



(c) Deterministic Giving: a is increasing (darker)



(d) Deterministic Giving: ρ is decreasing (darker)



(e) Deterministic Giving: θ is increasing (darker)

Figure 16: Indifference curves under EPF, EPU, and DetGiv.

K Experiment Instructions

The full set of instructions appears below. We provided instructions before each type of task. After the initial instructions, we include control questions about sample tasks.

Task types are presented in various orders, randomly assigned within each session. Instructions below appear according to the following order: Ex-post fair tasks, Ex-post unfair tasks, Deterministic Giving and Sharing in Chances. Instructions were presented to participants in Spanish.

General Instructions

Hello and Welcome

General Instructions

Please pay careful attention to the instructions. You will be asked to make decisions in 4 different tasks. Precise rules will be provided before you start each task.

Your final earnings today will depend on your decisions, on the decisions of others, and on chance. At the end of the experiment, you will be paid in cash. All payments will be made in private using an envelope. You will also receive an additional $5.00 as a participation fee.

During the study we will use **Experimental Tokens** instead of dollars. At the end of the study, your earnings in tokens will be translated into dollars. You will receive 1 dollar for every 3 tokens you have earned in the session. It is important that you do not talk or in any way try to communicate with other people during the session. If you have a question, raise your hand and we will attend to your station to answer your question privately. The experiment should be finished in approximately one hour.

Pairs

At the beginning of the study, you will be randomly and anonymously matched with another participant to form a pair. Participants will not learn the identity of the counterpart at any moment of the session.

Within each pair there will be one person with the role of Decider and another person with the role of Non-decider. These roles will be assigned randomly within each pair, and will remain for the entire session. Only choices made by the Decider will determine final earnings. Non-decider choices will not determine anyone's earnings. Roles will not be disclosed until the last decision has been made.

Tasks, Decision Rounds, Earnings

This study consists of 4 different tasks. Each task has multiple decisions rounds. The entire experiment has 48 decision rounds. Before the start of each task you will be given instructions specific to that task.

Your role (Decider or Non-Decider) will be disclosed after you and your counterpart are finished making decisions. Then, the computer will randomly select one decision round as the Decision-That-Counts. As all decisions are equally likely to be chosen and you could be the Decider, it is in your interest to treat each decision as if it is the Decision-That-Counts. The Decision-That-Counts made by the Decider will generate earnings for your pair (Decider and Non-Decider) based on the Decider's choice. Finally, earnings will be disclosed to both participants.

Instructions for Deterministic Giving Tasks

In this task, you will be given a set of possible Token combinations for you and your Partner. You are asked to choose one combination and submit your decision.

The figure below shows an example of this type of task. You will see a graph on a white background representing token allocations. In every round of this task, your tokens will be indicated on the horizontal axis and your Partner's tokens on the vertical axis. This information is also displayed in the table on the right of the screen. Throughout this session, your tokens will be represented in blue and Partner's tokens in orange.

For example, the point (3.9, 30.8) depicted in the graph by a red square indicates that you will receive 3.9 tokens and your counterpart will receive 30.8 tokens. This is also indicated in the label "You: 3.9, Partner: 30.8" next to the red square and in the table next to the figure.

At the beginning of each task, the red square will appear at a random location. To make a choice, drag the red square with the mouse to any position on the gray line. Only combinations along this line are feasible, valid choices. You can try as many combinations as you want before you make your decision. Once you have the red square at your preferred position, press the Submit Decision button and continue to the next round. Please make your decisions carefully.

Please press on the Next button to proceed.

[Screenshots of the graphical interface are shown in Appendix B]

Instructions for Ex-post Fair Risks

In each decision round, there will be two probable states: State A and State B. Think of these states as the weather: it could be either sunny or cloudy. However, when you make your decision, it is uncertain which state will occur.

Your decision will be represented by a point on a graph like the one in the figure below. In this graph, the horizontal axis indicates Tokens paid if State A occurs, and the vertical axis indicates Tokens paid if State B occurs. The chance that each state will occur is displayed in parentheses on the corresponding axis label. In the example screen-shot, for instance, the probability of State A is 50%, and the probability of State B is 50%, as well. Be aware that these probabilities might be different in different decision rounds.

The position of a point on this graph represents a lottery. For example, the point on location (40, 20) indicates that if State A is realized, the lottery pays 40 tokens; and if State B is realized, the lottery pays 20 tokens. Your lottery will be depicted by a blue square, and your partner's lottery by an orange circle.

Your task:

In each decision of this task, both you and your Partner share the same fortune. That is, in each state, the amount of tokens both receive is the same. See figure below for an example. Since in this task both face the same fate, your blue square and Partner's orange circle are always located on the same position. In the example, they both are on location (50, 16.7). This means, if State A happens, you and your Partner will each get 50 tokens and, if State B occurs, each of you get 16.7 tokens. At the beginning of each decision round, the square and the circle will appear at a random position. To make a choice drag the square or the circle to your chosen location. The other shape will follow.

Only combinations on the lines are feasible choices. The same information displayed in the graph is shown in the table next to it. Each row represents one possible state and the tokens paid if such state occurs. You can try as many combinations as you want before you decide. Once you make your choice, press the Submit Decision button and continue to the next round.

[Screenshots of the graphical interface are shown in Appendix B]

Instructions for Ex-post Unfair Risks Tasks

In each decision round, there will be two probable states: State A and State B. Think of these states as the weather: it could be either sunny or cloudy. However, when you make your decision, it is uncertain which state will occur.

Your decision will be represented by a point on a graph like the one in the figure below. In this graph, the horizontal axis indicates Tokens paid if State A occurs, and the vertical axis indicates Tokens paid if State B occurs. The chance that each state will occur is displayed in parentheses on the corresponding axis label. In the example screen-shot, for instance, the probability of State A is 50%, and the probability of State B is 50%, as well. Be aware that these probabilities might be different in different decision rounds.

The position of a point on this graph represents a lottery. For example, the point on location (40, 20) indicates that if State A is realized, the lottery pays 40 tokens; and if State B is realized, the lottery pays 20 tokens. Your lottery will be depicted by a blue square, and your partner's lottery by an orange circle.

Your task:

In each decision of this task, both you and your Partner have opposite fortunes. That is, what you would get in State A equals what your partner would receive in State B. Similarly, what you would get in State B is what your partner would get in State A. Since you and your partner face reverse fates, your blue square and your partner's orange circle are always in mirror positions on the graph. In the example screen shot, for example, your square is at (State A = 50, B = 16.7) and Partner's circle at (State A = 16.7, B = 50). This means, if State A happens, You receive 50 and Partner gets 16.7; and, if State B happens, you get 16.7 and your Partner gets 50. At the beginning of each decision screen, the square and the circle will appear at a random position. To make a choice drag the square and the circle to your chosen locations. The other shape will self adjust to the mirror position.

Only combinations on the lines are feasible choices. The same information displayed in the graph is shown in the table next to it. Each row represents one possible state and the tokens paid if such state occurs. You can try as many combinations as you want before you decide. Once you make your choice, press the Submit Decision button and continue to the next round.

[Screenshots of the graphical interface are shown in Appendix B]

Instructions for Sharing in Chances Tasks

In this task, you are given two fixed, mutually exclusive outcomes. You are then asked to decide the probabilities of these two outcomes. See graph below.

In the left-side graph, your tokens are represented on the horizontal axis and your partner's tokens on the vertical axis. Each of the two possible outcomes is represented by a dot or a bubble in this graph. You must decide what the probability of each outcome is in percentage (%) terms. The figure below shows a sample screen¬shot of this task where the two fixed outcomes are A = (You: 70, Partner: 10) and B = (You: 10, Partner: 80).

On the right side of the screen you have a slider tool where you can choose the chance of Outcome A, from 0% to 100%. Drag the slider onto your chosen percentage. For example: If you choose 40, outcome A will occur with a 40% chance and outcome B with a 60% chance. If you choose 100, outcome A will occur for sure. If you choose 0, outcome B will occur for sure.

In the graph, the size of each outcome bubble will increase with the chances it is given. You can try as many combinations as you want before you submit your decision. The same information of the graph is displayed in the table below the slider bar. Throughout this session, your tokens will be represented in blue and your partner's tokens in orange.

Once you make your choice, press the submit decision button and continue to the next round. If you do not make a choice (do not move the slider to your preferred position, the computer will return an error. Please think through your decisions carefully. If, at the end of the session, this Decision Round is selected to be paid, the computer will randomly choose outcome A or B according with the probabilities you decided and will pay you and your partner accordingly.

[Screenshots of the graphical interface are shown in Appendix B]