TCER Working Paper Series

Stochastic Choice and Social Preferences: Inequity Aversion versus Shame Aversion

Yosuke Hashidate
Keisuke Yoshihara

January 2021

Working Paper E-155
https://www.tcer.or.jp/wp/pdf/e155.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH
1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

In this paper, we propose a theory to identify the motivations behind altruistic or prosocial behavior. We focus on inequity aversion and shame aversion as social image concerns. To study these, we characterize two additively perturbed utility models, that is, the sum of expected utility and a non-linear cost function. First, we examine how to distinguish between stochastic inequity-averse behavior and stochastic shame-averse behavior. Next, we show that additively perturbed inequity-averse utility captures the general class of inequity-averse preferences, including ex-ante and ex-post fairness. Finally, we consider the relationship between our models and random utility, one of the most common stochastic choice models.

Yosuke Hashidate
Waseda University
Graduate School of Business and Finance
1-6-1 Nishi-Waseda, Shinjuku-ku Tokyo, 169-8050, JAPAN
yosukehashidate@gmail.com

Keisuke Yoshihara
Gunma University
Center for Mathematics and Data Science
4-2 Aramaki-machi, Maebashi-city, Gunma, 371-8510, JAPAN
yoshi@gunma-u.ac.jp

Stochastic Choice and Social Preferences: Inequity Aversion versus Shame Aversion*

YOSUKE HASHIDATE† KEISUKE YOSHIHARA‡

First Draft: April 27, 2020; Current Draft: January 10, 2021

Abstract

In this paper, we propose a theory to identify the motivations behind altruistic or prosocial behavior. We focus on inequity aversion and shame aversion as social image concerns. To study these, we characterize two additively perturbed utility models, that is, the sum of expected utility and a nonlinear cost function. First, we examine how to distinguish between stochastic inequity-averse behavior and stochastic shame-averse behavior. Next, we show that additively perturbed inequity-averse utility captures the general class of inequity-averse preferences, including ex-ante and ex-post fairness. Finally, we consider the relationship between our models and random utility, one of the most common stochastic choice models.

Keywords: Inequity Aversion; Ex-Ante Fairness; Ex-Post Fairness; Deliberate Randomization; Shame Aversion; Additive Perturbed Utility.

JEL Classification Numbers: D63; D64; D81; D91.

*We thank Ryota Iijima, Tetsuya Kawamura, Yusuke Osaki, and the participants at CTW Summer Camp (Hirosaki University), ABEF (Nagoya University of Commerce and Business), Game Theory Workshop (Komazawa University, Online), and ASSA 2021 (Virtual), for their helpful comments and suggestions. Hashidate acknowledges financial support from the Tokyo Center for Economic Research (TCER) and JSPS (KAKENHI Grant Number JP19J01049). All remaining errors are ours.

†*Corresponding Author*: Email: yosukehashidate@gmail.com; Graduate School of Business and Finance, Waseda University and Japan Society for the Promotion of Science; Address: 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, JAPAN.

‡Email: yoshi@gunma-u.ac.jp; Center for Mathematics and Data Science, Gunma University; Address: 4-2 Aramaki-machi, Maebashi-city, Gunma, 371-8510, JAPAN.

1 Introduction

In this study, we develop two axiomatic stochastic choice models for social contexts.[1] Using *stochastic choice functions* as primitives, we study social preferences, particularly *inequity aversion* and *shame aversion*. Under plausible conditions, axiomatization makes it possible to distinguish inequity-averse behavior from shame-averse behavior, and vice versa.

Background and Motivation. It is challenging to explain why people engage in altruistic behavior or behave prosocially in social contexts.[2] One explanation for such altruistic behavior is that the decision maker has a preference for fairness (Fehr and Schmidt, 1999). If the decision maker is inequity-averse, they compare their payoff with the other (passive) agents' payoffs, and the difference between payoffs causes *envy* or *guilt*. To avoid these feelings the decision maker chooses altruistic behavior.

Recent experimental evidence such as Dana et al. (2006) has suggested an alternative explanation; that is, the decision maker acts altruistically out of concern for their social image. In other words, they might feel social pressure to behave generously. A decision maker with social image concerns cares about how other agents perceive their choice behavior. Such a belief affects the trade-off between selfish motivation and personal norms, which impacts behavior (Dillenberger and Sadowski, 2012).[3]

We cannot easily identify the motivation behind observed behavior. Even if we observe that an altruistic action has been taken, we cannot be sure whether this was due to fairness concerns or image concerns. Consider the dictator game, where the dictator chooses from a set of options $\{(1,0),(0,1)\}$.[4] Suppose that the decision maker chooses $(1,0)$ or $(0,1)$ in a deterministic manner. The former choice exhibits *selfishness*, whereas the latter exhibits *altruism*.[5]

Altruistic choice behavior can be captured by *deliberate randomization*, experimentally studied in Agranov and Ortoleva (2017). Decision makers deliberately choose items randomly because it is *optimal* for them to do so due to trembling

[1] This paper uses the term "social context" to refer to social decision-making, scenarios where decision makers choose distributions of payoffs or lotteries of payoffs. The dictator game is a typical example of such decision-making.

[2] In economics, it is postulated that decision makers are *selfish* in the sense that they maximize their individual payoffs. However, experimental evidence often indicates that subjects tend to engage in altruistic or prosocial behavior (Camerer, 2003).

[3] Personal norms might differ from social norms. Hashidate (2020a) studies such a social decision-making with *reference-dependent* preferences.

[4] The allocation $(1,0)$ states that the dictator obtains the payoff 1, and the recipient obtains the payoff 0.

[5] In the presence of *social pressures*, the choice behavior may exhibit *impure selfishness* or *impure altruism*. See Saito (2015a), Hashidate (2020a), etc.

hands with implementation costs (Fudenberg et al., 2015), Allais-style lottery preferences (Cerreia-Vioglio et al., 2019; Machina, 1985), hedging against ambiguity (Saito, 2015b), regret minimization (Dwenger et al., 2018), and so on.

In social contexts, one possible explanation of deliberately stochastic choice behavior is that the decision maker is inequity-averse, that is, the decision maker dislikes "unfair" items.[6] If a menu does not include "fair" items, then the decision maker randomizes the items to explore *equity for opportunities* (Fudenberg and Levine, 2012).[7] In the example of $\{(1,0),(0,1)\}$, the dictator may delegate their decision-making to a "fair" coin flip to obtain the same expected payoff.[8]

Alternatively, there is the possibility of deliberately stochastic choice behavior. The decision maker may be perceived as selfish by other agents if they deterministically engage in selfish behavior. This makes the image-conscious decision maker averse to openly displaying such deterministic selfish behavior. Therefore, they may choose to engage in deliberately stochastic choice behavior.

Identification of the two incentives of behavior is difficult as the two theories adopt different approaches. The axiomatic study of *inequity aversion* takes preferences over items as primitives. On the other hand, the axiomatic study of *social image* takes preferences over choice sets, that is, *menus* of items as primitives. Since the primitives differ, it is not easy to compare the axioms of inequity-aversion with the axioms of social image concerns.

Objective. The goal of this study is to provide a unified framework to compare inequity-averse preferences with image-conscious preferences. We use stochastic choice functions as primitives, and axiomatize the two stochastic choice models stemming from inequity aversion and shame aversion. The objective is to distinguish inequity-averse behavior from shame-averse behavior, and vice versa. By doing this we seek to gain a deeper understanding of the motivations behind altruistic or prosocial behavior.

The contributions of this paper to the literature are threefold. First, we axiomatically examine how to identify the motivations behind altruistic or prosocial behavior. Second, we provide an axiomatic foundation for the general class of inequity-averse preferences consistent with experimental evidence (Brock et al., 2013; Miao and Zhong, 2018; Sandroni et al., 2013). In particular, we allow for inequity-averse behavior in Miao and Zhong (2018), which is not consistent with Saito (2013), a seminal axiomatic model in *inequity-averse preferences*. Third, we consider the relationship between our two models and *Random Utility* (henceforth, RU), one of the most common models used to understand stochastic decision-making.

[6]We call allocations of payoffs *items* in this paper. We say that an item is *fair* if every agent obtains the same payoff.

[7]The *equity for opportunities* is interpreted as *ex-ante fairness*.

[8]We say that a coin is *fair* if heads and tails occur with the same probability 0.5.

Outline. The remainder of this paper is organized as follows. In Section 2, we briefly explain the results. In Section 3, we characterize the additive perturbed inequity-averse utility (APU(IA)). In Section 4, we characterize the additive perturbed shame-averse utility (APU(SA)). In Section 5, we compare APU(IA) with APU(SA). In 6, we discuss the experimental evidence and the relationship between our models and RU. In Section 7, we provide a literature review. In section 8 we conclude the paper. All proofs are provided in the Appendix.

2 Summary of Results

In this section, we provide a brief summary of the results of this study. First, we explain APU(IA). Next, we explain APU(SA). Finally, we compare APU(IA) with APU(SA).

Stochastic Inequity-Averse Choice Behavior. To axiomatize deliberately stochastic inequity-averse behavior, we consider the *acyclic* condition on fairness concerns (*Inequity-Averse Mixing*). We require that there is no cycle on the monotonicity for "fair" items.

Consider a finite sequence of stochastic choices over menus $\{A_i\}_{i=1}^{k}$ $(i = 1, \cdots, k)$. Suppose that the stochastic choice from menu A_{i+1} could also be obtained from menu A_i. The standard condition in the revealed preference theory requires that the stochastic choice from A_i must be at least as good as anything available from A_{i+1}. The "stochasticity" can thus be interpreted as a lottery. Certainly, we cannot observe preference relations over lotteries of items. However, if the decision maker is inequity-averse, there cannot be anything in A_k that dominates the stochastic choice from A_1 in the sense of monotonicity for "fair" items.

The key axiom, *Inequity-Averse Mixing*, in addition to the standard conditions in inequity-averse preferences, characterizes stochastic inequity-averse choice behavior (Theorem 1). The model is in the class of additive perturbed utility (APU) (Fudenberg et al., 2015). APU has the two building blocks of the model: (i) a utility function and (ii) a convex perturbed cost function that can reward the decision maker for deliberate randomization. Our model, *additive perturbed inequity-averse utility* (APU(IA)), corresponds to the case where the utility function is inequity-averse (Fehr and Schmidt, 1999), and the cost function is *menu-dependent*.

APU(IA) captures a general class of inequity-averse preferences, including not only *ex-post fairness* but also *ex-ante fairness*. We characterize the cost functions that stem from not only *ex-post fairness* (Proposition 2), but also *ex-ante fairness* (Proposition 3). In the latter case, the cost function using deliberate randomization can be related to the difference between the expected payoff on ex-post fairness and the expected payoff for ex-ante fairness in Saito (2013).[9] Although Saito (2013) studied

[9] He axiomatizes *expected inequity-averse* (EIA) utility, which is a convex combination of ex-post

a linear relationship between ex-post and ex-ante fairness, we allow for non-linear relationships, which can lead to a broader understanding of behavioral patterns.

APU(IA) deviates from *Regularity*, one of the most common conditions in stochastic choices. Regularity states that the choice probabilities of items decrease as menus increases in the sense of set inclusions. Since RU satisfies *Regularity*, APU(IA) has a different motivation behind stochastic choice behavior. This feature also appears in Cerreia-Vioglio et al. (2019).

Stochastic Shame-Averse Choice Behavior. If they are image-conscious, the selfish decision maker dislikes revealing their image or type. The decision maker generally has a personal normative criterion. With image-conscious preferences, personal or social norms affect behavior under social pressures. The feeling of *shame* from selfish acts occurs when the individual deviates from the personal or social norm.

To axiomatize deliberately stochastic shame-averse behavior, we consider the *acyclic* condition of shame-averse preferences (*Shame-Averse Acyclicity*). If we add a (weakly) normatively better item into arbitrary menus, the choice probabilities of less normative items decrease. This stochastic choice behavior captures shame aversion.

Consider a finite sequence of menus that is susceptible to shame ($\{A_i\}_{i=1}^k$). Suppose that a menu A_i is more susceptible to shame than A_{i+1} ($i = 1, \cdots, k$). For such a sequence, choice probabilities do not have a cyclical nature.

The axiom *Shame-Averse Acyclicity*, in addition to the standard conditions, characterizes a stochastic shame-averse choice (Theorem 2). Our model corresponds to the case of the Menu-Invariant APU (Fudenberg et al., 2014), that is, a utility function and a *item-dependent* cost function. In particular, the model, *additive perturbed shame-averse utility* (APU(SA)), corresponds to the case in which the utility function is selfish, and the cost function is *item-dependent*, which is related to the psychological costs of shame aversion. APU(SA) can be interpreted as a stochastic version of Dillenberger and Sadowski (2012).

We use stochastic choice functions as primitives, while Dillenberger and Sadowski (2012) take preferences over menus as primitives. Fudenberg and Strzalecki (2015) argue that the two approaches can have different implications for menu choices. Fudenberg and Strzalecki (2015) identify the attitude toward *choice aversion*, which is affected by the extent to which the decision maker would like to add new items to the menu, as in "preferences for flexibility" (Kreps, 1979) and "preferences for commitment" (Gul and Pesendorfer, 2001). Choice aversion implies that the decision maker prefers removing items from a menu if their ex-ante value level is below a certain threshold. The psychological costs of shame aversion can be a specification of choice aversion in Fudenberg and Strzalecki (2015), in that they

---

fairness and ex-ante fairness.

could lead the decision maker to remove normatively better items.

APU(SA) is consistent with *Regularity*. In particular, if the shame function is *linear*, we can describe APU(SA) as RU. In other words, if the item-dependent cost functions are linear, stochastic choice behavior does not stem from deliberate randomization.

Inequity Aversion vs. Shame Aversion. We theoretically examine how to distinguish inequity aversion from shame aversion given observed behavior, and vice versa. We can do this by testing each axiom individually. However, we have taken a different approach. We provide some interesting cases that show us how to distinguish between inequity-averse behavior and shame-averse behavior.

We study the differences between inequity-aversion and shame aversion under the same (fixed) level of social pressures. In some previous studies, subjects engaged in selfish or altruistic behaviors with or without social pressures (see, for example, DellaVigna et al. (2012)). In other experiments, subjects choose the opportunity to take actions with or without social pressures (see, for example, Dana et al. (2006)[10]). Here, we identify the motivation of altruistic/prosocial behavior under the same level of social pressure.

First, by using the property of *Regularity*, we study the changes in choice probabilities by adding "unfair" items into menus. For those with inequity-averse preferences, because of "unfair" items, "fair" items become more attractive, so the choice probabilities of them can increase, which implies violations of *Regularity*.

For those with shame-averse preferences, the trade-off between selfishness and personal norms matters. Added "unfair" items affect the resulting behavior. However, since APU(SA) is consistent with *Regularity*, as menus gain more choice sets, the choice probabilities of all items decrease. The behavioral patterns make it clear that inequity-aversion and shame aversion have different motivations behind altruistic or prosocial behavior.

We study the decoy and menu-size effects on stochastic choice. The behavioral patterns are also related to the *Regularity*. Since APU(SA) is consistent with the property, the choice probabilities of items decrease as the menu-size increases. On the other hand, in APU(IA), some items can become more attractive as the menu-size increases. As a result, the choice probabilities of such items increase, a violation of *Regularity*.

Moreover, we study the (dis)advantageous cases: whether the resulting behavior is stochastic or deterministic. In the advantageous case, where the dic-

[10]Dana et al. (2006) study two-stage dictator games by introducing the *exit option* $(\$9, \$0)$. In the first stage, subjects choose to play the dictator game to allocate \$10, or to take the exit option $(\$9, \$0)$. The exit option was not observed by the recipients, but the experimenter did observe it. The subject has no incentive to increase the observer's welfare, nor do they affect the experimenter's welfare. If they choose to play the dictator game, then in the second stage, they play it in the usual manner. About one-third of the subjects chose to leave the dictator game.

tator (decision maker) obtains higher payoffs than the recipient (passive agent), APU(IA) exhibits almost deterministic behavior because of the linearity of indifference curves. On the other hand, APU(SA) exhibits stochastic behavior because of the trade-off between private ranking and personal norm. In the same way, in the disadvantageous case, where the dictator (decision maker) obtains smaller payoffs than the recipient (passive agent) does, APU(IA) exhibits almost deterministic behavior, whereas APU(SA) exhibits stochastic behavior.

3 Additive Perturbed Inequity-Averse Utility

Set-Up. Let $I = \{1, 2\}$ be a set of individuals where 1 is the decision maker, and 2 is the other (passive) agent.[11] We assume that the set of payoffs is $\mathbb{R}$. A vector $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ is called an *allocation* of payoffs among individuals, yielding payoff $x_i \in \mathbb{R}$ for each $i \in I$. Let $X \subseteq \mathbb{R}^2$ be the *compact* set of allocations. We call the allocations *items*. A choice set, that is, *menu*, is a non-empty subset of X. Let $\mathcal{A}$ be the collection of all non-empty finite subsets of X.[12] The elements in $\mathcal{A}$ are denoted by $A, B, C \in \mathcal{A}$.

3.1 Preliminary

Utility Representation. Fehr and Schmidt (1999) introduce the model of *preferences for fairness*, which is one of the most well-known *inequity-averse preferences* in the behavioral economics literature. Let $\succsim_i$ on X be a binary relation of decision maker 1 over *items* (i represents items).[13]

Definition 1. There exists a pair (α, β), where $\alpha \geq 0$, and $\beta \geq 0$ such that $\succsim_i$ is represented by the function $u : X \to \mathbb{R}$ defined by

$$u_{IA}(\boldsymbol{x}) = x_1 - \alpha \max\{x_2 - x_1, 0\} - \beta \max\{x_1 - x_2, 0\}.$$

We explain the interpretations of the model. First, the term $\alpha \max\{x_2 - x_1, 0\}$ captures the disutility of *envy* if $x_1 \leq x_2$, that is, the decision maker 1's payoff is lower than agent 2's payoff. Next, the term $\beta \max\{x_1 - x_2, 0\}$ captures the disutility of *guilt* if $x_1 \geq x_2$, which occurs when the decision maker 1's payoff is higher than agent 2's payoff.

[11] We can extend the n-th agents' case.

[12] We follow from the setting in Fudenberg et al. (2014).

[13] See Rohde (2010) for the axiomatization of the Fehr and Schmidt (1999)'s model. For recent developments, see Hashidate (2020b).

3.2 Axioms

We axiomatically study a stochastic choice rule ρ, which maps a *menu* A to a probability distribution over the *items* in A, denoted by $\rho(A)$. Formally, we denote a stochastic choice rule by $\rho : \mathcal{A} \to \Delta(X)$, where $\Delta(X)$ is the set of probability distributions over X with finite support. For any $A \in \mathcal{A}$, let $\text{supp}(\rho(A))$ be the support of A, which is a subset of A. Given a menu $A \in \mathcal{A}$ with $\boldsymbol{x} \in A$, let us denote the probability that an item $\boldsymbol{x}$ is chosen from the menu A by $\rho(\boldsymbol{x}, A)$. For example, consider a menu $A = \{\boldsymbol{x}, \boldsymbol{y}\}$. Then, $\rho(A) = (\rho(\boldsymbol{x}, A), \rho(\boldsymbol{y}, A))$. Analysts also observe that $\overline{\rho(A)} = \sum_{x \in A} \rho(\boldsymbol{x}, A)\boldsymbol{x}$.

Stochastic Inequity-Averse Choice. What are the properties of stochastic inequity-averse behavior? Analysts observe stochastic choice data for each menu, that is, they observe probability distributions over several menus. Even if we observe that the decision maker engages in stochastic choice behavior, we do not know why their behavior is stochastic.[14]

We present the axioms of *inequity-averse preferences*. The first axiom is a *continuity* condition. The next two axioms are closely related to a class of *inequity-averse preferences*. The last axiom is an *acyclic* condition in revealed preference theory.

First, we state the basic axiom on *Continuity*. This axiom guarantees that the utility functions are continuous.

Axiom 1. (Continuity): For any menu $\{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}$ with sequences of items $\lim_{n\to\infty} \boldsymbol{x}_n^k \to \boldsymbol{x}^i$ for each $k = 1, \cdots, m$,

$$\lim_{n\to\infty} \rho(\boldsymbol{x}_n^k, \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\}) = \rho(\boldsymbol{x}^k, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}).$$

Motivation of Deliberate Randomization: the Case of Inequity-Averse Preference. Consider (deterministic) dictator games where there are two agents ($n = 2$). A significant number of subjects prefer allocating an endowment altruistically, rather than selfishly consuming the entire endowment by themselves (Camerer, 2003), that is, $(\frac{1}{2}, \frac{1}{2}) \succ_i (1, 0)$. However, the same subjects often prefer to consume the whole endowment themselves rather than give it away to a single recipient, that is, $(1, 0) \succ_i (0, 1)$. If the decision maker's preference $\succ_i$ satisfies *Transitivity*, then $(\frac{1}{2}, \frac{1}{2}) \succ_i (1, 0) \succ_i (0, 1)$. However, by choosing $(1, 0)$ with probability $\frac{1}{2}$ and $(0, 1)$ with probability $\frac{1}{2}$, we can obtain $(\frac{1}{2}, \frac{1}{2})$ under the outcome-mixture.

The source of such a violation is the fact that the rankings of the consequences are *opposite*; the other individual is better off than the decision maker in $(0, 1)$ and worse off in $(1, 0)$. In this case, the *inequality-averse* decision maker may have an *incentive* to use randomization, which leads to the violation of the *Independence*

[14] A coin flip as the cause of deliberately stochastic behavior is not observable. Preference relations over lotteries (of items) are not observable.

axiom.[15] Normative conditions such as *independence* hold as long as the two consequences are *not* opposite.

To distinguish the former case from the latter, we define the following: The rankings of the payoffs of individual $i \neq 1$ concerning decision maker 1 are not opposite between $\boldsymbol{x}$ and $\boldsymbol{y}$.

Definition 2. We say that two items $\boldsymbol{x}$ and $\boldsymbol{y} \in X$ are *quasi-comonotonic* if there exists no $i \neq 1$ such that $x_i > x_1$ and $y_i < y_1$.

We introduce the axiom of *Quasi-Comonotonic Additivity* under stochastic choice. This axiom requires the scale of choosing probability from menus, and does not change if the two items are *quasi-comotononic*. For each $A \in \mathcal{A}$ and $\boldsymbol{z} \in X$, let $A + \{\boldsymbol{z}\} := \{\boldsymbol{x} + \boldsymbol{z} \in X | \boldsymbol{x} \in A\}$, where $\boldsymbol{x} + \boldsymbol{z} = (x_i + z_i)_{i \in I}$.[16]

Axiom 2. (Quasi-Comonotonic Additivity): For any $A \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y} \in A$ and $\boldsymbol{z} \in X$, if $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are pair-wise quasi-comonotonic, then

$$\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{y}, A) \Rightarrow \rho(\boldsymbol{x} + \boldsymbol{z}, A + \{\boldsymbol{z}\}) > \rho(\boldsymbol{y} + \boldsymbol{z}, A + \{\boldsymbol{z}\}).$$

The next axiom is also related to *inequity-averse preferences*. The first condition is a property of *envy*, that is, $\rho((0,1), A) < \rho((0,0), A)$. The second condition is a property of *guilt*, that is, $\rho((0,-1), A) < \rho((0,0), A)$, where $i \neq 1$.[17]

Axiom 3. (Inequity Aversion): ρ satisfies (i) *envy* and (ii) *guilt*:

(i) (Envy): For any $A \in \mathcal{A}$ with $(0,1), \mathbf{0} \in A$,

$$\rho((0,1), A) < \rho(\mathbf{0}, A);$$

(ii) (Guilt): For any $A \in \mathcal{A}$ with $(-1,0), \mathbf{0} \in A$,

$$\rho((0,-1), A) < \rho(\mathbf{0}, A).$$

Finally, we consider an acyclic condition for *inequity-averse preferences*. This axiom is a modified version of the *Rational Mixing* introduced in Cerreia-Vioglio et al. (2019). Moreover, the axiom is a weaker version of *Item Acyclicity*, which was

[15] The *Independence* axiom is particularly used in the axiomatization of expected utility theory, which is required to be *linear* in probability.

[16] We can construct a simple experiment, and test this axiom. For example, according to Miao and Zhong (2018), consider a doubleton $\{\boldsymbol{x}, \boldsymbol{y}\}$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are *quasi-comotononic*. Moreover, take an item $\boldsymbol{z} \in X$ such that $\boldsymbol{x} + \boldsymbol{z}$ and $\boldsymbol{y} + \boldsymbol{z}$ are *quasi-comotononic*.

[17] By applying the property of the *monotonicity* with respect to fair items, we can easily test this axiom.

introduced in Fudenberg et al. (2015).[18] It is along the lines of the *rationalizability* conditions (Richter, 1966) and the *Strong Axiom of Revealed Preference* (SARP) (Houthakker, 1950). Since we focus on *inequity-averse preferences*, the acyclic condition includes a form of coherence with *fairness*. Note that various models of *fairness concerns* satisfy monotonicity with respect to equal allocations. In the presence of equal allocations, which are *optimal*, there is no incentive to engage in randomization.

Axiom 4. (Inequity-Averse Mixing): For each $k \in \mathbb{N} \setminus \{1\}$ and $A_1, \cdots, A_k \in \mathcal{A}$ if

$$\overline{\rho(A_2)} \in \mathrm{co}(A_1), \cdots, \overline{\rho(A_k)} \in \mathrm{co}(A_{k-1})$$

then

$$(x, x) \in A_k \Rightarrow (x, x) \not\succ \overline{\rho(A_1)},$$

3.3 Result

Utility Representation. For notational convenience, for each $\boldsymbol{x} \in A$, we write $\rho(\boldsymbol{x})$ instead of $\rho(\boldsymbol{x}, A)$.

Theorem 1. *The following statements are equivalent:*

(a) *ρ satisfies Continuity, Quasi-Comonotonic Additivity, Inequity Aversion, and Inequity-Averse Mixing.*

(b) *There exists a pair $\langle(\alpha, \beta), (c_A)_{A\in\mathcal{A}}\rangle$ where $\alpha \geq 0$, $\beta \geq 0$, and $(c_A)_{A\in\mathcal{A}}$ is a profile of cost functions for each menu $A \in \mathcal{A}$ such that*

$$\rho(A) = \arg\max_{\rho\in\Delta(A)} \sum_{\boldsymbol{x}\in A} \Big(u_{IA}(\boldsymbol{x})\rho(\boldsymbol{x}) - c_A(\rho(\boldsymbol{x})) \Big)$$

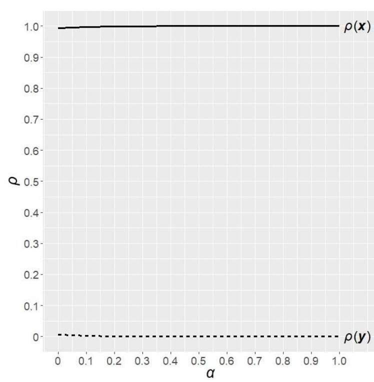
where $u_{IA}(\boldsymbol{x}) = x_1 - \alpha \max\{x_2 - x_1, 0\} - \beta \max\{x_1 - x_2, 0\}$.

We call the model in Theorem 1 the *additive perturbed inequity-averse utility* (APU(IA)) if the axioms in Theorem 1 are satisfied.

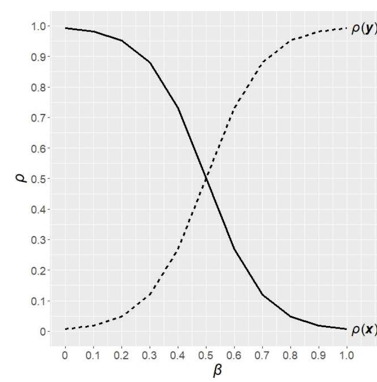
Inequity Aversion and Preference for Randomization. Inequity aversion leads to deliberately stochastic behavior. In Fehr and Schmidt (1999), the parameter α captures *envy*, and the parameter β captures *guilt*. We consider the relationship between each of parameter and deliberate randomization.

[18]See Appendix A.2 (Lemma 2). Let us introduce the relation $\geq^*$ on $[0, 1]$ defined by $p \geq^* q$ if and only if $p > q$ or $p = q \in (0, 1)$. ρ satisfies *Item Acyclicity* if

$$\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_2, A_1), \rho(\boldsymbol{x}_k, A_k) \geq^* \rho(\boldsymbol{y}_{k+1}, A_k) \text{ (for } 1 < k < n) \Rightarrow \rho(\boldsymbol{x}_n, A_n) \not\geq^* \rho(\boldsymbol{x}_1, A_n).$$



(a) $\boldsymbol{\alpha}$: Envy. Consider a doubleton $\{(5,5),(0,10)\}$, where $\boldsymbol{x} = (5,5)$ and $\boldsymbol{y} = (0,10)$. By fixing the menu and $\boldsymbol{\beta}$, we study how the level of *envy* affects choice probabilities.



(b) $\boldsymbol{\beta}$: Guilt. Consider a doubleton $\{(10,0),(5,5)\}$, where $\boldsymbol{x} = (10,0)$ and $\boldsymbol{y} = (5,5)$. By fixing the menu and $\boldsymbol{\alpha}$, we study how the level of *guilt* affects choice probabilities.

Figure 1: Inequity Aversion and Deliberate Randomization: We assume that in Fehr and Schmidt (1999)'s preferences, $\alpha = 0.8$ and $\beta = 0.3$. We consider the cost function to be menu-dependent in the sense that the parameter η depends on the maximizer of the Gini index of items in menus (Example 3.1).

Remark 1. *In the numerical examples, we use the logistic-type cost function with menu-dependence. For each $A \in \mathcal{A}$, let*

$$c_A(\rho(\boldsymbol{x})) := \eta(A)\rho(\boldsymbol{x})\log\rho(\boldsymbol{x}),$$

Interestingly, the parameter $\boldsymbol{\beta}$ of *guilt* affects deliberately stochastic choice behavior in Figure 1. If the decision maker exhibits *envy*, there is little effect on the choice probabilities on $\boldsymbol{x}$ and $\boldsymbol{y}$. On the other hand, if the decision maker exhibits *guilt*, as the level of *guilt* increases, the choice probability of $(10, 0)$ from $\{(10, 0),(5, 5)\}$ decreases. In addition, the choice probability of $(5, 5)$ from $\{(10, 0),(5, 5)\}$ increases.

Item-Invariant Utility. In APU(IA), the cost function depends on *menus*, that is, APU(IA) is an item-invariant APU (Fudenberg et al., 2014). Intuitively, different menus have different implementation costs. On the one hand, in a menu that includes a "fair" item, the inequity-averse decision maker may not have enough incentive to engage in randomization. On the other hand, in a menu that does not include "fair" items, the inequity-averse decision maker can have a strong incentive to engage in randomization. Therefore, one's attitude toward deliberate randomization relies on what menus they are presented with.

Violations of Regularity. APU(IA) generally deviates from the *Regularity*. The violation of *Regularity* also occurs in Brock et al. (2013) and Saito (2013).[19]

Axiom 5. (Regularity): For any $A, B \in \mathcal{A}$ with $\boldsymbol{x} \in A \subseteq B$,

$$\rho(\boldsymbol{x}, A) \geq \rho(\boldsymbol{x}, B).$$

Example 1. *(Violation of Regularity)*
Suppose that APU(IA) is represented by a logistic-type cost function. We assume that the parameter η is menu-dependent. Consider $I = \{1, 2\}$. Let $\boldsymbol{x} = (3, 4), \boldsymbol{y} = (4, 1)$, and $\boldsymbol{z} = (1, 5)$. For an inequity-averse preference, let $\alpha = 1$ and $\beta = 0.6$ (Fehr and Schmidt, 1999). Suppose that $\eta(\{(3, 4),(4, 1)\}) = 0.6$ and $\eta(\{(3, 4),(4, 1),(1, 5)\}) = 1$. Then, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \approx 0.417$, and $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\}) \approx 0.583$. Moreover, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \approx 0.449$, $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \approx 0.548$, and $\rho(\boldsymbol{z}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \approx 0.003$. We have $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \approx 0.417 < 0.449 \approx \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})$, which is not consistent with Regularity.

Uniqueness Result. We need the following richness conditions to achieve the uniqueness result.

[19] Hashidate and Yoshihara (2021) study a deliberately stochastic choice behavior stemming from Saito (2013).

Axiom 6. (Positivity): For each $A \in \mathcal{A}$ and $\boldsymbol{x} \in A$,

$$\rho(\boldsymbol{x}, A) > 0.$$

Axiom 7. (Richness): For any $\boldsymbol{x} \in X$ and any $p, q \in (0,1)$ such that $p + q \leq 1$, there exists $\boldsymbol{y}, \boldsymbol{z} \in X$ such that

$$\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = p \text{ and } \rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = q.$$

This axiom *Richness* states that the range of the utility function u equals $\mathbb{R}$. $u(X)$ is unbounded (either below or above), and there are at least three items with each utility level. These items should be unique. For example, consider inequity-averse preferences, where $|I| = 2$ (Fehr and Schmidt, 1999). Consider $\boldsymbol{x} \in X$. Without loss of generality, we can assume $\boldsymbol{x} = (x, \cdots, x)$, where $x \in \mathbb{R}$. Suppose that there exist $\boldsymbol{y}, \boldsymbol{z} \in X$ such that $y_1 > x_1 > z_1$ and $z_2 > x_2 > y_2$. Assume that in Axiom 6, $p = q = \frac{1}{2}$. Then, we can identify the pair of parameters (α_2, β_2).

Proposition 1. *Suppose that ρ satisfies Positivity and Richness in addition to the axioms in Theorem 1. If $\langle(\alpha, \beta), (c_A)_{A \in \mathcal{A}}\rangle$ and $\langle(\widehat{\alpha}, \widehat{\beta}), (\widehat{c}_A)_{A \in \mathcal{A}}\rangle$ represent the same ρ, then the following holds:*

(i) $(\alpha, \beta) = (\widehat{\alpha}, \widehat{\beta})$*;*

(ii) For each menu $A \in \mathcal{A}$, $c_A = a\widehat{c}_A + b_A p + c_A$ for some $a > 0, b_A, c_A \in \mathbb{R}$.

3.4 Ex-Post Fairness and Costs of Randomization

In this section we characterize the menu-dependent cost function that stems from *inequity aversion*. In Theorem 1, cost functions are menu-dependent, but the forms are not restrictive. Here, we study inequity-averse behavior in the case that randomization is costly with respect to "unfair" items.

Throughout this subsection, we assume that each menu-dependent cost function c_A is twice-differentiable.

unfair items are defined as follows: unfairness increases as the difference in the absolute value between 1's payoff and 2's payoff becomes larger.

Definition 3. We say that an item $\boldsymbol{x}$ is more *unfair* than $\boldsymbol{y}$ if $|x_1 - x_2| > |y_1 - y_2|$.

We present a new axiom to capture the relationship between inequity-aversion and deliberate randomization.

Axiom 8. (Ex-Post Fairness-Seeking): For any $\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{z}' \in X$ such that $\boldsymbol{z}$ is more unfair than $\boldsymbol{z}'$, if the following conditions hold:

(i) $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = \rho(\boldsymbol{y}, \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\})$; and

(ii) $\rho(x, \{x, x'\}) > \frac{1}{2}$,

then

$$\rho(x, \{x, y, z\}) \geq \rho(x, \{x, y, z'\})$$

The interpretation of the axiom is as follows. Condition (i) requires that both $\{x, y, z\}$, and $\{x', y, z'\}$ have the same menu strength under ρ. Condition (ii) requires that the item x is preferred to the item x', that is, the item x is fairer than the item x'. Then, the statement that the choice probability of x from $\{x, y, z\}$ is larger than that from $\{x, y, z'\}$ stems from the degree of "unfairness."

We obtain the following result.

Proposition 2. *Suppose that ρ is represented by an APU(IA) $\langle(\alpha, \beta), (c_A)_{A\in\mathcal{A}}\rangle$. Then, ρ is Ex-Post Fairness-Seeking if and only if $(c_A)_{A\in\mathcal{A}}$ satisfies the following property: for any $A \in \mathcal{A}$ and $z, z' \in X$ such that z is more unfair than z',*

$$c''_{A\cup\{z\}}(\cdot) \geq c''_{A\cup\{z'\}}(\cdot).$$

Menu-Dependence and Examples. We provide three examples of logistic-type cost functions (Remark 1). First, we develop a cost function in which the parameter η depends on the Gini index for each item x. Next, we develop a cost function in which the parameter η depends on Theil (1967)'s index of inequality. This index is characterized by *entropy*. Finally, we develop a cost function in which the parameter η depends on Atkinson (1970)'s index of inequality. In each index, as unfair items are added, the inequity index increases. The penalty of deliberate randomization captured by η also increases.

Example 3.1. (Gini index): Let $\mathcal{G}(x)$ for each $x \in X$ be the *Gini coefficient* of the items. The Gini coefficient is the comparison of cumulative proportions of the population against cumulative proportions of income they receive. It is normalized, ranging from 0 (*perfect equality*) and 1 (*perfect inequality*).

Here, consider the menu-dependence such as

$$\eta(A) = \delta \max_{x'\in A} \mathcal{G}(x') + (1-\delta) \min_{y'\in A} \mathcal{G}(y')$$

where $\delta \in [0, 1]$.

We can consider various types of menu-dependence, such as the average of the Gini coefficient denoted by $\overline{\mathcal{G}}$:

$$\eta(A) = \overline{\mathcal{G}}(A).$$

Example 3.2. (Theil's Entropy index): Consider the menu-dependence such as

$$\eta(A) = \delta \max_{\boldsymbol{x}' \in A} \mathcal{T}(\boldsymbol{x}') + (1-\delta) \min_{\boldsymbol{y}' \in A} \mathcal{T}(\boldsymbol{y}')$$

where $\mathcal{T}$ evaluates the *Theil's index* of items, defined by

$$\mathcal{T} = \frac{1}{n} \sum_{i \in I} \mu \ln\Big(\frac{1}{\mu}\Big) - \sum_{i \in I} x_i \ln\Big(\frac{1}{x_i}\Big),$$

where μ is the average payoff of items. This index captures the distortion from the maximal entropy at the payoff of items.

Example 3.3. (Atkinson index) Consider the menu-dependence such as

$$\eta(A) = \delta \max_{\boldsymbol{x}' \in A} \mathcal{A}i(\boldsymbol{x}') + (1-\delta) \min_{\boldsymbol{y}' \in A} \mathcal{A}i(\boldsymbol{y}')$$

where $\mathcal{A}i$ evaluates the *Atkinson's index* of items, defined by

$$\mathcal{A}i = 1 - \frac{1}{\mu} \left(\frac{1}{n} \sum_{i \in I} x_i^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}},$$

where μ is the average payoff of items, and $\varepsilon \neq 1$ captures the inequity index ($\varepsilon \geq 0$).

3.5 Ex-Ante Fairness and Costs of Randomization

Brock et al. (2013) experimentally show that it is essential to care not only about *ex-post fairness* but also about *ex-ante fairness*. Saito (2013) provides an axiomatic foundation for a convex combination of *ex-ante fairness* and *ex-post fairness* (*Expected Inequity-Averse* (EIA) utility).[20] APU(IA) has the utility function of ex-post fairness concerns. Here, we study the relationship between ex-ante fairness and the costs of deliberate randomization.

Define the stochastic EIA choice in the following.

Definition 4. A stochastic choice rule ρ is a stochastic EIA choice if there exists a tuple $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma)$ $\boldsymbol{\alpha} := (\alpha_i)_{i \in S}$ is a profile with $\alpha_i \geq 0$ for each $i \in S$, $\boldsymbol{\beta} := (\beta_i)_{i \in S}$ is a profile with $\beta_i \geq 0$ for each $i \in S$, $\gamma \in [0,1]$ is a parameter such that ρ is represented by

$$\rho_{\text{Saito}}(A) = \arg \max_{\rho \in \Delta(A)} \Big(\gamma u_{IA} \Big(\sum_{\boldsymbol{x} \in A} x\rho(\boldsymbol{x}) \Big) + (1-\gamma) \sum_{\boldsymbol{x} \in A} u_{IA}(\boldsymbol{x}) \rho(\boldsymbol{x}) \Big)$$

where $u_{IA}(\boldsymbol{x}) = x_1 - \sum_{i=2}^{n} \Big(\alpha_i \max\{x_i - x_1, 0\} + \beta_i \max\{x_1 - x_i, 0\} \Big)$.

[20]See also Miao and Zhong (2018), who mentions a relationship between their model and Saito (2013).

To study the randomization with *ex-ante fairness*, consider the following axiom that characterizes the costs of randomization stemming from *ex-ante fairness*.

Axiom 9. (Ex-Ante Fairness-Seeking): For any $\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{z}' \in X$ such that $\boldsymbol{x}$ and $\boldsymbol{z}$ are not quasi-comonotonic, $\boldsymbol{x}$, and $\boldsymbol{z}'$ are not quasi-comonotonic, and $d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{z}')$, if the following conditions hold:

(i) $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = \rho(\boldsymbol{y}, \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\})$; and

(ii) $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{x}'\}) > \frac{1}{2}$,

then

$$\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \geq \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}'\}).$$

Conditions (i) and (ii) are the same as the axiom of *Ex-Post Fairness-Seeking*. The key point is that $\boldsymbol{x}$ and $\boldsymbol{z}$, as well as $\boldsymbol{x}$ and $\boldsymbol{z}'$ are not quasi-comonotonic. Moreover, $d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{z}')$ holds. In inequity-averse preferences, deliberate randomization stemming from ex-ante fairness can lead to "fair" payoffs for each agent. In the procedure of deliberate randomization, randomization between more unfair items is more costly than randomization between less unfair items.

We get the following result:

Proposition 3. *Suppose that ρ is represented by an APU(IA) $\langle(\boldsymbol{\alpha}, \boldsymbol{\beta}), (c_A)_{A\in\mathcal{A}}\rangle$. Then, ρ exhibits Ex-Ante Fairness if and only if $(c_A)_{A\in\mathcal{A}}$ satisfies the following property: for any $A \in \mathcal{A}$ and $\boldsymbol{z}, \boldsymbol{z}' \in X$ such that (i) for any $\boldsymbol{x} \in \arg\max_{\boldsymbol{x}\in A} u_{\mathrm{IA}}(\boldsymbol{x})$, both the pair $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{x}, \boldsymbol{z}')$ are not quasi-comonotonic, and (ii) $\min_{\boldsymbol{x}\in A} d(\boldsymbol{x}, \boldsymbol{z}) \geq \min_{\boldsymbol{x}\in A} d(\boldsymbol{x}, \boldsymbol{z}')$,*

$$c''_{A\cup\{\boldsymbol{z}\}}(\cdot) \geq c''_{A\cup\{\boldsymbol{z}'\}}(\cdot).$$

APU(IA) is more general than Saito (2013) in the following sense. Although in Saito (2013), a relationship between ex-post fairness and ex-ante fairness is described by a convex combination, we allow for non-linear relationships.

An Example of Cost Functions. We provide an example of a menu-dependent cost function stemming from ex-ante fairness concerns.

Example 3.4. (Ex-Ante Fairness): Let d be the Euclidean distance. For each $A \in \mathcal{A}$, $c_A(\rho(\boldsymbol{x})) := \eta(A)\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})$, where $\boldsymbol{x} \in A$ and

$$\eta(A) = d(\boldsymbol{x}_+, \boldsymbol{x}_-)$$

with $\boldsymbol{x}_+ \in \arg\max_{\boldsymbol{x}\in A_+} u_{\mathrm{IA}}(\boldsymbol{x})$ and $\boldsymbol{x}_- \in \arg\max_{\boldsymbol{x}\in A_-} u_{\mathrm{IA}}(\boldsymbol{x})$. We denote $A_+ := \{\boldsymbol{x} \in A | x_1 - x_2 \geq 0\}$ and $A_- := \{\boldsymbol{x} \in A | x_1 - x_2 \leq 0\}$. The menu-dependent penalty η increases as the distance increases, that is, deliberate randomization is costly.

4 Additive Perturbed Shame-Averse Utility

4.1 Preliminary

Utility Representation. Dillenberger and Sadowski (2012) introduce the model of *shame of acting selfishly*, which is the seminal axiomatic model in *social image concerns*.[21] We describe their *shame-averse* model (Theorem 1; p. 106) in our setting.

We consider the case of $|I| = 2$, where 1 is a decision maker and 2 is a passive agent.[22] As in Dillenberger and Sadowski (2012), assume that $X = (k, +\infty) \times (k, +\infty)$, where $k \in \mathbb{R} \cup \{-\infty\}$. Throughout this paper, we assume that $k > 0$. Dillenberger and Sadowski (2012) apply the preferences-over-menus framework into social contexts; that is, they take a binary relation $\succsim_m$ on $\mathcal{A}$ as a primitive.

Definition 5. We say that decision maker 1 is *susceptible to shame* if there exist A and B such that $A \succ_m A \cup B$.

They elicit a personal norm ranking $\succ_n$ on X from the primitive of their model $\succsim_m$ in the following way:

Definition 6. Suppose that decision maker 1 is susceptible to shame. We say that the decision maker deems $\boldsymbol{y}$ to be *normatively better than* $\boldsymbol{x}$; that is, $\boldsymbol{y} \succ_n \boldsymbol{x}$ if there exists $A \in \mathcal{A}$ with $\boldsymbol{x} \in A$ such that $A \succ_m A \cup \{\boldsymbol{y}\}$.

Moreover, let us denote the definition of "more selfish than" between two functions. Let u and φ be real-valued functions of X. For all $\boldsymbol{x} \in X$, $\triangle_1$ and $\triangle_2$, consider $(x_1 - \triangle_1, x_2 - \triangle_2) \in X$.

Definition 7. We say that u is *more selfish than* φ if for any $\boldsymbol{x} \in X$, $\triangle_1$, and $\triangle_2$ such that $(x_1 - \triangle_1, x_2 - \triangle_2)$,

(i) $u(\boldsymbol{x}) = u(x_1 - \triangle_1, x_2 + \triangle_2)$ implies $\varphi(\boldsymbol{x}) \leq \varphi(x_1 - \triangle_1, x_2 + \triangle_2)$

(ii) $u(\boldsymbol{x}) = u(x_1 + \triangle_1, x_2 - \triangle_2)$ implies $\varphi(\boldsymbol{x}) \geq \varphi(x_1 + \triangle_1, x_2 - \triangle_2)$

with strict inequality for at least one pair $(\triangle_1, \triangle_2)$.

We describe the *shame-averse* utility model.

[21] There are some related axiomatic studies on social image concerns. Saito (2015a) axiomatizes a general (*image-conscious*) utilitarian model in which the decision maker exhibits *shame* for acting selfishly, *pride* for acting altruistically, and *temptation* to act selfishly. Hashidate (2020a) generalizes the Saito (2015a)'s model in the case that social image concern is *context-dependent*, and his model allows for various social emotions such as *spitefulness* and *regret*.

[22] We can easily extend the general case: $|S| \geq 2$.

Definition 8. There exists a tuple (u, φ, g), where $u : X \to \mathbb{R}$ is a continuous function, which is weakly increasing[23] and more selfish than φ, $\varphi : X \to \mathbb{R}$ is a continuous function[24], and $g : \varphi(X) \times X \to \mathbb{R}$ is a continuous function that is strictly increasing in the first argument, such that the function $V : \mathcal{A} \to \mathbb{R}$, defined by

$$V(A) = \max_{\boldsymbol{x} \in A} \left[u(\boldsymbol{x}) - g\left(\max_{\boldsymbol{y} \in A} \varphi(\boldsymbol{y}), \boldsymbol{x} \right) \right].$$

represents $\succsim_m$ and φ represents $\succ_n$.

We interpret the model as follows. The value function V states that the decision maker chooses the maximal element from the menu A after considering two terms. The first term $u(\boldsymbol{x})$ is the private ranking of the items. For example, let $u(\boldsymbol{x}) = x_1$ for each $\boldsymbol{x} \in X$. The second term captures the private cost of the *shame* involved in acting selfishly. The function g is interpreted as the *shame* from choosing $\boldsymbol{x}$ instead of an item that fulfils the decision maker's personal norms represented by φ. For example, we particularly study the shame function $g : \mathbb{R} \to \mathbb{R}$ by $g(\varphi(\boldsymbol{x}))$ for each $\boldsymbol{x} \in A$. We consider specific personal norm functions such as *utilitarian* ($\varphi(\boldsymbol{x}) = x_1 + x_2$) and *Nash products* ($\varphi(\boldsymbol{x}) = x_1 x_2$).

4.2 Axioms

In this section we axiomatically examine a stochastic *image-conscious* choice driven by *shame aversion*.

Overview. First, we present the axioms of *shame aversion*. In the same way as *inequity-averse preferences*, we impose on *Continuity* (see Axiom 1).

The next axiom is related to *self-interest*. This axiom states that the ranking on item is *monotone*.

Axiom 10. (Self-Interest): For any $A \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y} \in A$ such that $\boldsymbol{x} \geq \boldsymbol{y}$,

$$\rho(\boldsymbol{x}, A) \geq \rho(\boldsymbol{y}, A).$$

Define the following.

Definition 9. We say that an item $\boldsymbol{y}$ is *weakly normative better* than item $\boldsymbol{x}$; that is, $\boldsymbol{y} \succsim_n^{\rho} \boldsymbol{x}$ if

$$\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\})$$

for some $A \ni \boldsymbol{x}$ and $\boldsymbol{y} \in X \setminus A$ with $x_1 x_2 < y_1 y_2$.

[23]For any $\boldsymbol{x}, \boldsymbol{y} \in X$ with $\boldsymbol{x} \geq \boldsymbol{y}$, $u(\boldsymbol{x}) \geq u(\boldsymbol{y})$.

[24]We call the function φ a *personal norm* function.

This definition is weaker than Definition 6 (Dillenberger and Sadowski, 2012). The definition has two conditions. The first condition is a *Regularity* condition. The choice probability of $\boldsymbol{x}$ from A is higher than that from $A \cup \{\boldsymbol{y}\}$. By adding $\boldsymbol{y}$ to A, the choice probability of $\boldsymbol{x}$ decreases. The second condition is that for the added item $\boldsymbol{y}$, $x_1 x_2 < y_2 y_2$ holds. This condition means that $\boldsymbol{y}$ is more *altruistic* than $\boldsymbol{x}$, and affects choice probabilities, that is, $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\})$. Such an altruistic item causes a trade-off between selfishness and altruism.

Weakly normatively better items are related to *deliberate randomization*. The definition states that the choice probability of the item $\boldsymbol{x}$ from $A \cup \{\boldsymbol{y}\}$ is lower than that from A. The decision maker is averse to be perceived as selfish by others. Adding an *altruistic* item $\boldsymbol{y}$ into menus increases psychological costs when choosing the item $\boldsymbol{x}$, which reduces the choice probability of the item $\boldsymbol{x}$ from the menu $A \cup \{\boldsymbol{y}\}$.

Finally, we consider a weaker acyclic condition for stochastic choices. We say that a menu A is more *susceptible to shame* than B if for all $\boldsymbol{x} \in B$, there exists a *weakly normative better* item $\boldsymbol{y} \in A$, that is, $\boldsymbol{y} \succ_n^\rho \boldsymbol{x}$. In the following axiom, we consider a sequence of menus that is *susceptible to shame*. A_{k+1} is more *susceptible to shame* than A_k for $k = 1, \cdots, n-1$. Let us denote $\geq^*$ on $[0,1]$, by $p \geq^* q$ if $p > q$ or $p = q \in (0,1)$. This axiom states that *menu acyclicity* holds with shame aversion.

Axiom 11. (Shame-Averse Acyclicity): For any finite sequence $\{A_k\}_{k=1}^n$ that is *susceptible to shame* if

$$\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_1, A_2), \rho(\boldsymbol{x}_k, A_k) \geq^* \rho(\boldsymbol{x}_{k+1}, A_{k+1}) \text{ (for } 1 < k < n)$$

then $\rho(\boldsymbol{x}_n, A_n) \not\geq^* \rho(\boldsymbol{x}_n, A_1)$.

4.3 Result

Utility Representation.

Theorem 2. *The following statements are equivalent:*

(a) *ρ satisfies Continuity, Self-Interest, and Shame-Averse Acyclicity.*

(b) *There exists a tuple $\langle u, \varphi, g \rangle$ where $u : X \to \mathbb{R}$ is a continuous and weakly increasing function that is more selfish than φ, and $\varphi : X \to \mathbb{R}$ is a continuous function that represents $\succsim_n^\rho$, and $g : \mathbb{R} \to \mathbb{R}$ is a strictly convex function, such that*

$$\rho(A) = \arg\max_{\rho \in \Delta(A)} \sum_{\boldsymbol{x} \in A} \Big(u(\boldsymbol{x})\rho(\boldsymbol{x}) - c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) \Big)$$

where $c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) = g\big(\varphi(\boldsymbol{x})\big)\big(\rho(\boldsymbol{x})\big)$ for each $\boldsymbol{x} \in A$.

We call the model in Theorem 2 the *additive perturbed shame-averse utility* (henceforth, APU(SA)) if ρ satisfies the axioms in Theorem 2.

Menu-Invariant Utility. In the representation result, the cost function depends on items (*allocations*). Intuitively, in the theory of *social image concerns*, the decision maker cares about how others perceive their decision-making. Different items have different implementation costs due to image concerns, so the choice behavior itself impacts social image concerns. Thus, the decision maker has an incentive to engage in randomization. If they choose the most selfish item with probability 1, they would be perceived as a selfish person. However, if it perceived that there was an element of chance in them acquiring the most selfish item, the social image costs are lessened.

Concealing one's Image/Type. APU(SA) supposes that the utility function is *selfish*.[25] The desire for randomization stems from the fact that randomization mitigates the impact of the decision on the decision maker's image. Thus, the decision maker willingly opts for stochastic choice behavior.

Regularity. The APU(SA) is an example of a menu-invariant APU model, so the model satisfies *Regularity* (Fudenberg et al., 2014). This property differs from that of APU(IA). In this study, the incentive to randomization distinguishes *ex-ante fairness* from *social image concerns*. In the former case, we characterize the model as an application of the item-invariant APU. In the latter case, we characterize the model as an application of the menu-invariant APU.

Remark 2. *Consider the logistic-type item-dependent cost function as follows: For each* $\boldsymbol{x} \in X$,

$$c_{\boldsymbol{x}}(\rho(\boldsymbol{x}) = \eta(\boldsymbol{x})\rho(\boldsymbol{x}) \log \rho(\boldsymbol{x}),$$

The item-dependent parameter η *captures the level of shame aversion. As items are increasingly selfish or altruistic, the penalties on deliberate randomization increase. For example, let* $\eta(\boldsymbol{x}) = \eta^{\frac{1}{x_1}+\frac{1}{x_2}}$ *and* $\eta(\boldsymbol{x}) = \eta^{\frac{1}{x_1 x_2}}$. *Then, fewer normative items have higher penalties.*

Violations of Luce's IIA. The APU(SA) deviates from *Luce's Independence of Irrelevant Alternatives* (Luce's IIA). Luce states that the likelihood of choosing an item $\boldsymbol{x}$ relative to $\boldsymbol{y}$ is independent of what other items are available in the menu A.

Axiom 12. (Luce's IIA): For any $A, B \in \mathcal{A}$ and $\boldsymbol{x}, \boldsymbol{y} \in A \cap B$,

$$\frac{\rho(\boldsymbol{x}, A)}{\rho(\boldsymbol{y}, A)} = \frac{\rho(\boldsymbol{x}, B)}{\rho(\boldsymbol{y}, B)}.$$

We provide an example of the violation of Luce's IIA.

[25] Generally, in economic theory, we do not assume preferences for generosity such as *fairness concerns*. Saito (2015a) and Hashidate (2020a) study the attitude toward pure selfishness and altruism. In APU(SA), we allow for non-selfish utility functions.

Example 2. *(A Violation of Luce's IIA)*
Suppose that APU(SA) is represented by a logistic-type cost function with item-dependence. We assume that the parameter η is item-dependent and is defined by $c_{\boldsymbol{x}}(\rho(\cdot)) = \eta^{\frac{1}{x_1 x_2}}\rho(\cdot)\log\rho(\cdot)$ with $\eta = 10$. Consider $I = \{1,2\}$. Let $\boldsymbol{x} = (3,6), \boldsymbol{y} = (4,1)$, and $\boldsymbol{z} = (1,7)$. Then, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \approx 0.3679$, and $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\}) \approx 0.6321$. Moreover, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \approx 0.3679$, $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \approx 0.537$, and $\rho(z, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \approx 0.0952$. Thus, we have

$$\frac{\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\})}{\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\})} = \frac{0.3679}{0.6321} \neq \frac{0.3679}{0.5370} = \frac{\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})}{\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}).}$$

Stochastic Transitivity. Next, we consider the two conditions of *stochastic transitivity*. APU(SA) deviates from the *strong stochastic transitivity* .

Axiom 13. (Weak Stochastic Transitivity): For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$, if $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \geq 0.5$ and $\rho(\boldsymbol{y}, \{\boldsymbol{y}, \boldsymbol{z}\}) \geq 0.5$, then $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{z}\}) \geq 0.5$.

Axiom 14. (Strong Stochastic Transitivity) For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$, if $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \geq 0.5$ and $\rho(\boldsymbol{y}, \{\boldsymbol{y}, \boldsymbol{z}\}) \geq 0.5$, then $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{z}\}) \geq \max\{\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}), \rho(\boldsymbol{y}, \{\boldsymbol{y}, \boldsymbol{z}\})\}$.

First, we provide an example of a violation of *strong stochastic transitivity* . This counter example states that *Shame-Averse Acyclicity* is weaker than *Strong Stochastic Transitivity*.

Example 3. *(A Violation of Strong Stochastic Transitivity)*
Suppose that APU(SA) is represented by a logistic-type cost function with item-dependence. We assume that the parameter η is item-dependent and is defined by

$$c_{\boldsymbol{x}}(\rho(\cdot)) = \eta^{\frac{1}{x_1+1}+\frac{1}{x_2+1}}\rho(\cdot)\log\rho(\cdot)$$

with $\eta = 5$. Consider $I = \{1,2\}$. Let $\boldsymbol{x} = (10,0), \boldsymbol{y} = (9,1)$, and $\boldsymbol{z} = (8,2)$. We have $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \approx 0.5018 > 0.5$, and $\rho(\boldsymbol{y}, \{\boldsymbol{y}, \boldsymbol{z}\}) \approx 0.5880 > 0.5$. However, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{z}\}) \approx 0.5556 < 0.5880 \approx \max\{\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}), \rho(\boldsymbol{y}, \{\boldsymbol{y}, \boldsymbol{z}\})\}$, which is not consistent with Strong Stochastic Transitivity.

Uniqueness Result. We impose on *Positivity* and *Richness*, and then the following uniqueness result is obtained.

Proposition 4. *Suppose that ρ satisfies Positivity and Richness in addition to the axioms in Theorem 2. If $\langle u, \varphi, g\rangle$ and $\langle \widehat{u}, \widehat{\varphi}, \widehat{g}\rangle$ represent the same ρ, then the following holds: $a > 0$, $b_u, b_\varphi \in \mathbb{R}$ such that*

(i) $\widehat{u} = au + b_u$ *and* $\widehat{\varphi} = a\varphi + b_\varphi$*; and*

(ii) $\widehat{g} = ag$.

The uniqueness result states that both self-utility u and personal norm utility φ are unique up to positive affine transformations with the same unit a and $g' = ag$ with the same unit a. In general, menu-invariant APU is not unique. In APU(SA), cost functions have a psychological structure concerning shame aversion. This specific structure enables the identification of the model.

Remark 3. *Take an arbitrary menu $A \in \mathcal{A}$. Then, in APU(SA),*

$$\begin{aligned}
\rho(A) &= \arg\max_{\rho \in \Delta(A)} \sum_{\boldsymbol{x} \in A} \left[\widehat{u}(\boldsymbol{x})\rho(\boldsymbol{x}) - \widehat{g}\big(\widehat{\varphi}(\boldsymbol{x})\big)\rho(\boldsymbol{x})\right] \\
&= \arg\max_{\rho \in \Delta(A)} \sum_{\boldsymbol{x} \in A} \left[(au(\boldsymbol{x}) + b_u)\rho(\boldsymbol{x}) - \widehat{g}\big(a\varphi(\boldsymbol{x}) + b_\varphi\big)\rho(\boldsymbol{x})\right], \\
&= \arg\max_{\rho \in \Delta(A)} a \sum_{\boldsymbol{x} \in A} u(\boldsymbol{x})\rho(\boldsymbol{x}) + b_u - \sum_{\boldsymbol{x} \in A} \widehat{g}\big(a\varphi(\boldsymbol{x}) + b_\varphi\big)\rho(\boldsymbol{x}), \\
&= \arg\max_{\rho \in \Delta(A)} a \sum_{\boldsymbol{x} \in A} u(\boldsymbol{x})\rho(\boldsymbol{x}) + b_u - a \sum_{\boldsymbol{x} \in A} g\big(\varphi(\boldsymbol{x})\big)\rho(\boldsymbol{x}) + b^* \\
&= \arg\max_{\rho \in \Delta(A)} a \left(\sum_{\boldsymbol{x} \in A} \Big(u(\boldsymbol{x})\rho(\boldsymbol{x}) - g\big(\varphi(\boldsymbol{x})\big)\rho(\boldsymbol{x})\Big)\right) + b, \\
&= \arg\max_{\rho \in \Delta(A)} a \left(\sum_{\boldsymbol{x} \in A} \Big(u(\boldsymbol{x})\rho(\boldsymbol{x}) - c_{\boldsymbol{x}}(\rho(\boldsymbol{x}))\Big)\right) + b,
\end{aligned}$$

where $b = b_u + b^*$.

4.4 Deliberate Randomization stemming from Shame Aversion

In this subsection, we study the implications of APU(SA). In APU(SA), the cost function is *item or outcome-dependent* and stems from shame aversion. First, we show the corollary of the cost function in the APU(SA). Second, we provide numerical examples of the cost functions.

Corollary 1. *Suppose that ρ is represented by the APU(SA). Then, for any $\boldsymbol{x}, \boldsymbol{y} \in X$, if $\boldsymbol{y} \succsim_n^\rho \boldsymbol{x}$, then $c''_{\boldsymbol{y}}(\cdot) \leq c''_{\boldsymbol{x}}(\cdot)$.*

This corollary shows that less personally normative items have higher penalties of deliberate randomization.

4.5 Preference Reversals in Social Contexts

APU(SA) can provoke preference reversals such as the *Attraction* and *Compromise* effects. In the logistic-type model, whether preference reversals occur depends on the parameter η.

Example 4. *Consider the doubleton* $\{\boldsymbol{x}, \boldsymbol{y}\}$*, where* $\boldsymbol{x} = (5,5)$ *and* $\boldsymbol{y} = (7,2)$*. Suppose that APU(SA) is represented by the logistic-type cost function* $\eta(\boldsymbol{x}) = \eta^{\frac{1}{x_1}+\frac{1}{x_2}}$*. Suppose that* η *is sufficiently large.*[26] *Then,* $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) > \rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\})$*. However, by adding the decoy item* $\boldsymbol{z}$*,* $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) < \rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})$ *(Figure 2).*[27]

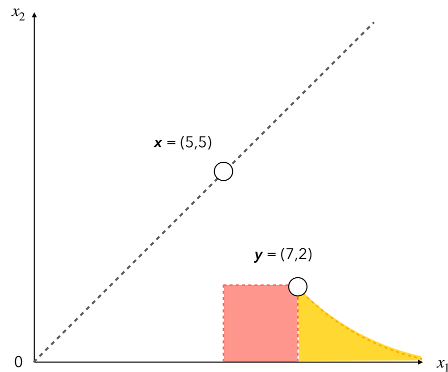


Figure 2: APU(SA) and Context Effects: When the "decoy" item $\boldsymbol{z}$ in the red zone is added, the *Attraction* effect occurs. When the "decoy" item $\boldsymbol{z}$ in the orange zone is added, the *Compromise* effect occurs.

5 Inequity Aversion versus Shame Aversion

As mentioned in the Introduction, it is not easy to identify the source of altruistic or prosocial behavior. Here, we compare those with inequity-averse preferences to those with image-conscious preferences. We consider how to identify the difference between them through observed behavior. In the following subsections, we study the difference between APU(IA) and APU(SA) in terms of *Luce's IIA* (preference reversals), *Regularity* (*Decoy* and *Menu-Size* Effects), and decision-making in the advantageous and disadvantageous cases.

5.1 IIA Revisited and Violations of Regularity

In this subsection, we provide an example of the preference reversal behavior of APU(SA), which does not occur in APU(IA). We take arbitrary items $\boldsymbol{x}$ and $\boldsymbol{y}$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) > \rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\})$, irrespective of social preferences such as inequity aversion and shame aversion. We consider the case in which adding a

[26] If $\eta \leq 41.43$, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \leq \rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\})$. In this case, preference reversals do not occur.

[27] These preference reversals can occur due to the attitude toward shame aversion (deliberate randomization). The behavioral patterns differ from *reference-dependence* and *limited attention*.

new item z in APU(IA), $\rho(x, \{x, y, z\}) > \rho(y, \{x, y, z\})$ holds, but $\rho(x, \{x, y, z\}) < \rho(y, \{x, y, z\})$ holds in APU(SA).

Deliberate Stochastic Behavior. Fix a doubleton $\{x, y\}$, where $x = (3, 4)$ and $y = (4, 1)$ with $\rho(x, \{x, y\}) > \rho(y, \{x, y\})$, irrespective of social preferences.[28] We add a new item $z = (2, 6)$ to the doubleton $\{x, y\}$ (Figure 7 in Appendix D).

First, we compare the choice probability of the item x from the menu $\{x, y, z\}$ with the choice probability of the item y from the menu $\{x, y, z\}$. Second, we compare the choice probability of the item x from the menu $\{x, y\}$ with the choice probability of the item x from the menu $\{x, y, z\}$.

Inequity Aversion. Consider the case in which stochastic choice behavior exhibits inequity-averse preferences. As stated above, the item $z = (2, 6)$ is the most altruistic item in the menu $\{x, y, z\}$. Moreover, *bmz* can be interpreted as the most "unfair" item in the menu. Let $\succsim_i^{\rho}$ be the binary relation over X that satisfies *inequity-averse preferences*. We obtain $x \succ_i^{\rho} y \succ_i^{\rho} z$. By definition, we also obtain $\rho(x, \{x, y, z\}) > \rho(y, \{x, y, z\}) > \rho(z, \{x, y, z\})$; that is, we have

$$\rho(x, \{x, y, z\}) > \rho(y, \{x, y, z\}).$$

Consider a numerical example with a logistic-type cost function. Let $\alpha = 0.9$ and $\beta = 0.7$.

No Preference Reversal. We study several cases of stochastic inequity-averse behavior. First, we consider the case of menu-dependent cost functions stemming from ex-post fairness explained in Example 3.1. Here, we assume that the parameter η is menu-dependent and is set as the mean of the Gini index of the items in menus. Then, we have $\rho(x, \{x, y\}) \approx 0.63$, and $\rho(y, \{x, y\}) \approx 0.37$. We infer that the decision maker prefers the item x to the item y.

Second, we consider the case of menu-dependent cost functions stemming from ex-ante fairness explained in Example 3.4. Then, we have $\rho(x, \{x, y\}) \approx 0.52$, and $\rho(y, \{x, y\}) \approx 0.48$.

Third, we consider the case of inequity-averse preferences in Saito (2013). In the same way, assume that $\alpha = 0.9$, $\beta = 0.7$, and $\gamma = 0.5$. Then, we have $\rho(x, \{x, y\}) = 0.75$, and $\rho(y, \{x, y\}) = 0.25$.

Violations of Regularity. The *inequity-averse* decision maker deviates from *Regularity*. For example,

$$\rho(y, \{x, y\}) < \rho(y, \{x, y, z\}),$$

[28] We can design such a menu, irrespective of inequity-averse or shame-averse preference. For example, we can collect data on (α, β) in (pilot) experiments. See Appendix D.

which deviates from *Regularity*. Consider the above numerical example. In particular, the case of cost functions stemming from ex-post fairness is not consistent with *Regularity*. We have $\rho(x, \{x, y, z\}) \approx 0.6183$, $\rho(y, \{x, y, z\}) \approx 0.3816$, and $\rho(z, \{x, y, z\}) \approx 0.0001$. Thus, we have $\rho(y, \{x, y\}) \approx 0.37 < 0.38 \approx \rho(y, \{x, y, z\})$. In the other cases stated above, *Regularity* is satisfied.[29]

Social Image Concerns: the Case of Shame Aversion. Next, we consider the case in which stochastic choice behavior exhibits shame-averse preferences. As mentioned in Section 4, the decision maker has two utility functions $\langle u, \varphi \rangle$, where u is the self-utility and φ is the personal-norm utility. The decision maker wants to act selfishly, but cares about how other passive agents perceive their behavior. They feel *shame* as they deviate from their ethical norm. In sum, the decision maker faces the trade-off between their selfishness (u) and their norms (φ).

Preference Reversal. In this case as the added item $z = (2, 6)$ is more altruistic than $x = (3, 4)$ and $y = (4, 1)$, the item z makes the trade-off between private ranking and personal norms (altruism). On the other hand, by adding the item $z = (2, 6)$, the decision maker may have an incentive to engage in selfish behavior; that is,

$$\rho(x, \{x, y, z\}) \leq \rho(y, \{x, y, z\}),$$

which is a different stochastic behavior stemming from inequity-averse preferences. For example, consider a logistic-type cost function. We assume that the parameter η is item-dependent. Then, we have $c_x(\rho(x)) = \eta^{\frac{1}{x_1} + \frac{1}{x_2}} \rho(x) \log \rho(x)$. Let $\eta = 10$. We have $\rho(x, \{x, y\}) \approx 0.57$, and $\rho(y, \{x, y\}) \approx 0.43$. Moreover, we obtain $\rho(x, \{x, y, z\}) \approx 0.34$, $\rho(y, \{x, y, z\}) \approx 0.38$, and $\rho(z, \{x, y, z\}) \approx 0.28$.

Figure 3 shows that as η increases, the decision maker takes deliberately stochastic behavior. As a result, preference reversals can occur.

Consistency with Regularity. Since APU(SA) is a special case of the Menu-Invariant APU, the model is consistent with the axiom of *Regularity*, that is,

$$\rho(y, \{x, y\}) \geq \rho(y, \{x, y, z\}).$$

Consider the above numerical example. We have $\rho(y, \{x, y\}) \approx 0.43 \geq 0.38 \approx \rho(y, \{x, y, z\})$.

Remark 4. *In social contexts, whether Regularity is satisfied depends on the motivation behind deliberate randomization. If social preferences are outcome-based, as is the case with*

[29] First, in the case of cost functions stemming from ex-ante fairness, resulting stochastic choice behavior is as follows. We have $\rho(x, \{x, y, z\}) \approx 0.44$, $\rho(y, \{x, y, z\}) \approx 0.42$, and $\rho(z, \{x, y, z\}) \approx 0.14$. Second, in the case of Saito (2013), the resulting stochastic choice behavior is as follows. We have $\rho(x, \{x, y, z\}) = 0.75$, $\rho(y, \{x, y, z\}) = 0.25$, and $\rho(z, \{x, y, z\}) = 0.00$.

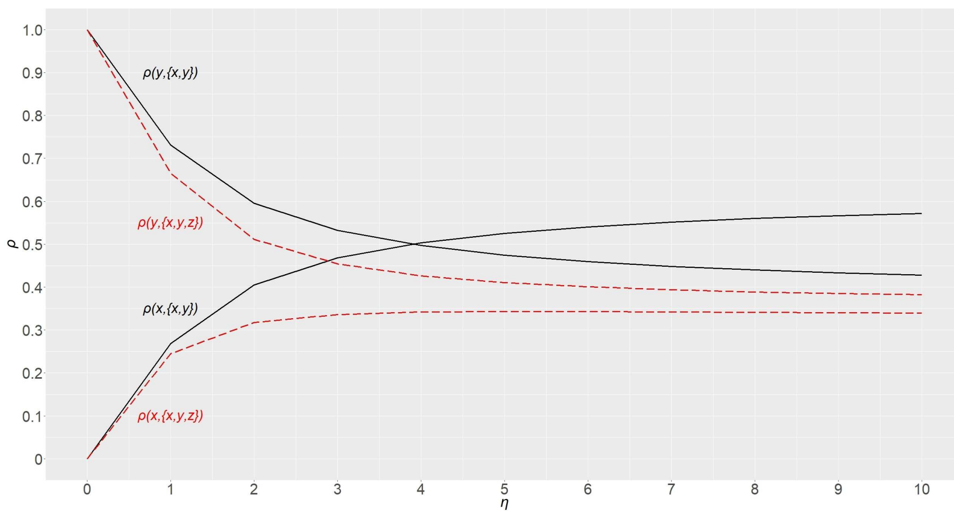


Figure 3: APU(SA) and Preference Reversals: The solid lines depict the choice probabilities from the doubleton $\{x, y\}$. The red dashed lines depict the choice probabilities from the menu $\{x, y, z\}$.

inequity aversion, then Regularity is violated. If social preferences are image-conscious, as is the case with shame aversion, then Regularity is satisfied.

Given stochastic choice data, we can statistically test whether the observed behavior stems from the outcome-based preferences by using the property of Regularity.[30]

The procedure consists of two steps. First, we check whether the stochastic choice data is represented by additive perturbed utility (APU) by testing whether the data satisfy the property of Acyclicity.[31] *Subsequently, we perform the one-sided test under the hypotheses* $H_0 : \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) - \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) = 0$ *and* $H_1 : \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) - \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) > 0$*. If* H_0 *is rejected, we can say that the observed behavior stems from outcome-based preferences.*

5.2 Decoy Effects

Choice probabilities can differ according to the motivations behind observed behavior, that is, *inequity aversion* or *shame aversion*. Here, we consider a doubleton $\{\boldsymbol{x}, \boldsymbol{y}\}$ where $\boldsymbol{x} = (5, 5)$ and $\boldsymbol{y} = (7, 4)$, and studies the change in choice probabilities when an "unfair" item is added into the doubleton (Figure 4).[32] We measure the "unfairness" of items by using the *Gini* index.

In APU(IA), the probability of choosing the fair item $\boldsymbol{x}$ increases. For the third added item, the fair item seems to be *attractive*. On the other hand, APU(SA) is consistent with *Regularity*, the choice probability of choosing $\boldsymbol{x}$ decreases.

5.3 Menu Size Effects

We explain the implications of the menu-size effect on our findings. Here, we consider a sequence of "selfish" and "altruistic" items. We consider the sequence of menus such that an item $\boldsymbol{z}$ is sequentially added in ascending order of its "unfairness."

First, we consider the sequence of altruistic menus. Let $A_2 = \{(5, 5), (7, 4)\}$. The next menu is denoted by $A_3 = A_2 \cup \{\boldsymbol{z}_3\}$, where $\boldsymbol{z}_3 = (3, 8)$. The menu A_4 is constructed by $A_4 = A_3 \cup \{\boldsymbol{z}_4\}$, where $\boldsymbol{z}_4 = (3, 9)$ is more unfair than $\boldsymbol{z}_3$. In this way, we add a more "unfair" item sequentially.

As Figure 5 shows, the choice probability of $\boldsymbol{x}$ in APU(IA) with the *ex-post* cost function increases. In APU(IA) with the *ex-ante* cost function, the choice probability of $\boldsymbol{x}$ remains at the same level. On the other hand, the choice probability of $\boldsymbol{x}$ in APU(SA) decreases. The difference mainly stems from the property of *Regularity*.

[30] To rigorously identify the source of observed behavior, for example, whether the observed behavior stems from the inequity-averse preferences, we need to statistically test not only Regularity but all other axioms on APU(IA) one by one.

[31] Fudenberg et al. (2015) shows that a stochastic choice data satisfies acyclicity if and only if it is represented by a APU.

[32] For *inequity-averse* preferences, assume $\alpha = 1, \beta = 0.6$, and η is the average of the Gini index of items in menus. For *shame-averse* preferences, assume that $u(\boldsymbol{x}) = x_1$ and $\varphi(\boldsymbol{x}) = x_1 x_2$ for all $\boldsymbol{x} \in X$. Let $c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) = \eta^{1/x_1 x_2} \rho(\boldsymbol{x}) \log \rho(\boldsymbol{x})$ for all $\boldsymbol{x} \in X$. Assume that $\eta = 2$.

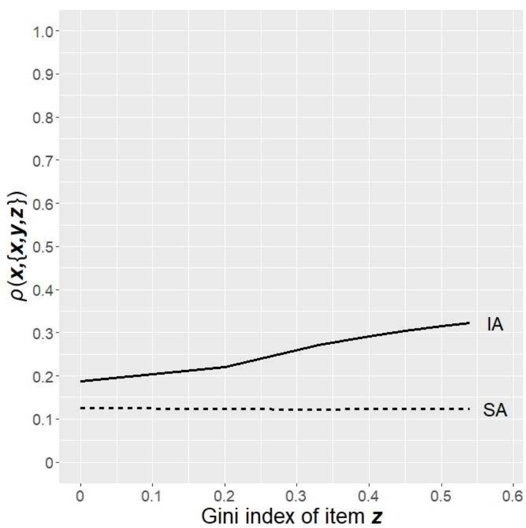


Figure 4: We have $\rho_{\text{IA}}(x, \{x, y\}) \approx 0.1874$ and $\rho_{\text{SA}}(x, \{x, y\}) \approx 0.1248$. As the Gini index of the third item increases, the third item is getting unfair. Note that in Figure 4, the x-axis = 0 depicts the choice probability of x from the menu $\{x, y\}$. For example, in the case of the Gini index 0.3, the item with a Gini coefficient 0.3, denoted by z, is added to the menu $\{x, y\}$, and the y-axis depicts the choice probability of x from the menu $\{x, y, z\}$.

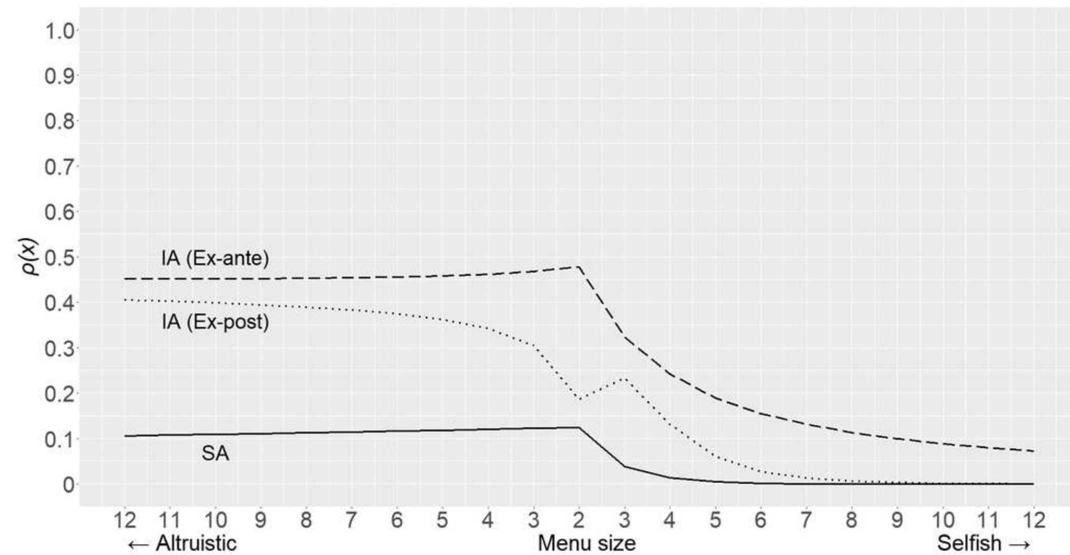


Figure 5: In APU(IA), we consider the two cases of cost functions. In APU(IA) with the ex-ante cost function studied in Example 3.4, the choice probability of the "fair" item x is consistent with *Regularity* in both selfish and altruistic menu sequences. In APU(IA) with the ex-post function used in Figure 4, the choice probability of x increases as the altruistic menu-size increases. In particular, the choice probability of x increases from approximately 20 % to approximately 40 %. In the case of the selfish menu sequence, at a menu size from 2 to 3, the choice probability increases. On the other hand, in APU(SA) with the cost function used in Figure 4, the choice probability is consistent with *Regularity*. In the altruistic menu sequence, the choice probability is approximately 10 %, and in the selfish menu sequence, the choice probability tends to 0 %.

Next, we consider the sequence of selfish menus. Let $A_2 = \{(5,5),(7,4)\}$. The next menu is denoted by $A_3' = A_2 \cup \{z_3'\}$, where $z_3' = (8,3)$. The menu A_4' is constructed by $A_4' = A_3' \cup \{z_4'\}$, where $z_4' = (9,3)$ is more unfair than z_3'. In this way, we add a more "unfair" item sequentially.

As Figure 5 shows, the choice probability of $\boldsymbol{x}$ decreases as the menu-size increases in the three models. This behavioral pattern is consistent with the property of the *Regularity*.

5.4 Deterministic or Stochastic? Advantageous Case

The behavioral patterns between APU(IA) and APU(SA) differ in the advantageous case; that is, doubletons that include items where the dictator's payoff is larger than the recipient's payoff.

Table 1: A Numerical Example of APU(IA): the Advantageous Case

menu	$\rho(\boldsymbol{x})$	$\rho(\boldsymbol{y})$
$\{(5,4),(6,3)\}$	0.1419	0.8581
$\{(5,4),(8,2)\}$	0.0145	0.9855
$\{(5,4),(10,1)\}$	0.0037	0.9963

Table 2: A Numerical Example of APU(SA): the Advantageous Case

menu	$\rho(\boldsymbol{x})$	$\rho(\boldsymbol{y})$
$\{(5,4),(6,3)\}$	0.4255	0.5745
$\{(5,4),(8,2)\}$	0.3189	0.6811
$\{(5,4),(10,1)\}$	0.4325	0.5675

First, we consider the case of APU(IA) (Table 1), where $\boldsymbol{x} = (5,4)$ and $\boldsymbol{y} = (6,3),(8,2),(10,1)$. Without loss of generality, we assume that the cost function of randomization is based on *ex-post fairness*.[33] In this case, inequity-aversion exhibits *guilt avoidance*. Since the indifference curves in this region are linear, APU(IA) leads to "almost" deterministic behavior.

Second, we consider the case of APU(SA) (Table 2). We assume that the *item-dependency* cost function is as follows: $c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) = \eta^{\frac{1}{x_1}+\frac{1}{x_2}}\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})$, where $\eta =$

[33] Assume that the Fehr and Schmidt (1999)'s inequity-averse preference with $\alpha = 0.9$ and $\beta = 0.3$ holds. The cost function is menu-dependent as the mean of the Gini index of items in menus (Example 3.2). As the parameter β increases, the choice behavior can be more and more stochastic (Figure 1b).

10. The resulting behavior is *stochastic* because of the trade-off between private ranking described by u and the personal norm described by φ.

5.5 Deterministic or Stochastic? the Disadvantageous Case

The behavioral patterns between APU(IA) and APU(SA) differ in the disadvantageous case; that is, doubletons that include items where the recipient's payoff is larger than the dictator's payoff.

Table 3: A Numerical Example of APU(IA): the Disadvantageous Case

menu	$\rho(\boldsymbol{x})$	$\rho(\boldsymbol{y})$
$\{(4,5),(3,6)\}$	0.9998	0.0002
$\{(4,5),(2,8)\}$	1.0000	0.0000
$\{(4,5),(1,10)\}$	1.0000	0.0000

Table 4: A Numerical Example of APU(SA): the Disadvantageous Case

menu	$\rho(\boldsymbol{x})$	$\rho(\boldsymbol{y})$
$\{(4,5),(3,6)\}$	0.5909	0.4091
$\{(4,5),(2,8)\}$	0.6612	0.3388
$\{(4,5),(1,10)\}$	0.6686	0.3314

First, we consider the case of APU(IA) (Table 3), where $\boldsymbol{x} = (4,5)$ and $\boldsymbol{y} = (3,6),(2,8),(1,10)$. We assume that the cost function of randomization is based on *ex-post fairness*.[34] In this case, inequity-aversion exhibits *envy*. Since the indifference curves in this region are linear, APU(IA) leads to (almost) deterministic behavior.

Second, we consider the case of APU(SA) (Table 4). We assume that the *item-dependency* cost function is as follows: $c_{\boldsymbol{x}}(\rho(\boldsymbol{x}) = \eta^{\frac{1}{x_1}+\frac{1}{x_2}}\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})$, where $\eta = 10$. In the same way as the advantageous case, the resulting behavior is *stochastic* because of the trade-off between private ranking described by u and the personal norm described by φ.

[34] Assume that the Fehr and Schmidt (1999)'s inequity-averse preference with $\alpha = 0.9$ and $\beta = 0.7$ holds. The cost function is menu-dependent as the maximizer of the Gini index of items in menus.

6 Discussions

6.1 Probabilistic Dictator Games

Social Preferences under Risk. Miao and Zhong (2018) experimentally investigate *preferences for randomization* in social contexts. By using the random lottery mechanism, they study eleven pairs of allocations (*items*).[35] The doubletons indicate several concerns regarding equity, efficiency, and selfishness. For example, in a menu $\{(20,0),(0,20)\}$, the decision maker may care about *ex-ante fairness* and thus employ randomization.

In their experiment, subjects tended to randomize items for almost all menus except the menu where $\{(20,0),(0,0)\}$. Interestingly, when *inequity* is less extreme, the decision maker tries to randomize items owing to *ex-ante fairness*. For example, subjects tend to randomize items in $\{(16,4),(4,16)\}$ more than in $\{(20,0),(0,20)\}$. We can interpret this tendency as a menu-dependent deliberate randomization. In APU(IA), this stochastic choice behavior is captured by menu-dependent cost functions.[36]

In the doubletons in Miao and Zhong (2018), Saito (2013) predicted that the decision maker would randomize items in all of the menus except for $\{(20,0),(0,20)\}$ and $\{(16,4),(4,16)\}$.[37] On the other hand, APU(IA) allows for stochastic choice behavior that is consistent with Miao and Zhong (2018).

Miao and Zhong (2018) argue that some subjects' behavioral patterns (42.0%) are not consistent with *inequity-averse preferences*. There are several motivations, such as social image concerns. In behavioral patterns beyond Miao and Zhong (2018), subjects mainly have either selfish or efficiency concerns. Social image concerns such as shame aversion capture such stochastic behavior as the trade-off between selfishness (u) and efficiency ($\varphi = x_1 + x_2$).

6.2 Relation to Random Utility

In this subsection we compare APU(IA) and APU(SA) with RU. First, we describe the relationship between APU(IA) and Random Inequity-Averse Utility. Second,

[35] Miao and Zhong (2018) provide experimental evidence against a number of weak assumptions, including (i) *first-order stochastic dominance* (FOSD), (ii) proportional monotonicity of ex-post fairness, and (iii) monotonicity for both ex-ante and ex-post fairness.

[36] Consider a logistic-type cost function with menu-dependence ($c_A(\rho(\boldsymbol{x})) = \eta(A)\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})$). Let $\eta(\{(20,0),(0,20)\}) = 25$ and $\eta(\{(16,4),(4,16)\}) = 50$. Then, the decision maker randomizes the items in the two menus, and in particular, the choice behavior is more deliberately stochastic in $\{(16,4),(4,16)\}$, compared to $\{(20,4),(0,20)\}$. For an inequity-averse preference, let $\alpha = 1$ and $\beta = 0.6$. We have $\rho(\{(20,0),(0,20)\}) \approx (0.75, 0.25)$, and $\rho(\{(16,4),(4,16)\}) \approx (0.58, 0.42)$, respectively.

[37] If γ is not sufficiently large, then the decision maker takes deterministic behavior. Assume that the parameters of inequity-aversion $\alpha = 1$ and $\beta = 0.6$. Then, if $\gamma \leq 0.875$, then the decision maker randomizes them.

we describe the relationship between APU(SA) and Random Shame-Averse Utility.

Random Inequity-Averse Utility. In RU, preferences are stochastic, and the resulting behavior differs from deliberately stochastic choice behavior. RU satisfies *Regularity*, but APU(IA) generally does not. Following Fudenberg et al. (2015), we consider the case in which APU(IA) satisfies *Regularity*; that is, APU(IA) is in the class of Random Inequity-Averse Utility. Let u_{IA} be an inequity-averse utility with the parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. $\mathcal{U}_{IA}$ is the set of inequity-averse utility functions. Let $\boldsymbol{\mu}$ be a probability measure on $\mathcal{U}_{IA}$.

Formally, we define *Random Inequity-Averse Utility* as follows.

Definition 10. A stochastic choice rule ρ is *Random Inequity-Averse Utility* if there exists a pair $(\mathcal{U}_{IA}, \boldsymbol{\mu})$, where the set of inequity-averse utility function u_{IA} and $\boldsymbol{\mu}$ is a probability measure on $\mathcal{U}_{IA}$, such that

$$\rho(\boldsymbol{x}, A) = \mu(\{u_{IA} \in \mathcal{U}_{IA} | u_{IA}(\boldsymbol{x}) \geq u_{IA}(\boldsymbol{y}) \text{ for all } \boldsymbol{y} \in A\})$$

where u_{IA} is represented by $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ with $\boldsymbol{\alpha} := (\alpha_i)_{i\in S}$ is a profile with $\alpha_i \geq 0$ for each $i \in S$, and $\boldsymbol{\beta} := (\beta_i)_{i\in S}$ is a profile with $\beta_i \geq 0$ for each $i \in S$.

The model states that inequity-averse utility changes stochastically. The choice probability of items depends on how frequently items are preference-maximal conditional on inequity-averse utilities.

We consider the case in which APU(IA) corresponds to RU. Taking u_{IA}, we assume that the "average" level of utility of items is evaluated by $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. However, the actual utility is $u_{IA}(\boldsymbol{x}) + \varepsilon_{\boldsymbol{x}}$. Let $\{\varepsilon_{\boldsymbol{x}}\}$ be a distribution, where $\varepsilon_{\boldsymbol{x}} \in \mathbb{R}^X$ is a random variable. We say that $\{\varepsilon_{\boldsymbol{x}}\}$ is *exchangeable* if for any permutation π, $(\varepsilon_1, \cdots, \varepsilon_n)$ and $(\varepsilon_{\pi(1)}, \cdots, \varepsilon_{\pi(n)})$ have the same distribution. We say that ρ is a *Symmetric* Random Inequity-Averse Utility if there exists a pair $\langle(\boldsymbol{\alpha}, \boldsymbol{\beta}), \{\varepsilon_{\boldsymbol{x}}\}\rangle$, where $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ with $\boldsymbol{\alpha} := (\alpha_i)_{i\in S}$ is a profile with $\alpha_i \geq 0$ for each $i \in S$, and $\boldsymbol{\beta} := (\beta_i)_{i\in S}$ is a profile with $\beta_i \geq 0$ for each $i \in S$, and $\{\varepsilon_{\boldsymbol{x}}\}$ is *exchangeable* such that

$$\rho(\boldsymbol{x}, A) = \mathbb{P}\left\{u_{IA}(\boldsymbol{x}) + \varepsilon_{\boldsymbol{x}} \geq \max_{\boldsymbol{y}\in A} u_{IA}(\boldsymbol{y}) + \varepsilon_{\boldsymbol{y}}\right\}$$

where u_{IA} is represented by$(\boldsymbol{\alpha}, \boldsymbol{\beta})$. For simplicity, suppose $I = \{1, 2\}$. Consider an item $\boldsymbol{x} \in X$ with $x_1 > x_2$. Let $\tilde{\alpha} = \alpha + \varepsilon_\alpha$, where $\varepsilon_\alpha \in \mathbb{R}$ and $\varepsilon_\alpha(x_1 - x_2) = \varepsilon_{\boldsymbol{x}}$. Then, random inverse utility is *symmetric*.

Finally, we describe the relationship between APU(IA) and *random expected utility* . Gul and Pesendorfer (2006) provides an axiomatic foundation for *Random Expected Utility*. In addition to *Regularity*, Gul and Pesendorfer (2006) requires that ρ satisfies *Linearity*, *Extremeness*, and *Mixture-Continuity*. Here, we provide the example that APU(IA) deviates from *Linearity*. For each $\lambda \in (0, 1)$, $\lambda\boldsymbol{x} + (1 - \lambda)\boldsymbol{y} = (\lambda x_i + (1 - \lambda)y_i)_{i\in I}$,

Axiom 15. (Linearity): For any $A \in \mathcal{A}$ with $\boldsymbol{x} \in A$, $\boldsymbol{y} \in A$, and $\lambda \in (0,1)$,

$$\rho(\boldsymbol{x}, A) = \rho(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda A + (1-\lambda)\{\boldsymbol{y}\}).$$

Example 5. *(Violation of Linearity)*
Suppose that APU(IA) is represented by a logistic-type cost function. We assume that the parameter η is menu-dependent and is set as the maximizer of the Gini index of the items in menus. Consider $I = \{1,2\}$. Let $\boldsymbol{x} = (10,0), \boldsymbol{x}' = (4,6)$, and $\boldsymbol{y} = (0,10)$. For an inequity-averse preference, let $\alpha = 1$ and $\beta = 0.6$ (Fehr and Schmidt, 1999). Let $A = \{\boldsymbol{x}, \boldsymbol{x}'\}$. Then, $\rho(\boldsymbol{x}, A) \approx 0.88$, and $\rho(\boldsymbol{x}', A) \approx 0.12$. By letting $\lambda = 0.5$, consider $\lambda A + (1-\lambda)\{\boldsymbol{y}\} = \{(5,5),(2,8)\}$. Then, $\rho(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda A + (1-\lambda)\{\boldsymbol{y}\}) = 1$ and $\rho(\lambda \boldsymbol{x}' + (1-\lambda)\boldsymbol{y}, \lambda A + (1-\lambda)\{\boldsymbol{y}\}) = 0$. This is not consistent with Linearity.

Random Shame-Averse Utility. We consider the case of *Random Shame-Averse Utility*. For simplicity, assume that ρ satisfies *Linearity*.[38] Remember that APU(SA) satisfies *Regularity*. We say that ρ is a Random Shame-Averse Utility if there exists a tuple $(u, \varphi, \{\varepsilon_{\boldsymbol{x}}\})$ such that

$$\rho(\boldsymbol{x}, A) = \mathbb{P}\left\{ u(\boldsymbol{x}) + \varphi(\boldsymbol{x}) + \varepsilon_{\boldsymbol{x}} \geq \max_{\boldsymbol{y} \in A} u(\boldsymbol{y}) + \varphi(\boldsymbol{y}) + \varepsilon_{\boldsymbol{y}} \right\}.$$

Stochastic choice behavior stems from the stochasticity of both self-utility u and personal norm utility φ. It is interesting that the random variable $\varepsilon_{\boldsymbol{x}}$ stems from the stochasticity of the norm-utility φ. Compared with social norms, personal norms are not stable. The multiple norm utilities lead to stochastic choice behavior.

7 Related Literature

In this section we provide a literature review. First, we cover the literature on deliberate randomization (stochastic choices). Second, we cover the literature on inequity aversion. Third, we cover the literature on image concerns.

Deliberate Randomization. Our models are related to *additive perturbed utility* (APU) in Fudenberg et al. (2015). We applied their concept of APU to social contexts. APU(IA) is an example of an item-invariant APU (Fudenberg et al., 2014). We introduced a weaker version of the *acyclic* conditions in stochastic choices. With inequity-averse preferences, we characterized APU(IA), and that the model is in the class of item-invariant APU.

In the same way, APU(SA) is an example of menu-invariant APU (Fudenberg et al., 2014). We relaxed the acyclic condition in Fudenberg et al. (2015) to characterize APU(SA). We recovered the identification of APU(SA).

[38]The shame function $g : \mathbb{R}_+ \to \mathbb{R}$ is *linear* if $g(a) = a$ for all $a \in \mathbb{R}_+$.

Cerreia-Vioglio et al. (2019) characterize a deliberately stochastic choice model. They provide a new acyclic condition with *first-order stochastic dominance* (FOSD) called *Rational Mixing*. Our axiom, *Inequity-Averse Mixing*, is similar to their key axiom. Their axiom requires a lack of incentive to randomization in the two lotteries p and q, if p is first-order stochastically dominates q. Meanwhile, we require that there is no incentive to randomization in the two items $\boldsymbol{x}$ and $\boldsymbol{y}$ if $\boldsymbol{x}$ and $\boldsymbol{y}$ are monotone with respect to equal allocations, that is, $x \geq y \Rightarrow (x, \cdots, x) \succsim_i (y, \cdots, y)$.

Inequity Aversion. APU(IA) is in the class of inequity-averse preferences. Saito (2013) develops an axiomatic model of a convex combination of *ex-ante* and *ex-post* fairness, called the *expected inequity averse* model (EIA).[39] As mentioned in Section 3, APU(IA) is more general than EIA.

In Miao and Zhong (2018), they experimentally studied inequity-averse preferences, including ex-ante and ex-post fairness. They observed a behavioral pattern that was not consistent with Saito (2013). As mentioned in subsection 6.1, we can allow for the behavioral pattern in Miao and Zhong (2018).

Social Image Concerns: Shame Aversion. Dillenberger and Sadowski (2012) is the seminal axiomatic study on social image concerns. APU(SA) is more generalizable than the shame-averse model of Dillenberger and Sadowski (2012) because (i) we characterize a stochastic shame-averse choice model, and (ii) we allow for item-dependent shame aversion. We find that APU(SA) is consistent with *Regularity*, one of the most common properties of stochastic choices.

8 Concluding Remarks

In this study, we applied APU (Fudenberg et al., 2015) in social contexts. We studied the motivation of deliberately stochastic choice behavior in social contexts and characterized the two sources of deliberate randomization (see Table 5). One source of such behavior is an altruistic preference, in particular, inequity-averse preference (APU(IA)), and the other is an image-conscious preference, in particular, shame-averse preference (APU(SA)).

Table 5: Summary of Results

APU	Acyclic Condition	Axioms
APU(IA)	Item Acyclicity	Inequity-Averse Mixing
APU(SA)	Menu Acyclicity	Shame-Averse Acyclicity

[39] See, for the axiomatic study of *ex-post* fairness, Rohde (2010).

The axiomatizations made in this study have made it possible for us to identify the motivations behind stochastic choice behavior. We applied the property of *Regularity* to provide an example to distinguish between inequity-averse behavior and shame-averse behavior. Different motivations for deliberate randomization in social contexts have different behavioral patterns in choice probabilities.

In this study, the behavioral patterns of preference reversals depend on the penalty associated with deliberate randomization. The motivation can differ from *reference points* and *limited attention*. Thus, the sources of preference reversals differ.[40]

There are some limitations in our study, which are opportunities for further research. Although some experimental evidence such as DellaVigna et al. (2012) shows that the existence of social pressures affects prosocial behavior, there is little evidence to test the motivation behind prosocial behavior. Miao and Zhong (2018) study inequity-averse preferences including both ex-post and ex-ante fairness. Researchers can also test image-conscious preferences in similar experimental environments. Estimating cost functions in both inequity-averse preferences and image-conscious preferences in future research would add greatly to the findings of this study.

Table 6: Summary of Characterizations of Costs of Randomization

Utility	Cost's type	Sources	Results
Inequity aversion	Menu-dependence	Ex-post fairness	Proposition 2
Inequity Aversion	Menu-Dependence	Ex-Ante Fairness	Proposition 3
Selfishness	Item-Dependence	Shame Aversion	Corollary 1

As shown in Table 6, we have characterized APU(SA) with the utility that represents private ranking. However, we do not characterize the case of purely selfish utility functions. The decision maker maximizes their own payoffs, but their behavior is affected by *social pressures*, which can lead to prosocial behavior. The behavioral foundation is a further task.

Further tasks to apply our models to game theory also remain. For example, in *quantal response equilibria* (Goeree et al., 2016), we mainly assume that players have *random utility* (RU). We can apply APU(IA) or APU(SA) into a strategic setting and obtain different implications. We leave this application as a recommendation for future research.

[40]The characterizations of APU with *reference points* and *limited attention* are further tasks.

A Proofs of Theorems

A.1 Proof Outline of Theorem 1

We provide the proof outline of Theorem 1. In Step 1, we show that ρ is represented by an item-variant APU (Fudenberg et al., 2014) where $u : X \to \mathbb{R}$ is continuous function that is *monotone* with respect to equal items, i.e., for any $x, y \in \mathbb{R}$ with $x \geq y$,

$$(x, \cdots, x) \succsim_i (y, \cdots, y) \Leftrightarrow u(x, \cdots, x) \geq u(x, \cdots, x).$$

First, we show that ρ satisfying *Inequity-Averse Mixing* represents $\trianglerighteq$-maximization where the binary relation $\trianglerighteq$ over Δ associated with ρ (Definition 11) (Lemma 1). This captures a preference maximization as a *deliberate randomization*. Second, in Lemma 2, we show that if ρ satisfies *Inequity-Averse Mixing*, then ρ satisfies *item-acyclicity*. Third, we show that if ρ satisfies *Inequity-Averse Mixing*, then ρ is represented by an item-invariant APU in which u is a continuous monotone function with respect to equal items (Lemma 3).

In Step 2, we show that $\succsim_i^{\rho}$ satisfies the inequity-averse preference (Fehr and Schmidt, 1999) (Lemma 4). Then, u represents the Fehr and Schmidt (1999)'s inequity-averse preference. Thus, we obtain the desired utility representation.

A.2 Proof of Theorem 1

Proof of the Sufficiency Part

We show the sufficiency part. Suppose that ρ satisfies the axioms in Theorem 1.[41]

Step 1.

Deliberate Randomization. Define a binary relation $\trianglerighteq$ on Δ that captures a ranking of *deliberate randomization*. Let $\overline{\rho(A)} = \sum_{x \in A} \rho(x, A)x$, for each $A \in \mathcal{A}$.

Definition 11. We say p is *stochastically preferred to* q, i.e.,

$$p \trianglerighteq q \text{ if there exists } A \in \mathcal{A} \text{ such that } p = \overline{\rho(A)} \text{ and } q \in \text{co}(A).$$

We describe *deliberate randomization* as a binary relation $\trianglerighteq$ on Δ. Note that *outcome-mixtures* are as follows; for any $\boldsymbol{x}, \boldsymbol{y} \in X$ and $\lambda \in [0, 1]$,

$$\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} := (\lambda x_1 + (1 - \lambda)y_1, \cdots, \lambda x_n + (1 - \lambda)y_n).$$

Lemma 1. *Suppose that ρ satisfies Inequity-Averse Mixing. Then, ρ is represented by the $\trianglerighteq$-maximization associated with ρ.*

[41]In this paper, we study the case of $|I| = 2$, but in the proof, we consider the general case of $|I| = n$.

Proof. Suppose that $\trianglerighteq$ satisfies *monotonicity with respect to equal allocations*. Take $x \in \mathbb{R}$ and $\boldsymbol{y} \in X$ such that $\delta_{(x,\cdots,x)} \triangleright \delta_{\boldsymbol{y}}$. Note that $\trianglerighteq$ implies that, for any $\lambda \in [0,1)$, $\delta_{(x,\cdots,x)} \triangleright \lambda\delta_{(x,\cdots,x)} + (1-\lambda)\delta_{\boldsymbol{y}}$. Then, we have $\rho(\{(x,\cdots,x),\boldsymbol{y}\}) = \{(x,\cdots,x)\}$. This implies $(x,\cdots,x) \triangleright \boldsymbol{y}$. Thus, $\neg(\delta_{\boldsymbol{y}}\ tran(\trianglerighteq)\ \delta_{(x,\cdots,x)})$ where $tran(\trianglerighteq)$ is the smallest transitive binary relation $\trianglerighteq$ on Δ (Chambers and Echenique, 2016).

Fix $(x,\cdots,x)$ and $\boldsymbol{y}$ in X. Let $A_1 = \{(x,\cdots,x),\boldsymbol{y}\}$ and $A_2 = \{(x,\cdots,x)\}$. By *Inequity-Averse Mixing*, $\overline{\rho(A_1)} = \delta_{(x,\cdots,x)}$. Hence, we obtain $\delta_{(x,\cdots,x)} \triangleright \delta_{\boldsymbol{y}}$. Consider a finite sequence $\{A_i\}_{i=1}^k$ with $k \in \mathbb{N}$ with $\overline{\rho(A_1)} = \delta_{(x,\cdots,x)}$ and $\overline{\rho(A_i)} \in \mathrm{co}(A_{i-1})$ for $i = 2,\cdots,k$. Then, by *Inequity-Averse Mixing*, $\delta_{\boldsymbol{y}} \notin \mathrm{co}(A_k)$. Thus, we cannot have that $\delta_{\boldsymbol{y}}\ tran(\trianglerighteq)\ \delta_{(x,\cdots,x)}$ where $tran(\trianglerighteq)$. $tran(\trianglerighteq)$ is an extension of $\triangleright$. We can take $\triangleright$ as a complete extension of $tran(\trianglerighteq)$. By definition, for any $q \in \mathrm{co}(A)$, we have $\overline{\rho(A)}\ tran(\trianglerighteq)\ q \Rightarrow \overline{\rho(A)} \trianglerighteq q$. Hence, ρ satisfying *Inequity-Averse Mixing* is represented by $\trianglerighteq$-maximization. □

Binary Relation over X. Consider a set

$$\mathcal{D} := \{(\boldsymbol{x}, A) \in X \times A | \boldsymbol{x} \in A,\ A \in \mathcal{A}\}.$$

We induce a binary relation $\succsim_i^\rho$ over X, i.e., the ranking of *items*.

Definition 12. We say that $\boldsymbol{x}$ is preferred to $\boldsymbol{y}$, i.e.,

$$\boldsymbol{x} \succ_i^\rho \boldsymbol{y} \text{ if } \rho(\boldsymbol{x}, A) > \rho(\boldsymbol{y}, A) \text{ for some } A \ni \boldsymbol{x}, \boldsymbol{y}.$$

In the similar way, we say that $\boldsymbol{x}$ is indifferent to $\boldsymbol{y}$, i.e.,

$$\boldsymbol{x} \sim_i^\rho \boldsymbol{y} \text{ if } \rho(\boldsymbol{x}, A) = \rho(\boldsymbol{y}, A) \in (0,1) \text{ for some } A \ni \boldsymbol{x}, \boldsymbol{y}.$$

Let $\succsim_i^\rho$ be the union of $\succ_i^\rho$ and $\sim_i^\rho$, i.e., $\succsim_i^\rho := \succ_i^\rho \cup \sim_i^\rho$. Notice that $\succsim_i^\rho$ is associated with ρ. We consider the following axiom.

Axiom 16. (Item Acyclicity*): There does not exist a sequence of items $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m$ such that

$$\boldsymbol{x}_1 \succsim_i^\rho \boldsymbol{x}_2 \succsim_i^\rho \cdots \succsim_i^\rho \boldsymbol{x}_m \succ_i^\rho \boldsymbol{x}_1.$$

Continuity. By the continuity of ρ, $\succsim_i^\rho$ is also continuous. Since $\succsim_i^\rho$ is *monotone with respect to equal allocations*, we obtain the property of *Constant Equivalence*; for any $\boldsymbol{y} \in X$, there exists $x \in \mathbb{R}$ such that

$$\boldsymbol{y} \sim (x,\cdots,x).$$

Item Acyclicity. We show Lemma 2, regarding the *item-acyclic* condition.

Lemma 2. *If ρ satisfies Inequity-Averse Mixing, then $\succsim_i^\rho$ associated with ρ satisfies Item Acyclicity*.*

Proof. Suppose that ρ satisfies *Inequity-Averse Mixing*. Notice that $\trianglerighteq^*$ on Δ. We consider degenerate lotteries, which correspond to elements in X. For any $\boldsymbol{x}, \boldsymbol{y} \in X$,

$$\boldsymbol{x} \succsim_i^\rho \boldsymbol{y} \Rightarrow \delta_{\boldsymbol{x}} \trianglerighteq \delta_{\boldsymbol{y}}.$$

Remember that, for each menu $A \in \mathcal{A}$, $\overline{\rho(A)} = \sum_{\boldsymbol{x} \in A} \rho(\boldsymbol{x}, A)\boldsymbol{x}$. Take a menu $A_1 \ni \boldsymbol{x}_1, \boldsymbol{x}_2$. Then, $\overline{\rho(A)} = \sum_{\boldsymbol{x} \in A_1} \rho(\boldsymbol{x}, A_1)\boldsymbol{x}$. Without loss of generality, suppose $\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_2, A_1)$. By definition, we have $\boldsymbol{x}_1 \succ_i^\rho \boldsymbol{x}_2$. In the same way, we can construct $\overline{\rho(A_k)} \in \text{co}(A_{k-1})$ for $k \in \mathbb{N}$ with $\boldsymbol{x}_i \succ_i^\rho \boldsymbol{x}_{i+1}$, by using *constant equivalence*. Then, we have $\boldsymbol{x}_1 \succ_i^\rho \boldsymbol{x}_2 \succ_i^\rho \cdots \succ_i^\rho \boldsymbol{x}_k$. We can construct such a finite sequence of menus $\{A_i\}_{i=1}^k$. Suppose $(x, \cdots, x) \in \text{co}(A_k)$. Then, by *Inequity-Averse Mixing*, $(x, \cdots, x) \not\triangleright \overline{\rho(A_1)}$. Without loss of generality, assume $(x, \cdots, x) \in A_k$. Then, *Inequity-Averse Mixing* implies that $(x, \cdots, x) \not\succ_i^\rho \boldsymbol{x}$ for all $\boldsymbol{x} \in A_1$. Thus, *Item Acyclicity** is satisfied. □

Let us introduce some notation. For each $p, q \in [0, 1]$, we write $p \geq^* q$ if $p > q$ or $p = q$.

Axiom 17. (Item Acyclicity): There is no sequence $\{A_i\}_{i=1}^n$ such that if $\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_2, A_1)$ and $\rho(\boldsymbol{x}_k, A_k) \geq^* \rho(\boldsymbol{x}_{k+1}, A_k)$ for $1 < k < n$, then $\rho(\boldsymbol{x}_n, A_n) \not\geq^* \rho(\boldsymbol{x}_1, A_n)$.

Corollary 2. $\succsim_i^\rho$ *associated with ρ satisfies Item Acyclicity* if and only if ρ satisfies Item Acyclicity.*

Proof. This is clear by definition. Consider a sequence $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m$, and suppose that $\succsim_i^\rho$ is associated with ρ. Suppose that ρ satisfies *Item Acyclicity*. We have a menu $A_1 \in \mathcal{A}$ with $\boldsymbol{x}_1, \boldsymbol{x}_2 \in A_1$ such that $\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_1, A_1)$. By definition, $\boldsymbol{x}_1 \succsim_i^\rho \boldsymbol{x}_2$. In the same way, for $1 < k < m$, we can obtain $\boldsymbol{x}_k \succsim_i^\rho \boldsymbol{x}_{k+1}$. Hence, there does not exist a cycle such as $\boldsymbol{x}_m \succ_i^\rho \boldsymbol{x}_1$. □

Fudenberg et al. (2014) show Lemma 3. We use the result by applying Lemma 2.

Definition 13. We say that ρ has an *item-invariant* additive perturbed utility (APU) form if there exists a pair $\langle u, (c_A)_{A \in \mathcal{A}} \rangle$ where $u : X \to \mathbb{R}$ and, for each $A \in \mathcal{A}$, $c_A : [0, 1] \to \mathbb{R} \cup \{\infty\}$ is strictly convex and C^1 over $(0, 1)$, such that ρ is represented by

$$\rho(A) = \arg \max_{\rho \in \Delta(A)} \sum_{\boldsymbol{x} \in A} \Big(u(\boldsymbol{x})\rho(\boldsymbol{x}) - c_A(\rho(\boldsymbol{x})) \Big)$$

Lemma 3. *The following statements are equivalent.*

(i) ρ *is represented by item-invariant APU.*

(ii) *There exists a function* $u : X \to \mathbb{R}$ *such that* $u(\boldsymbol{x}) > u(\boldsymbol{y})$ *if* $\boldsymbol{x} \succ_i^\rho \boldsymbol{y}$, *and* $u(\boldsymbol{x}) = u(\boldsymbol{y})$ *if* $\boldsymbol{x} \sim_i^\rho \boldsymbol{y}$.

(iii) ρ *satisfies Item Acyclicity.*

Proof. The proof is based on Fudenberg et al. (2014)'s Proposition 8. The sufficiency part is as follows. Suppose that $u : X \to \mathbb{R}$ is continuous and monotone with respect to equal allocations. Take an arbitrary menu $A \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y} \in A$. Let $u(\boldsymbol{x}) > u(\boldsymbol{y})$ if $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{y}, A)$, and $u(\boldsymbol{x}) = u(\boldsymbol{y})$ if $\rho(\boldsymbol{x}, A) = \rho(\boldsymbol{y}, A)$. Let

$$\underline{w}(A) = \begin{cases} 0 & \text{if } \rho(\boldsymbol{x}, A) > 0, \forall \boldsymbol{x} \in A \\ \max\{u(\boldsymbol{x}) + \lambda(A) \mid A \in \mathcal{A}, \rho(\boldsymbol{x}, A) = 0\} & \text{otherwise.} \end{cases}$$

Construct a strictly increasing and continuous function $f_A \to \mathbb{R}$ for each $A \in \mathcal{A}$, such that (i) $f_A(0) = \underline{w}(A)$, and (ii) $f_A(\rho(\boldsymbol{x}, A)) = \lambda(A)$ if $\rho(\boldsymbol{x}, A) \in (0, 1)$. For each $A \in \mathcal{A}$, define a strictly convex C^1 function $c_A : [0, 1] \to \mathbb{R}$ by $c_A(q) = \int_0^q g_A(p)dp$. We have the optimality condition:

$$u(\boldsymbol{x}) - c'_A(\rho(\boldsymbol{x}, A)) + \lambda(A) \begin{cases} \geq 0 & \rho(\boldsymbol{x}, A) = 1 \\ = 0 & \rho(\boldsymbol{x}, A) \in (0, 1) \\ \leq 0 & \rho(\boldsymbol{x}, A) = 0. \end{cases}$$

The optimality condition holds for each menu $A \in \mathcal{A}$ with $\lambda(A) = 0$. □

Continuity of u. Remember that for each $A \in \mathcal{A}$, c_A is C^1 over $(0, 1)$. Take arbitrary $\boldsymbol{x}, \boldsymbol{y} \in X$. Consider the doubleton $\{\boldsymbol{x}, \boldsymbol{y}\}$. In Lemma 3, by the FOC of item-invariant APU, we have $u(\boldsymbol{x}) - u(\boldsymbol{y}) = c'_{\{\boldsymbol{x},\boldsymbol{y}\}}(\rho(\boldsymbol{x})) - c'_{\{\boldsymbol{x},\boldsymbol{y}\}}(1 - \rho(\boldsymbol{x}))$. For each $A \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y} \in A$, we have $\rho(\boldsymbol{x}, A) \geq \rho(\boldsymbol{y}, A) \Leftrightarrow u(\boldsymbol{x}) - c'_A(\rho(\boldsymbol{x}, A)) \geq u(\boldsymbol{y}) - c'_A(\rho(\boldsymbol{y}, A))$. By the continuity of ρ and c'_A, u is continuous.

Thus, u is continuous and monotone with respect to equal allocations. ρ satisfying *Inequity-Averse Mixing* and *Continuity* is represented by u and $(c_A)_{A \in \mathcal{A}}$.

Step 2.

We show that $\succsim_i^\rho$ is represented by the Fehr and Schmidt (1999)'s model. We show that $\succsim_i^\rho$ associated with ρ satisfies the axioms of Rohde (2010), which axiomatizes the Fehr and Schmidt (1999)'s *inequity-averse* utility model.

First, we state the axioms of Rohde (2010). We define the following due to the third axiom.

Definition 14. We say that an item $\boldsymbol{x} \in X$ is *constant* if $\boldsymbol{x} = (x, \cdots, x)$ where $x \in \mathbb{R}$.[42]

Moreover, we define the following due to the fifth axiom.

Definition 15. We say that two items $\boldsymbol{x}$ and $\boldsymbol{y}$ are *covalent*, if for every individual $i = 2, \cdots, n$, the deviation has the same sign in both distributions: $d_i(\boldsymbol{x})d_i(\boldsymbol{y}) \geq 0$ where $d_i(\boldsymbol{x}) = x_i - x_1$.

Axiom 18. (Rohde, 2010): $\succsim_i^\rho$ satisfies the following.

(i) (Weak Order): $\succsim_i^\rho$ is *complete* and *transitive*.

(ii) (Continuity): The sets $\{\boldsymbol{x} \in X \mid \boldsymbol{x} \succsim_i^\rho \boldsymbol{y}\}$ and $\{\boldsymbol{x} \in X \mid \boldsymbol{y} \succsim_i^\rho \boldsymbol{x}\}$ are closed in X.

(iii) (Constant Monotonicity): For any $x, y \in \mathbb{R}$,

$$x \geq y \Rightarrow (x, \cdots, x) \succsim_i^\rho (y, \cdots, y).$$

(iv) (Constant Equivalence): For any $\boldsymbol{x} \in X$, there exists $z \in \mathbb{R}$ such that $\boldsymbol{x} \sim_i^\rho (z, \cdots, z)$.

(v) (Covalent Additivity): For any items $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ that are pairwise covalent,

$$\boldsymbol{x} \succ_i^\rho \boldsymbol{y} \Leftrightarrow \boldsymbol{x} + \boldsymbol{z} \succ_i^\rho \boldsymbol{y} + \boldsymbol{z}.$$

(vi) (inequity Aversion): $\succsim_i$ satisfies (a) *envy* and (b) *guilt*: For each $j \in S$,

  (a) (Envy): For any $z \in \mathbb{R}_+$, $\mathbf{0} \succsim_i^\rho (0, (z, \cdots, z)_{k \in I \setminus \{j\}})$ where $\mathbf{0} := (0, \cdots, 0)$.

  (b) (Guilt): For any $z \in \mathbb{R}_-$, $\mathbf{0} \succsim_i^\rho (0, (z, \cdots, z)_{k \in I \setminus \{j\}})$ where $\mathbf{0} := (0, \cdots, 0)$.

Lemma 4. *$\succsim_i^\rho$ associated with ρ satisfies Axiom 18.*

Proof. We show that $\succsim_i^\rho$ associated with ρ satisfies the Rohde (2010)'s axioms. First, we verify that $\succsim_i^\rho$ is a weak order. By definition, $\succsim_i^\rho$ is complete.

By *Item Acyclicity* and Lemma 3, $\succsim_i^\rho$ is transitive. Take arbitrary $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in X$. Without loss of generality, suppose $\boldsymbol{x} \succsim_i^\rho \boldsymbol{y}$ and $\boldsymbol{y} \succsim_i^\rho \boldsymbol{z}$. By Lemma 3, we obtain $u(\boldsymbol{x}) \geq u(\boldsymbol{y}) \geq u(\boldsymbol{z})$, so $u(\boldsymbol{x}) \geq u(\boldsymbol{z}) \Leftrightarrow \boldsymbol{x} \succsim_i^\rho \boldsymbol{z}$. Notice that $\succsim_i^\rho := \succ_i^\rho \cup \sim_i^\rho$, where $\succ_i^\rho$ denotes the asymmetric (strict) part, and $\sim_i^\rho$ denotes the symmetric (indifference) part. Since $\succsim_i^\rho$ is complete, $\succsim_i^\rho$ is transitive (Chambers and Echenique, 2016).

[42] In this paper, we call *constant* items *fair* or *equity* items.

By Step 1, we have already verified that the *continuity* of $\succsim_i^\rho$ is satisfied.[43] The forth axiom, *Constant Equivalence*, is verified by *Continuity* and *Inequity-Averse Mixing* (see Step 1).

In the similar way, we can show that the third axiom *Constant Monotonicity* is satisfied; $\succsim_i^\rho$ is monotone with respect to equal allocations. Take an arbitrary menu $A \in \mathcal{A}$ with $(x_1, x_S), (y_1, x_S) \in A$ where $x_1 > y_1$, $\rho((x_1, x_S), A) > \rho((y_1, x_S), A)$. By definition $(x_1, x_S) \succ_i^\rho (y_1, x_S)$. By *Continuity*, we can find a constant item for each item (x_1, x_S) and (y_1, x_S). Let $(z, \cdots, z)$ be a constant item for (x_1, x_S), and $(z', \cdots, z')$ be a constant item for (y_1, x_S). Then, we have $(z, \cdots, z) \succ_i^\rho (z', \cdots, z')$. If $z < z'$, then $\succsim_i^\rho$ violates *Transitivity*. Hence, $z \geq z'$. Thus, *Constant Monotonicity* is satisfied.

We verify that *Covalent Additivity* is satisfied. By *Quasi-Comonotonic Independence*, if two items $\boldsymbol{x}$ and $\boldsymbol{y}$ are quasi-comonotonic, then the choice probabilities of them do not change. Suppose that two items $\boldsymbol{x}$ and $\boldsymbol{y}$ are quasi-comonotonic, and that $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{y}, A)$. Then, by definition $\boldsymbol{x} \succ_i^\rho \boldsymbol{y}$. We can take $\boldsymbol{z} \in X$ such that both $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{y}, \boldsymbol{z})$ are pair-wise quasi-comonotonic. Then, $\boldsymbol{x} + \boldsymbol{z}$ and $\boldsymbol{y} + \boldsymbol{z}$ are also quasi-comonotonic. That is, $\boldsymbol{x} + \boldsymbol{z}$ and $\boldsymbol{y} + \boldsymbol{z}$ are covalent. By *Quasi-Comonotonic Independence*, $\boldsymbol{x} + \boldsymbol{z} \succ_i^\rho \boldsymbol{y} + \boldsymbol{z}$. The symmetric part is shown in the same way.

Finally, we verify that *Inequity Aversion* is satisfied. We immediately show this property by the axiom of ρ's *Inequity Aversion*. We omit it. □

Utility Representation. It is shown that $\succsim_i^\rho$ associated with ρ satisfies the Rohde (2010)'s axioms. Hence, $\succsim_i^\rho$ is represented by the Fehr and Schmidt (1999)'s inequity-averse utility model, i.e., $u = u_{IA}$ where u_{IA} is a pair $(\boldsymbol{\alpha}, \boldsymbol{\beta})$.[44] By applying this result into Lemma 3, we obtain the desired result: For any menus $A \in \mathcal{A}$,

$$\rho(A) = \arg\max_{\rho \in \Delta(A)} \sum_{\boldsymbol{x} \in A} \Big(u_{IA}(\boldsymbol{x})\rho(\boldsymbol{x}) - c_A(\rho(\boldsymbol{x})) \Big)$$

where $u_{IA}(\boldsymbol{x}) = x_1 - \sum_{i=2}^n \Big(\alpha_i \max\{x_i - x_1, 0\} + \beta_i \max\{x_1 - x_i, 0\} \Big)$. □

Proof of the Necessity Part

Suppose that ρ is represented by a pair $\langle (\boldsymbol{\alpha}, \boldsymbol{\beta}), (c_A)_{A \in \mathcal{A}} \rangle$.

[43] Take an arbitrary menu A^m such that for all $\boldsymbol{x}^m \in A^m$, $\boldsymbol{x}^m \to \boldsymbol{x}$. Take two items $\boldsymbol{x}^m, \boldsymbol{y}^m \in A^m$ with $\rho(\boldsymbol{x}^m, A^m) > \rho(\boldsymbol{y}^m, A^m)$. By definition, $\boldsymbol{x}^m \succ_i^\rho \boldsymbol{y}^m$. By *Continuity*, $\boldsymbol{x} \succ_i^\rho \boldsymbol{y}$ as $\boldsymbol{x}^m \to \boldsymbol{x}$ and $\boldsymbol{y}^m \to \boldsymbol{y}$.

[44] $\boldsymbol{\alpha} = (\alpha_i)_{i \in S}$ where $\alpha_i \geq 0$ for each $i \in S$, and $\boldsymbol{\beta} = (\beta_i)_{i \in S}$ where $\alpha_i \geq 0$ for each $i \in S$ hold. We allow for the heterogeneity of the parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$.

Quasi-Comonotonic Additivity. Take $A \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y} \in A$. Assume $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{y}, A)$. Then, $u_{IA}(\boldsymbol{x}) > u_{IA}(\boldsymbol{y})$. Take $\boldsymbol{z} \in X$ such that $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are *quasi-comonotonic*. Then, $u_{IA}(\boldsymbol{x}+\boldsymbol{z}) = u_{IA}(\boldsymbol{x}) + u_{IA}(\boldsymbol{z})$, and $u_{IA}(\boldsymbol{y}+\boldsymbol{z}) = u_{IA}(\boldsymbol{y}) + u_{IA}(\boldsymbol{z})$. Thus, we have $u_{IA}(\boldsymbol{x}+\boldsymbol{z}) > u_{IA}(\boldsymbol{y}+\boldsymbol{z})$. By definition, $\rho(\boldsymbol{x}+\boldsymbol{z}, A) > \rho(\boldsymbol{y}+\boldsymbol{z}, A)$ for each $A \in \mathcal{A}$ with $\boldsymbol{x}+\boldsymbol{z}, \boldsymbol{y}+\boldsymbol{z} \in A$.

Inequity Aversion. Take $(1,(0)_{-i}), (-1,(0)_{-i}) \in X$ where $i \neq 1$. Since $\alpha_i, \beta_i \geq 0$, $u_{IA}(1,(0)_{-i}) < u_{IA}(\boldsymbol{0})$ and $u_{IA}(-1,(0)_{-i}) < u_{IA}(\boldsymbol{0})$. Then, by definition, for each $A \in \mathcal{A}$ with $(1,(0)_{-i}), (-1,(0)_{-i}) \in A$, $\rho((1,(0)_{-i}), A) < \rho(\boldsymbol{0}, A)$ and $\rho((-1,(0)_{-i}), A) < \rho(\boldsymbol{0}, A)$.

Inequity-Averse Mixing. Take a finite sequence of menus $\{A_i\}_{i=1}^k$. Suppose that the conditions in the axiom of *Inequity-Averse Mixing* are satisfied. By the way of contradiction, suppose $(x, \cdots, x) > \overline{\rho(A_1)}$. By *Monotonicity with respect to Equal Allocations*, $u(x, \cdots, x) = x > u(\overline{\rho(A_1)})$. If $(x, \cdots, x) \in A_k$, then $\rho((x, \cdots, x), A_k) > \rho(\boldsymbol{x}, A_k)$ for all $\boldsymbol{x} \in A_k$, due to the conditions in *Inequity-Averse Mixing*. The inequity-averse utility u_{IA} with $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ states that

$$u_{IA}(\overline{\rho(A_1)}) \geq u_{IA}(\overline{\rho(A_2)}) \geq \cdots \geq u_{IA}(\overline{\rho(A_k)}) = x.$$

Hence, $u_{IA}(\overline{\rho(A_1)}) \geq u_{IA}(\overline{\rho(A_k)}) = x$. This is a contradiction.

Continuity. Take an arbitrary menu $\{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}$ with sequences of items $\lim_{n\to\infty} \boldsymbol{x}_n^i \to \boldsymbol{x}^i$ for each $i = 1, \cdots, m$. Let $\succsim_i^\rho$ be represented by u_{IA} with $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Take $\boldsymbol{x}_n^i, \boldsymbol{x}^j \in \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\}$ with $\rho(\boldsymbol{x}_n^i, \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\}) > \rho(\boldsymbol{x}_n^j, \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\})$. Then, we have $u_{IA}(\boldsymbol{x}_n^i) > u_{IA}(\boldsymbol{x}_n^j)$. Without loss of generality, suppose for each $i = 1, \cdots,$, $\lim_{n\to\infty} u_{IA}(\boldsymbol{x}_n^i) = u_{IA}(\boldsymbol{x}^i)$. By the way of contradiction, suppose that $\rho(\boldsymbol{x}^i, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}) \leq \rho(\boldsymbol{x}^j, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\})$. Since $\sum_{i=1}^m \rho(\boldsymbol{x}^i, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}) = 1$, there exists $\boldsymbol{x}^k \in \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}$ such that for each n, $u_{IA}(\boldsymbol{x}_n^i) < u_{IA}(\boldsymbol{x}_n^k)$, but $\rho(\boldsymbol{x}^i, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}) \geq \rho(\boldsymbol{x}^k, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\})$. This is a contradiction. Hence, if $\rho(\boldsymbol{x}_n^i, \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\}) > \rho(\boldsymbol{x}_n^j, \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\})$, then $\rho(\boldsymbol{x}^i, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\}) > \rho(\boldsymbol{x}^j, \{\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m\})$. The case of equality is similary verified.

A.3 Proof Outline of Theorem 2

We provide the proof outline of Theorem 2 (the sufficiency part). First, in Step 1, we verify the property of *menu-acyclic* conditions (Lemma 5). Then, we show that if ρ satisfies *Continuity*, then *Shame-Averse Acyclicity* implies *Menu Acyclicity* (Lemma 6). We thus show that ρ satisfying *Shame-Averse Acyclicity* is represented by a Menu-Invariant APU (Lemma 7). In Step 2, the binary relation $\succsim_n^\rho$ on X, which is called the *personal norm raking*, is represented by a continuous function

$\varphi : X \to \mathbb{R}$ (Lemma 8). In Step 3, we show that the selfish utility u (in the Menu-Invariant APU) is selfish than the personal norm utility φ (Lemma 9). In Step 4, we complete the desired utility representation, i.e., the personal norm utility φ is embedded into the item-dependent cost functions $(c_{\boldsymbol{x}})_{\boldsymbol{x} \in X}$ (in the Menu-Invariant APU).

A.4 Proof of Theorem 2

Proof of the Sufficiency Part

We show the sufficiency part. Suppose that ρ satisfies the axioms in Theorem 2. In Step 1, we show that if ρ satisfies *Shame-Averse Acyclicity*, then ρ is represented by a Menu-Invariant APU. In Step 2, we characterize an item-dependent cost function. We show that $\succsim_n^\rho$ is represented by φ. In Step 3, we show that a utility function u is more selfish than φ. In Step 4, finally, we obtain the desired utility representation.

Step 1.

In Step 1, we show that ρ satisfying the axioms in Theorem 2 has a *Menu-Invariant APU* (Fudenberg et al., 2014). We mainly use *Self-Interest* and *Shame-Averse Acyclicity*.

Menu Acyclicity. We induce a binary relation $\succsim_m^\rho$ over $\mathcal{A}$ as follows. We say that A is preferred to B, i.e., $A \succ_m^\rho B$ if

$$\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, B)$$

for some $\boldsymbol{x} \in A \cap B$. In the similar way, we say that A is indifferent to B, i.e., $A \sim_m^\rho B$ if

$$\rho(\boldsymbol{x}, A) = \rho(\boldsymbol{x}, B)$$

for some $\boldsymbol{x} \in A \cap B$.

Fix $A, B \in \mathcal{A}$. Let $B_0 = \{\varnothing\}, \cdots, B_m = \{\boldsymbol{x}_n^1, \cdots, \boldsymbol{x}_n^m\}$ for each $n, m \in \mathbb{N}$. Suppose $\boldsymbol{x}_n^m \to \boldsymbol{x}^m$, for each $n, m \in \mathbb{N}$. Then, $B_n^m \to B^m$. Let $B = \cup_{i=1}^m B^m$. Suppose, for all $\boldsymbol{y} \in A$, $\boldsymbol{x}_n^m \geq \boldsymbol{y}$. Let $A \cup B_n^m := (A \cup B_n^{m-1}) \cup \{\boldsymbol{x}_n^m\}$. Then, by *Self-Interest*, $A \cup B_n^m \succsim_m^\rho A \cup B_n^{m-1}$. Notice that for any $\boldsymbol{x}, \boldsymbol{y} \in X \subset \mathbb{R}_+^2$, if $\boldsymbol{x} \geq \boldsymbol{y}$, then $x_1 x_2 \geq y_1 y_2$. Hence, by *Shame-Averse Acyclicity*, for all m, $A \cup B_n^m \succsim_m^\rho A$ holds. By *Continuity*, $A \cup B \succsim_m^\rho A$. Thus, $\rho(\boldsymbol{x}, A \cup B) \geq \rho(\boldsymbol{y}, A)$.

Fudenberg et al. (2015) introduce the following axiom.

Axiom 19. (Menu Acyclicity): If

$$\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_1, A_2), \rho(\boldsymbol{x}_k, A_k) \geq^* \rho(\boldsymbol{x}_k, A_{k+1}) \text{ for } 1 < k < n,$$

then $\rho(\boldsymbol{x}_n, A_n) \not\geq^* \rho(\boldsymbol{x}_n, A_1)$.

Fudenberg et al. (2014) introduce the following axiom.

Axiom 20. (Menu Acyclicity*): There does not exist a sequence of menus $A_1, \cdots, A_n$ such that

$$A_1 \succsim_m^\rho A_2 \succsim_m^\rho \cdots \succsim_n^\rho A_n \succ_m^\rho A_1.$$

We show the following lemma regarding the menu-acyclic condition.

Lemma 5. *If ρ satisfies Menu Acyclicity, then $\succsim_m^\rho$ associated with ρ satisfies Menu Acyclicity*.*

Proof. This is clear by definition. Take a sequence of menus $A_1, \cdots, A_n$. Suppose that ρ satisfies *Menu Acyclicity*.

$$\rho(\boldsymbol{x}_1, A_1) > \rho(\boldsymbol{x}_1, A_2), \rho(\boldsymbol{x}_k, A_k) \geq^* \rho(\boldsymbol{x}_{k+1}, A_{k+1}) \text{ (for } 1 < k < n) \Rightarrow \rho(\boldsymbol{x}_n, A_n) \not\geq^* \rho(\boldsymbol{x}_1, A_n).$$

By definition, $A_1 \succ_m^\rho A_2$, and for $1 < k < n$, $A_k \succsim_m^\rho A_{k+1}$. Hence, there does not exist a cycle such as $A_n \succ_m^\rho A_1$. □

We show that *Shame-Averse Acyclicity* implies *Menu-Acyclicity*.

Lemma 6. *Suppose that ρ satisfies Continuity. If ρ satisfies Shame-Averse Acyclicity, then ρ satisfies Menu Acyclicity.*

Proof. Take a finite sequence of menus $\{A_i\}_{i=1}^n$ such that A_i is *more susceptible to* A_{i+1} ($i = 1, \cdots, n-1$). Take $\boldsymbol{x}, \boldsymbol{y} \in X$ such that $\boldsymbol{x} > \boldsymbol{y}$, and fix them. Take an arbitrary menu $A_1' \in \mathcal{A}$ such that $\boldsymbol{x} \in A$. Without loss of generality, suppose $A_1 \sim_m^\rho A_1'$. Since $\boldsymbol{x} \geq \boldsymbol{y}$ implies $x_1 x_2 \geq y_1 y_2$, we construct a menu A_2' with $\boldsymbol{x}, \boldsymbol{y} \in A_2'$ such that $A_2 \sim_m^\rho A_2'$. By *Continuity*, we can find a finite sequence of menus $\{A_i'\}_{i=1}^n$ that corresponds to the shame-averse menu sequence $\{A_i\}_{i=1}^n$, i.e., $A_i' \sim_m^\rho A_i$ for each i. By *Shame-Averse Acyclicity*, such a sequence $\{A_i'\}_{i=1}^n$ is *menu-acyclic*. □

Fudenberg et al. (2014) show the following result (Lemma 7). We use the result by applying Lemma 5.

Definition 16. We say that ρ has an *menu-invariant* additive perturbed utility (APU) form if there exists a pair $\langle u, (c_x)_{x \in X} \rangle$ where $u : X \to \mathbb{R}$ and, for each $x \in X$, $c_x : [0,1] \to \mathbb{R} \cup \{\infty\}$ is strictly convex and C^1 over $(0,1)$, such that ρ is represented by

$$\rho(A) = \arg \max_{\rho \in \Delta(A)} \sum_{\boldsymbol{x} \in A} \Big(u(\boldsymbol{x}) \rho(\boldsymbol{x}) - c_x(\rho(\boldsymbol{x})) \Big)$$

Lemma 7. *The following statements are equivalent.*

(i) *ρ is represented by menu-invariant APU.*

(ii) *There exists a function $\lambda : \mathcal{A} \to \mathbb{R}$ such that $\lambda(A) > \lambda(B)$ if $A \succ_m^\rho B$, and $\lambda(A) = \lambda(B)$ if $A \sim_m^\rho B$.*

(iii) *ρ satisfies Menu Acyclicity.*

Proof. See Fudenberg et al. (2014)'s Proposition 9. □

Step 2.

In Step 2, we show that the personal norm ranking $\succ_n$ on X is represented by a continuous function $\varphi : X \to \mathbb{R}$.

Lemma 8. *$\succsim_n^\rho$ on X is represented by a continuous function $\varphi : X \to \mathbb{R}$.*

Proof. By Step 1, $\succsim_m^\rho$ is *transitive*. Take an arbitrary menu $A \in \mathcal{A}$. Suppose $\boldsymbol{y} \in X \setminus A$ with $x_1x_2 < y_1y_2$ for some $\boldsymbol{x} \in A$. Then, $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\})$. By definition, $A \succ_m^\rho A \cup \{\boldsymbol{y}\}$. Remember that $\succ_n$ on X is defined as follows (Definition 9). We say that an item $\boldsymbol{y}$ is *weakly normatively better* than an item $\boldsymbol{x}$, i.e., $\boldsymbol{y} \succ_n \boldsymbol{x}$ if $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\})$ for some $A \ni \boldsymbol{x}$ and $\boldsymbol{y} \in X \setminus A$ with $x_1x_2 < y_1y_2$. By definition, since $\succsim_m^\rho$ is *acyclic*, $\succ_n$ is *acyclic*.

Define $\varphi : X \to \mathbb{R}$ such that φ represents $\succ_n$. Take arbitrary two items $\boldsymbol{x}, \boldsymbol{y} \in X$ with $x_1x_2 \leq y_1y_2$. Then,

$$\begin{aligned} y_1y_2 \geq x_1x_2 &\Leftrightarrow \boldsymbol{y} \succsim_n^\rho \boldsymbol{x} \\ &\Rightarrow \varphi(\boldsymbol{y}) \geq \varphi(\boldsymbol{x}) \end{aligned}$$

for some $\varphi : X \to \mathbb{R}$. Take an arbitrary item $\boldsymbol{x}' \in X$ with $x_1'x_2' = x_1x_2$. Then, $y_1y_2 \geq x_1'x_2'$ holds. By definition, $\boldsymbol{y} \succsim_n^\rho \boldsymbol{x}'$. Then, $\varphi(\boldsymbol{y}) \geq \varphi(\boldsymbol{x}')$. Since any items are defined on $\succ_n^\rho$, φ is well-defined.

By *Self-Interest*, take $\boldsymbol{x}, \boldsymbol{y} \in X$ with $\boldsymbol{x} \geq \boldsymbol{y}$. Then, $\boldsymbol{x} \geq \boldsymbol{y}$ implies $x_1x_2 \geq y_1y_2$. Thus, $\boldsymbol{x} \succsim_n \boldsymbol{y} \Rightarrow \varphi(\boldsymbol{x}) \geq \varphi(\boldsymbol{y})$. Hence, φ is *monotone*.

By *Continuity*, we show that φ is *continuous*. Take sequences of the two items $\boldsymbol{x}$ and $\boldsymbol{y}$ with $\boldsymbol{x}^n \to \boldsymbol{x}$ and $\boldsymbol{y}^n \to \boldsymbol{y}$. Without loss of generality, suppose that for each k, $x_1^kx_2^k \geq y_1^ky_2^k$. Then, we have $x_1x_2 \geq y_1y_2$ as $\boldsymbol{x}^k \to \boldsymbol{x}$. Thus, $\boldsymbol{x} \succsim_n \boldsymbol{y} \Rightarrow \varphi(\boldsymbol{x}) \geq \varphi(\boldsymbol{y})$. Since the monomial x_1x_2 for each $\boldsymbol{x} \in X$ is continuous, φ is *continuous*. □

Step 3.

In Step 3, we show that the selfish utility function u is more selfish than the personal norm utility φ

Lemma 9. *u is more selfish than φ. .*

Proof. Remember that φ represents $\succsim_n$ (personal norm ranking). We say that u is *more selfish than* φ if for any $\boldsymbol{x} \in X$ and $\triangle_1$ and $\triangle_2$ such that $(x_1 - \triangle_1, x_2 - \triangle_2)$,

(i) $u(\boldsymbol{x}) = u(x_1 - \triangle_1, x_2 + \triangle_2)$ implies $\varphi(\boldsymbol{x}) \leq \varphi(x_1 - \triangle_1, x_2 + \triangle_2)$

(ii) $u(\boldsymbol{x}) = u(x_1 + \triangle_1, x_2 - \triangle_2)$ implies $\varphi(\boldsymbol{x}) \geq \varphi(x_1 + \triangle_1, x_2 - \triangle_2)$

with strict inequality for at least one pair $(\triangle_1, \triangle_2)$. Here, we show the first condition (i). Remember that, in Step 1, there exists a self-utility $u : X \to \mathbb{R}$. Take $\boldsymbol{x} \in X$ and $\Delta_1, \Delta_2 > 0$ such that $u(\boldsymbol{x}) = u(x_1 - \Delta_1, x_2 + \Delta_2)$. Consider $(x_1 - \Delta_1)(x_2 + \Delta_2) = x_1x_2 + x_1\Delta_2 - x_2\Delta_1 - \Delta_1\Delta_2$ where $\boldsymbol{x} = (x_1, x_2)$.

Case (i). First, consider the case of $x_1\Delta_2 = x_2\Delta_1 + \Delta_1\Delta_2$; that is, for each menu $A \in \mathcal{A}$ with $\boldsymbol{x}, (x_1 - \Delta_1, x_2 + \Delta_2) \in A$, $\rho(\boldsymbol{x}, A) = \rho(x_1 - \Delta_1, x_2 + \Delta_2, A)$. Then, $\boldsymbol{x} \sim_n^\rho (x_1 - \Delta_1, x_2 + \Delta_2) \Leftrightarrow \varphi(\boldsymbol{x}) = \varphi(x_1 - \Delta_1, x_2 + \Delta_2)$. The desired result holds with equality.

Case (ii). Next, consider the case of $x_1\Delta_2 < x_2\Delta_1 + \Delta_1\Delta_2$. Then, $\boldsymbol{x} \succ_n^\rho (x_1 - \Delta_1, x_2 + \Delta_2) \Leftrightarrow \varphi(\boldsymbol{x}) > \varphi(x_1 - \Delta_1, x_2 + \Delta_2)$. Take an arbitrary $\boldsymbol{y} \in X$. Suppose that $y_1y_2 \leq (x_1 - \Delta_1)(x_2 + \Delta_2)$. By *Shame-Averse Acyclicity*, $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\}) < \rho(\boldsymbol{y}, \{\boldsymbol{y}, (x_1 - \Delta_1, x_2 + \Delta_2)\})$. Then, we obtain $u(\boldsymbol{x}) > u(x_1 - \Delta_1, x_2 + \Delta_2)$, which contradicts the assumption of $u(\boldsymbol{x}) = u(x_1 - \Delta_1, x_2 + \Delta_2)$.

Case (iii). Finally, consider the case of $x_1\Delta_2 > x_2\Delta_1 + \Delta_1\Delta_2$. We need to show that there does not exist $\boldsymbol{x} \in X$ such that $u(\boldsymbol{x}) = u(x_1 - \Delta_1, x_2 + \Delta_2)$ implies $\varphi(\boldsymbol{x}) = \varphi(x_1 - \Delta_1, x_2 + \Delta_2)$. If this holds, then we do not observe that $\rho(\boldsymbol{x}, A) = \rho(x_1 - \Delta_1, x_2 + \Delta_2, A)$ for each $A \in \mathcal{A}$. Suppose not. Assume $\varphi(\boldsymbol{x}) \neq \varphi(x_1 - \Delta_1, x_2 + \Delta_2)$. By definition, $x_1x_2 \neq (x_1 - \Delta_1)(x_2 + \Delta_2)$. Without loss of generality, assume that $x_1x_2 < (x_1 - \Delta_1)(x_2 + \Delta_2)$. Then, $\boldsymbol{x} \prec_n^\rho (x_1 - \Delta_1, x_2 + \Delta_2)$. Take an arbitrary menu $A \in \mathcal{A}$ with $\boldsymbol{x} \in A$. By *Shame-Averse Acyclicity*, $\rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{(x_1 - \Delta_1)(x_2 + \Delta_2)\})$. We also observe that $\rho(\boldsymbol{x}, A \cup \{(x_1 - \Delta_1)(x_2 + \Delta_2)\}) = \rho(x_1 - \Delta_1, x_2 + \Delta_2, A \cup \{(x_1 - \Delta_1)(x_2 + \Delta_2)\})$. By Step 1, we obtain a Menu-Invariant APU with $(u, (c_x)_{x \in X})$. Then, the following must hold: $u(\boldsymbol{x}) = u(x_1 - \Delta_1, x_2 + \Delta_2)$ and $\varphi(\boldsymbol{x}) = \varphi(x_1 - \Delta_1, x_2 + \Delta_2)$. This is a contradiction. □

Step 4.

In Step 4, we obtain the desired utility representation.

Take $\boldsymbol{x}, \boldsymbol{y} \in X$ such that $x_1x_2 < y_1y_2$. Assume $\boldsymbol{x} \in A$. Then,

$$\begin{aligned} \boldsymbol{y} \succ_n \boldsymbol{x} &\Leftrightarrow \varphi(\boldsymbol{y}) > \varphi(\boldsymbol{x}) \\ &\Leftrightarrow \rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\}) \\ &\Leftrightarrow A \succ_m^\rho A \cup \{\boldsymbol{y}\}. \end{aligned}$$

By Step 1, we have a Menu-Invariant APU with $(u, (c_x)_{x \in X})$. For each $\boldsymbol{x} \in X$, $c_{\boldsymbol{x}}$ is C^1 over $(0, 1)$. Then, by Lemma 7,

$$u(\boldsymbol{x}) + \lambda(A) > u(\boldsymbol{x}) + \lambda(A \cup \{\boldsymbol{y}\})$$

Without loss of generality, assume that λ takes values in $(0, 1)$. Take an arbitrary item $\boldsymbol{x} \in X$. Let

$$\overline{w}(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \rho(\boldsymbol{x}, A) < 1, \forall A \ni \boldsymbol{x} \\ \min\{\lambda(A) \mid A \in \mathcal{A}, \rho(\boldsymbol{x}, A) = 1\} & \text{otherwise.} \end{cases}$$

$$\underline{w}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \rho(\boldsymbol{x}, A) > 0, \forall A \ni \boldsymbol{x} \\ \max\{\lambda(A) \mid A \in \mathcal{A}, \rho(\boldsymbol{x}, A) = 0\} & \text{otherwise.} \end{cases}$$

Construct a strictly increasing and continuous function $f_{\boldsymbol{x}} : [0,1] \to \mathbb{R}$ for each $\boldsymbol{x} \in X$, such that (i) $f_{\boldsymbol{x}}(0) = \underline{w}(\boldsymbol{x})$, (ii) $f_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A)) = \lambda(A)$ if $\rho(\boldsymbol{x}, A) \in (0,1)$, and (iii) $f_{\boldsymbol{x}}(1) = \overline{w}(\boldsymbol{x})$. For each $\boldsymbol{x} \in X$, define a strictly convex C^1 function $c_{\boldsymbol{x}} : [0,1] \to \mathbb{R}$ by $c_{\boldsymbol{x}}(q) = \int_0^q g_{\boldsymbol{x}}(p)dp$. We have the optimality condition:

$$u(\boldsymbol{x}) - c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A)) + \lambda(A) \begin{cases} \geq 0 & \rho(\boldsymbol{x}, A) = 1 \\ = 0 & \rho(\boldsymbol{x}, A) \in (0,1) \\ \leq 0 & \rho(\boldsymbol{x}, A) = 0. \end{cases}$$

Thus, we have

$$\lambda(A) > \lambda(A \cup \{\boldsymbol{y}\}) \Leftrightarrow u(\boldsymbol{x}) - c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A)) < u(\boldsymbol{x}) - c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\})).$$

Then, we have

$$\begin{aligned} \lambda(A) > \lambda(A \cup \{\boldsymbol{y}\}) &\Leftrightarrow c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A)) > c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\}) \\ &\Leftrightarrow \rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, A \cup \{\boldsymbol{y}\}) \\ &\Leftrightarrow \varphi(\boldsymbol{x}) < \varphi(\boldsymbol{y}). \end{aligned}$$

Thus, we obtain

$$g(\varphi(\boldsymbol{x})) > g(\varphi(\boldsymbol{y})),$$

for some continuous function $g : \mathbb{R} \to \mathbb{R}$. We can show that if $c\boldsymbol{x}$ is convex, it is continuous. Since $c_{\boldsymbol{x}}$ is strictly convex, we obtain

$$g'(\varphi(\boldsymbol{x}))(\rho(\boldsymbol{x}, A)) > g'(\varphi(\boldsymbol{y}))(\rho(\boldsymbol{y}, A)).$$

We characterize the item-dependent, i.e., menu-invariant cost functions $(c_{\boldsymbol{x}})_{\boldsymbol{x} \in X}$. For each menu $\boldsymbol{x} \in X$, let

$$c_{\boldsymbol{x}}(\cdot) := g(\varphi(\boldsymbol{x}))(\cdot).$$

Thus, for each $\boldsymbol{x} \in A \cap B$, suppose $A \succ_m^\rho B$ such that for any $\boldsymbol{y}' \in A$ there exists $\boldsymbol{y} \in B$ with $\boldsymbol{y} \succ_n^\rho \boldsymbol{y}'$. By the additive separability of ρ (Lemma 7),

$$\begin{aligned} \rho(\boldsymbol{x}, A) > \rho(\boldsymbol{x}, B) &\Leftrightarrow u(\boldsymbol{x}) + \lambda(A) > u(\boldsymbol{y}) + \lambda(B) \\ &\Leftrightarrow u(\boldsymbol{x}) - c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, A)) > u(\boldsymbol{x}) - c'_{\boldsymbol{x}}(\rho(\boldsymbol{x}, B)) \\ &\Leftrightarrow u(\boldsymbol{x}) - g'(\varphi(\boldsymbol{x}))(\rho(\boldsymbol{x}, A)) > u(\boldsymbol{x}) - g'(\varphi(\boldsymbol{x}))(\rho(\boldsymbol{x}, B)). \end{aligned}$$

Representation. We obtain the desired result. There exists a pair $\langle u, \varphi, (g_{\boldsymbol{x}})_{\boldsymbol{x}\in X}\rangle$ where $u : X \to \mathbb{R}$, and $\varphi : X \to \mathbb{R}$, and $g_{\boldsymbol{x}} : \mathbb{R}_+ \to \mathbb{R}$ for each $\boldsymbol{x} \in X$, such that

$$\rho(A) = \arg\max_{\rho\in\Delta(A)} \sum_{\boldsymbol{x}\in A} \Big(u(\boldsymbol{x})\rho(\boldsymbol{x}) - c_{\boldsymbol{x}}(\rho(\boldsymbol{x}))\Big)$$

where $c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) = g\big(\varphi(\boldsymbol{x})\big)(\rho(\boldsymbol{x}))$ for each $\boldsymbol{x} \in A$. □

Proof of the Necessity Part

Suppose that ρ is represented by a pair$\langle u, \varphi, (g_{\boldsymbol{x}})_{\boldsymbol{x}\in X}\rangle$.

Continuity. Let u and φ be *continuous* functions, and $g_{\boldsymbol{x}}$ be a *strictly convex* function for each $\boldsymbol{x} \in X$. ρ is represented by

$$\rho_{u,\varphi,g_{\boldsymbol{x}}}(A) = \arg\max_{\rho\in\Delta(A)} \sum_{\boldsymbol{x}\in A} \Big(u(\boldsymbol{x})\rho(\boldsymbol{x}) - c_{\boldsymbol{x}}(\rho(\boldsymbol{x}))\Big)$$

where $c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) = g_{\boldsymbol{x}}\Big(\max_{\boldsymbol{y}\in A} \varphi(\boldsymbol{y}) - \varphi(\boldsymbol{x})\Big)(\rho(\boldsymbol{x}))$ for each $\boldsymbol{x} \in A$. Remember that X is compact. The inverse function $u^{-1}(\cdot)$ is bounded. $u^{-1}(\cdot)$ is a subset of the menu A, i.e., $u^{-1}(\cdot) \subseteq A$, and A is bounded and closed. Since u and $c_{\boldsymbol{x}}(\boldsymbol{x} \in A)$ are continuous, the inverse image of a closed set is closed. Hence, $\rho_{u,\varphi,g_{\boldsymbol{x}}}$ is closed.

Self-Interest. Take $\boldsymbol{x}, \boldsymbol{y} \in X$ with $\boldsymbol{x} \geq \boldsymbol{y}$. Consider the doubleton $\{\boldsymbol{x}, \boldsymbol{y}\}$. By the definition of $c_{\boldsymbol{x}}$, we have $u(\boldsymbol{x})\rho(\boldsymbol{x})c_{\boldsymbol{x}}(\rho(\boldsymbol{x})) \geq u(\boldsymbol{y})\rho(\boldsymbol{y})c_{\boldsymbol{y}}(\rho(\boldsymbol{y}))$ under the maximizer ρ in $\Delta(A)$. Thus, we have $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \geq \rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\})$. Since u is weakly increasing, this holds for any $A \in \mathcal{A}$ with $A \ni \boldsymbol{x}, \boldsymbol{y}$ with $\boldsymbol{x} \geq \boldsymbol{y}$.

Shame-Averse Acyclicity. The tuple $\langle u, \varphi, (g_{\boldsymbol{x}})_{\boldsymbol{x}\in X}\rangle$ has a Menu-Invariant APU representation. Then, the resulting stochastic choice behavior $\rho_{u,\varphi,g_{\boldsymbol{x}}}$ satisfies *Menu Acyclicity*. By the continuity of $\rho_{u,\varphi,g}$, we can find a finite sequence of menus $\{A_i\}(i = 1, \cdots, k)$ that is susceptible to shame. This sequence also satisfies *Menu-Acyclicity*. Hence, $\rho_{u,\varphi,g_{\boldsymbol{x}}}$ satisfies *Shame-Averse Acyclicity*.

B Proof of Propositions

B.1 Proof of Proposition 1

Proof of (i). By the way of contradiction, suppose that $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \neq (\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$.[45] Without loss generality, assume that $\alpha_i < \widehat{\alpha}_i$ for some $i \neq 1$. Then,

$$\begin{aligned} u_{IA}(1, (0)_{-i}) &= -\alpha_i \\ &= u_{IA}(-\alpha_i, \cdots, -\alpha_i) \\ &> -\widehat{\alpha}_i \\ &= u_{IA}(-\widehat{\alpha}_i, \cdots, -\widehat{\alpha}_i). \end{aligned}$$

Moreover, we have $\widehat{u}_{IA}(1, (0)_{-i}) = -\widehat{\alpha}_i = u_{IA}(-\widehat{\alpha}_i, \cdots, -\widehat{\alpha}_i)$. Thus, we have

$$(1, (0)_{-i}) \sim_i^{\rho} (-\alpha_i, \cdots, -\alpha_i) \succ_i^{\rho} (-\widehat{\alpha}_i \cdots, -\widehat{\alpha}_i) \sim_i^{\rho} (1, (0)_{-i}).$$

This is a contradiction. Hence, $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$.

Proof of (ii). Suppose that $(u_{IA}, (c_A)_{A \in \mathcal{A}})$ and $(\widehat{u}_{IA}, (\widehat{c}_A)_{A \in \mathcal{A}})$ represent the same ρ where u_{IA} has a pair $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, and $\widehat{u}_{IA}$ has a pair $(\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$. Note that for any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w} \in X$,

$$\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \geq \rho(\boldsymbol{z}, \{\boldsymbol{z}, \boldsymbol{w}\}) \Leftrightarrow u_{IA}(\boldsymbol{x}) - u_{IA}(\boldsymbol{y}) \geq u_{IA}(\boldsymbol{z}) - u_{IA}(\boldsymbol{w}).$$

By the proof of (i), $u_{IA} = \widehat{u}_{IA} \Leftrightarrow (\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}})$. Thus, $\widehat{u}_{IA}(\boldsymbol{x}) - \widehat{u}_{IA}(\boldsymbol{y}) \geq \widehat{u}_{IA}(\boldsymbol{z}) - \widehat{u}_{IA}(\boldsymbol{w})$.

Take $p, p' \in (0, 1)$. By *Richness*, we have $q < 1 - p, 1 - p'$ with $\rho(\boldsymbol{x}, A) = p, \rho(\boldsymbol{y}, A) = q, \rho(\boldsymbol{x}', A) = p'$, and $\rho(\boldsymbol{y}', A) = q$ for some $A \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}', \boldsymbol{y}' \in A$. Fix $A \in \mathcal{A}$.

By the first order condition (FOC) of APU(IA), we have

$$\begin{aligned} \widehat{c}'_A(p) - \widehat{c}'_A(p') &= \widehat{c}'_A(p) - \widehat{c}'_A(q) + \widehat{c}'_A(q) - \widehat{c}'_A(p') \\ &= \widehat{u}_{IA}(\boldsymbol{x}) - \widehat{u}_{IA}(\boldsymbol{y}) + \widehat{u}_{IA}(\boldsymbol{y}') - \widehat{u}_{IA}(\boldsymbol{x}') \\ &= a\big(u_{IA}(\boldsymbol{x}) - u_{IA}(\boldsymbol{y}) + u_{IA}(\boldsymbol{y}') - u_{IA}(\boldsymbol{x}')\big) \\ &= a\big(c'_A(p) - c'_A(q) + c'_A(q) - c'_A(p')\big) \\ &= a\big(c'_A(p) - c'_A(p')\big). \end{aligned}$$

Let $b_A = \widehat{c}'_A(\frac{1}{2}) - ac'_A(\frac{1}{2})$. Then, for any $p \in (\frac{1}{2}, 1)$, $\widehat{c}'_A(p) - \widehat{c}'_A(\frac{1}{2}) = a(c'_A(p) -$

[45] In the paper, we study the case of $|I| = 2$, but in the proof, we study the general case of $|I| = n$.

$c'_A(\frac{1}{2})$). Hence, $\widehat{c}'_A(p) = ac'_A(p) + b_A$. We have

$$\begin{aligned}\widehat{c}_A(p) - \widehat{c}_A\left(\frac{1}{2}\right) &= \int_{\frac{1}{2}}^{p} \widehat{c}'_A(q)dq \\ &= \int_{\frac{1}{2}}^{p} a(c'_A(q) + \gamma_A)dq \\ &= a\Big(c_A(p) - c_A\left(\frac{1}{2}\right)\Big) + \left(p - \frac{1}{2}\right) b_A.\end{aligned}$$

Thus, $\widehat{c}_A(p) = ac_A(p) + b_A p + c_A$ where $c_A := \widehat{c}_A(\frac{1}{2}) - ac_A(\frac{1}{2}) - \frac{b_A}{2}$. □

B.2 Proof of Proposition 2

Proof of the Sufficiency Part

Suppose that ρ satisfies *Ex-Post Fairness-Seeking*. Take $\boldsymbol{z}, \boldsymbol{z}' \in X$ such that $\boldsymbol{z}$ is more unfair than $\boldsymbol{z}'$. Let $\lambda(A)$ be the Lagrange multiplier at the menu A under APU(IA) (u_{IA}, c_A) where $u_{\text{IA}} = (\alpha, \beta)$.

Suppose, by the way of contradiction, $c''_{A\cup\{\boldsymbol{z}\}} \geq c''_{A\cup\{\boldsymbol{z}'\}}$ does not hold at some point $(0, 1)$. Then, there exists $(\underline{q}, \overline{q})$ such that $c''_{A\cup\{\boldsymbol{z}\}} < c''_{A\cup\{\boldsymbol{z}'\}}$ for all $q \in (\underline{q}, \overline{q})$.

Note that, by *Richness*, the range of u_{IA} is bounded. There exist $\boldsymbol{x}, \boldsymbol{y} \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) > \overline{q}$.

u_{IA} is *continuous*; $\{\rho(\boldsymbol{x}, \{\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})|\boldsymbol{z} \in X\}$ is connected. As $u_{\text{IA}}(\boldsymbol{z}) \to \infty$, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \to 0$. As $u_{\text{IA}}(\boldsymbol{z}) \to -\infty$, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \to \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\})$.

We can take $\boldsymbol{z} \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \in (\underline{q}, \overline{q})$. Likewise, take $\boldsymbol{z}' \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}'\}) = \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})$.

Fix $\boldsymbol{y} \in X$. For each $\boldsymbol{x}', \boldsymbol{z}' \in X$, note that $\lambda(\{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\})$ depends on $(u_{\text{IA}}(\boldsymbol{x}'), u_{\text{IA}}(\boldsymbol{z}'))$. We write it down by

$$\lambda(\{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\}) := f(u_{\text{IA}}(\boldsymbol{x}'), u_{\text{IA}}(\boldsymbol{z}')).$$

Thus, f is continuous and strictly decreasing in each argument. For a strictly decreasing sequence $\varepsilon_k \searrow 0$, we can find a strictly increasing sequence $\varepsilon'_k \nearrow 0$ such that

$$f(u_{\text{IA}}(\boldsymbol{x} + \varepsilon_k), u_{\text{IA}}(\boldsymbol{z} + \varepsilon'_k)) = f(u_{\text{IA}}(\boldsymbol{x}), u_{\text{IA}}(\boldsymbol{z}))$$

for all large enough k. Pick such k so that $c'^{-1}(u_{\text{IA}}(\boldsymbol{x} + \varepsilon_k) + f(u_{\text{IA}}(\boldsymbol{x}), u_{\text{IA}}(\boldsymbol{z})) < \overline{q}$.

By the *connectedness* and *continuity*, take $\boldsymbol{x}', \boldsymbol{z}' \in X$ such that $u_{\text{IA}}(\boldsymbol{x}') = u_{\text{IA}}(\boldsymbol{x}) + \varepsilon_k$ and $u_{\text{IA}}(\boldsymbol{z}') = u_{\text{IA}}(\boldsymbol{z}) + \varepsilon'_k$.

By construction, $\rho(\boldsymbol{x}', \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\}) < \overline{q}$ and $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = \rho(\boldsymbol{y}, \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\})$. This corresponds to the first condition of *Ex-Post Fairness-Seeking*. By the second condition of *Ex-Post Fairness-Seeking*, we have

$$
\begin{aligned}
u_{\mathrm{IA}}(x) > u_{\mathrm{IA}}(z) &\Rightarrow \rho(x, \{x, x'\}) > \frac{1}{2} \\
&\Rightarrow \rho(x, \{x, y, z\}) = \rho(x', \{x', y, z'\}).
\end{aligned}
$$

by the construction of g.

By FOC,

$$c'_{\{x,y,z\}}(\rho(x, \{x, y, z\}) = u_{\mathrm{IA}}(x) + \lambda(\{x, y, z\});$$

and

$$c'_{\{x',y,z'\}}(\rho(x', \{x', y, z'\}) = u_{\mathrm{IA}}(x') + \lambda(\{x', y, z'\});$$

Hence, $c'_{\{x,y,z\}}(\rho(x, \{x, y, z\}) = c'_{\{x',y,z'\}}(\rho(x, \{x, y, z\})$. We also have

$$\int_{\rho(x,\{x,y,z\})}^{\rho(x',\{x',y,z'\})} c''_{\{x,y,z\}}(p)dp = \int_{\rho(x,\{x,y,z\})}^{\rho(x',\{x',y,z'\})} c''_{\{x',y,z'\}}(p)dp.$$

Let $A = \{x, x', y\}$. Then, this contradicts that there exists $(\underline{q}, \overline{q})$ such that $c''_{A\cup\{z\}} < c''_{A\cup\{z'\}}$ for all $q \in (\underline{q}, \overline{q})$, because of $\rho(x, \{x, y, z'\}) = \rho(x, \{x, y, z\})$. □

Proof of the Necessity Part

To prove the necessity part, take $z, z' \in X$ such that z is more unfair than z', and $A \in \mathcal{A}$. Suppose that $c''_{A\cup\{z\}}(\cdot) \geq c''_{A\cup\{z'\}}(\cdot)$. Let $A = \{x, y\}$. By the FOC of APU(IA), since c_A for each $A \in \mathcal{A}$ is convex,

$$u_{IA}(x) - c'_{\{x,y,z\}}(\rho(x)) \geq u_{IA}(x) - c'_{\{x,y,z'\}}(\rho(x)).$$

In Lemma 3, we have

$$\lambda(\{x, y, z\}) \geq \lambda(\{x, y, z'\}).$$

Therefore, we have $\rho(x, \{x, y, z\}) \geq \rho(x, \{x, y, z'\})$. □

B.3 Proof of Proposition 3

Proof of the Sufficiency Part

Suppose that ρ satisfies *Ex-Ante Fairness-Seeking*. Take $A \in \mathcal{A}$ with $x, y \in A$ and fix it. And, take $z, z' \in X$ such that x, z are not *quasi-comonotonic*, and $d(x, z) < d(x, z')$.

Let $\lambda(A)$ be the Lagrange multiplier at the menu A under APU(IA) (u_{IA}, c_A) where $u_{\mathrm{IA}} = (\alpha, \beta)$. Suppose, by the way of contradiction, $c''_{A\cup\{z\}} \geq c''_{A\cup\{z'\}}$ does not hold at some point $(0, 1)$. Then, there exists $(\underline{q}, \overline{q})$ such that $c''_{A\cup\{z\}} < c''_{A\cup\{z'\}}$ for all $q \in (\underline{q}, \overline{q})$.

In the following, we use the same proof strategy in Proposition 2. Note that, by *Richness*, the range of u_{IA} is bounded. There exist $\boldsymbol{x}, \boldsymbol{y} \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) > \bar{q}$.

u_{IA} is *continuous*; $\{\rho(\boldsymbol{x}, \{\{\rho(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})|\boldsymbol{z} \in X\}$ is connected. As $u_{\text{IA}}(\boldsymbol{z}) \to \infty$, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \to 0$. As $u_{\text{IA}}(\boldsymbol{z}) \to -\infty$, $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \to \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\})$.

We can take $\boldsymbol{z} \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) \in (\underline{q}, \bar{q})$. Likewise, take $\boldsymbol{z}' \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}'\}) = \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})$.

Fix $\boldsymbol{y} \in X$. For each $\boldsymbol{x}', \boldsymbol{z}' \in X$, note that $\lambda(\{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\})$ depends on $(u_{\text{IA}}(\boldsymbol{x}'), u_{\text{IA}}(\boldsymbol{z}'))$. We write it down by

$$\lambda(\{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\}) := f(u_{\text{IA}}(\boldsymbol{x}'), u_{\text{IA}}(\boldsymbol{z}')).$$

Thus, f is continuous and strictly decreasing in each argument. For a strictly decreasing sequence $\varepsilon_k \searrow 0$, we can find a strictly increasing sequence $\varepsilon'_k \nearrow 0$ such that

$$f(u_{\text{IA}}(\boldsymbol{x} + \varepsilon_k), u_{\text{IA}}(\boldsymbol{z} + \varepsilon'_k)) = f(u_{\text{IA}}(\boldsymbol{x}), u_{\text{IA}}(\boldsymbol{z}))$$

for all large enough k. Pick such k so that $c'^{-1}(u_{\text{IA}}(\boldsymbol{x} + \varepsilon_k) + f(u_{\text{IA}}(\boldsymbol{x}), u_{\text{IA}}(\boldsymbol{z})) < \bar{q}$.

By the *connectedness* and *continuity*, take $\boldsymbol{x}', \boldsymbol{z}' \in X$ such that $u_{\text{IA}}(\boldsymbol{x}') = u_{\text{IA}}(\boldsymbol{x}) + \varepsilon_k$ and $u_{\text{IA}}(\boldsymbol{z}') = u_{\text{IA}}(\boldsymbol{z}) + \varepsilon'_k$.

By construction, $\rho(\boldsymbol{x}', \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\}) < \bar{q}$ and $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = \rho(\boldsymbol{y}, \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\})$. This corresponds to the first condition of *Ex-Ante Fairness-Seeking*. By the second condition of *Ex-Ante Fairness-Seeking*, we have

$$\begin{aligned} u_{\text{IA}}(\boldsymbol{x}) > u_{\text{IA}}(\boldsymbol{z}) &\Rightarrow \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{x}'\}) > \frac{1}{2} \\ &\Rightarrow \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = \rho(\boldsymbol{x}', \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\}). \end{aligned}$$

by the construction of g.

By FOC,

$$c'_{\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}}(\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = u_{\text{IA}}(\boldsymbol{x}) + \lambda(\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\});$$

and

$$c'_{\{\boldsymbol{x}',\boldsymbol{y},\boldsymbol{z}'\}}(\rho(\boldsymbol{x}', \{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\}) = u_{\text{IA}}(\boldsymbol{x}') + \lambda(\{\boldsymbol{x}', \boldsymbol{y}, \boldsymbol{z}'\});$$

Hence, $c'_{\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}}(\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}) = c'_{\{\boldsymbol{x}',\boldsymbol{y},\boldsymbol{z}'\}}(\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})$. We also have

$$\int_{\rho(\boldsymbol{x},\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\})}^{\rho(\boldsymbol{x}',\{\boldsymbol{x}',\boldsymbol{y},\boldsymbol{z}'\})} c''_{\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\}}(p)dp = \int_{\rho(\boldsymbol{x},\{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\})}^{\rho(\boldsymbol{x}',\{\boldsymbol{x}',\boldsymbol{y},\boldsymbol{z}'\})} c''_{\{\boldsymbol{x}',\boldsymbol{y},\boldsymbol{z}'\}}(p)dp.$$

Let $A = \{\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{y}\}$. Then, this contradicts that there exists $(\underline{q}, \bar{q})$ such that $c''_{A\cup\{\boldsymbol{z}\}} < c''_{A\cup\{\boldsymbol{z}'\}}$ for all $q \in (\underline{q}, \bar{q})$, because of $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}'\}) = \rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\})$. □

Proof of the Necessity Part

The proof is similar to the necessity part in Proposition 2. We omit it.

B.4 Proof of Proposition 4

By *Richness*, for any $\boldsymbol{x} \in X$ and $p \in (0,1)$, there exists $\boldsymbol{y} \in X$ such that $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) = p$. By letting $q = 1 - p$, $\rho(\boldsymbol{y}, \{\boldsymbol{x}, \boldsymbol{y}\}) = q$. By applying *Richness*, $p + q = 1$, we have $\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) = p$.

First, we show that u is *cardinal*. We use the following lemma based on Corollary 1 in Fudenberg et al. (2015) (Debreu, 1958).

Lemma 10. *Suppose that ρ satisfies Shame-Averse Acyclicity, Richness, and Continuity. Then, there exists $u : X \to \mathbb{R}$ such that, for any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w} \in X$,*

$$\rho(\boldsymbol{x}, \{\boldsymbol{x}, \boldsymbol{y}\}) \geq \rho(\boldsymbol{z}, \{\boldsymbol{z}, \boldsymbol{w}\}) \Leftrightarrow u(\boldsymbol{x}) - u(\boldsymbol{y}) \geq u(\boldsymbol{z}) - u(\boldsymbol{w}).$$

Moreover, u is unique up to positive affine transformations.

Next, we show the uniqueness result of *item-dependent* cost functions: For each $\boldsymbol{x} \in X$, there exist $a > 0$, $b_{\boldsymbol{x}}, c_{\boldsymbol{x}} \in \mathbb{R}$ such that $\widehat{c}_{\boldsymbol{x}}(p) = ac_{\boldsymbol{x}}(p) + b_{\boldsymbol{x}}p + c_{\boldsymbol{x}}$ for all $p \in (0,1)$. Fix $\boldsymbol{x} \in X$.

By APU(SA),

$$\widehat{c}_{\boldsymbol{x}}(p) = ac_{\boldsymbol{x}}(p) + b_{\boldsymbol{x}}p + c_{\boldsymbol{x}} = a(g(\varphi(\boldsymbol{x}))(p) + b_{\boldsymbol{x}}p + c_{\boldsymbol{x}}.$$

This holds if $\widehat{\varphi} = a\varphi + b_\varphi$ and $\widehat{g} = ag$ with the same unit $a > 0$ and $b_\varphi \in \mathbb{R}$.

Take $p, p' \in (0,1)$. By *Richness*, we have $q < 1 - p, 1 - p'$ with $\rho(\boldsymbol{x}, A) = p, \rho(\boldsymbol{y}, A) = q, \rho(\boldsymbol{x} + \boldsymbol{\varepsilon}, A') = p'$, and $\rho(\boldsymbol{y}, A') = q$ for some $A, A' \in \mathcal{A}$ with $\boldsymbol{x}, \boldsymbol{y} \in A \cap A'$, and $\boldsymbol{\varepsilon} \in \mathbb{R}^2_+$.

By the first order condition (FOC) of APU(SA), we have

$$\begin{aligned}
\widehat{c}'_{\boldsymbol{x}}(p) - \widehat{c}'_{\boldsymbol{x}+\boldsymbol{\varepsilon}}(p') &= \widehat{c}'_{\boldsymbol{x}}(p) - \widehat{c}'_{\boldsymbol{y}}(q) + \widehat{c}'_{\boldsymbol{y}}(q) - \widehat{c}'_{\boldsymbol{x}+\boldsymbol{\varepsilon}}(p') \\
&= \widehat{u}_{IA}(\boldsymbol{x}) - \widehat{u}_{IA}(\boldsymbol{y}) + \widehat{u}_{IA}(\boldsymbol{y}) - \widehat{u}_{IA}(\boldsymbol{x} + \boldsymbol{\varepsilon}) \\
&= a\big(u_{IA}(\boldsymbol{x}) - u_{IA}(\boldsymbol{y}) + u_{IA}(\boldsymbol{y}) - u_{IA}(\boldsymbol{x} + \boldsymbol{\varepsilon})\big) \\
&= a\big(c'_{\boldsymbol{x}}(p) - c'_{\boldsymbol{y}}(q) + c'_{\boldsymbol{y}}(q) - c'_{\boldsymbol{x}+\boldsymbol{\varepsilon}}(p')\big) \\
&= a\big(c'_{\boldsymbol{x}}(p) - c'_{\boldsymbol{x}+\boldsymbol{\varepsilon}}(p')\big).
\end{aligned}$$

By *Continuity*, as $\boldsymbol{\varepsilon} \to \mathbf{0}$, $u(\boldsymbol{x}) = u(\boldsymbol{x} + \boldsymbol{\varepsilon})$. Without loss of generality, assume that $\boldsymbol{x} \sim^\rho_n \boldsymbol{x} + \boldsymbol{\varepsilon}$. as $\boldsymbol{\varepsilon} \to \mathbf{0}$, $\boldsymbol{x} \sim^\rho_n \boldsymbol{x}$. Then, Remember that g is continuous, and that φ is continuous. Thus, we obtain $c_{\boldsymbol{x}} = c_{\boldsymbol{x}+\boldsymbol{\varepsilon}}$.

Let $b_{\boldsymbol{x}} = \widehat{c}'_{\boldsymbol{x}}(\frac{1}{2}) - ac'_{\boldsymbol{x}}(\frac{1}{2})$. Then, for any $p \in (\frac{1}{2}, 1)$, $\widehat{c}'_{\boldsymbol{x}}(p) - \widehat{c}'_{\boldsymbol{x}}(\frac{1}{2}) = a(c'_{\boldsymbol{x}}(p) - c'_{\boldsymbol{x}}(\frac{1}{2}))$. Hence, $\widehat{c}'_{\boldsymbol{x}}(p) = ac'_{\boldsymbol{x}}(p) + b_{\boldsymbol{x}}$. We have

$$\begin{aligned}
\widehat{c}_{\boldsymbol{x}}(p) - \widehat{c}_{\boldsymbol{x}}\left(\frac{1}{2}\right) &= \int_{\frac{1}{2}}^{p} \widehat{c}'_{\boldsymbol{x}}(q)dq \\
&= \int_{\frac{1}{2}}^{p} a(c'_{\boldsymbol{x}}(q) + b_{\boldsymbol{x}})dq \\
&= a\Big(c_{\boldsymbol{x}}(p) - c_{\boldsymbol{x}}\left(\frac{1}{2}\right)\Big) + \left(p - \frac{1}{2}\right) b_{\boldsymbol{x}}.
\end{aligned}$$

Thus, $\widehat{c}_x(p) = ac_x(p) + b_x p + c_x$ where $c_x := \widehat{c}_x(\frac{1}{2}) - ac_x(\frac{1}{2}) - \frac{b_x}{2}$.

By $\widehat{g} = ag$, we have $\widehat{g} = ag(\varphi)$. □

C Costs of Randomization: Ex-Post Fairness

C.1 Example 3.1: Gini Index

We verify that the cost function of APU(IA) in Example 3.1 satisfies Axiom 8. Suppose that the conditions (i) and (ii) in Axiom 8 holds. By the way of contradiction, assume that $\rho(x, \{x, y, z\}) < \rho(x, \{x, y, z'\})$. Take a menu $A \in \mathcal{A}$. Suppose that $z, z' \in X \setminus A$ (such that z is unfair than z'). Then, by definition, we have

$$\eta(A \cup \{z\}) \geq \eta(A \cup \{z'\}).$$

We have

$$\lambda(A \cup \{z\}) \geq \lambda(A \cup \{z'\}).$$

Therefore, we have $\rho(x, A \cup \{z\}) \geq \rho(x, A \cup \{z'\})$. This is a contradiction. □

C.2 Example 3.2 and 3.3: generalized Entropy

Consider the generalized entropy as follows:

$$\mathcal{I}_\alpha = \frac{1}{\alpha(1-\alpha)} \frac{1}{|I|} \sum_{i \in I} \left(1 - \left(\frac{x_i}{\overline{x}}\right)^\alpha\right),$$

where $\alpha \in (-\infty, +\infty)$ with $\alpha \neq 0, 1$, and $\overline{x}$ is the average payoff of x between agents. Example 3.2 (Theil's Entropy) corresponds to $\alpha = 1$, and Example 3.3 (Atkinson) corresponds to $\alpha \leq 1$ by monotone transformations. In the same way as Example 3.1, we can verify that the cost function of APU(IA) in Example 3.2 and 3.3 satisfies Axiom 8.

D Inequity Aversion and Shame Aversion

In this Appendix, we provide an additional explanation to study the difference between IA and SA.

D.1 Choice Set that Both IA and SA exhibits the same Stochastic Choice Behavior

Attitudes toward Inequity. We examine the choice situation that can distinguish inequity aversion from shame aversion. Let $I = \{1, 2\}$, and we fix it. We can experimentally identify the parameters (α, β) in Fehr and Schmidt (1999). Experimenters can ask the subjects as dictators (decision makers) real numbers α and β such that

(i) $(-\alpha, -\alpha) \sim_i (0, 1)$; and

(ii) $(-\beta, -\beta) \sim_i (0, -1)$ (Figure 6).

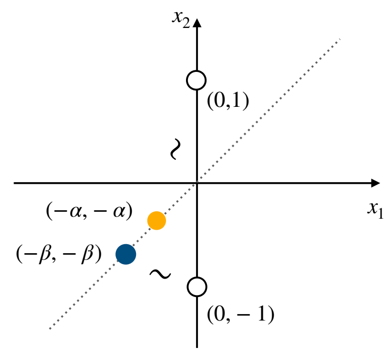


Figure 6: An Example of the Elicitation of Inequity-Averse Preferences

Notice that analysts or experimenters may not know if subjects are inequity-averse or image-conscious, particularly shame-averse. Even if subjects are shame-averse decision makers, they might have fairness concerns a priori.[46]

Thus, we can prepare the items that the decision makers exhibit the same behavior irrespective of IA and SA. For example, in Figure 7, we can find the items $\boldsymbol{x} = (3, 4)$ and $\boldsymbol{y} = (4, 1)$ such that $\rho(\boldsymbol{x}) > \rho(\boldsymbol{y})$ for both IA and SA.

[46] If the decision maker is purely selfish, then $\alpha = \beta = 0$.

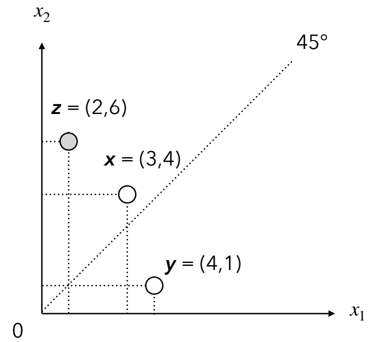


Figure 7: An Example of the Difference between Inequity Aversion and Shame Aversion

References

AGRANOV, M., AND ORTOLEVA, P. (2017): "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, 125(1), 40-68.

ATKINSON, A. B. (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2(3), 244-263.

BROCK, J. M., LANGE, A., AND OZBAY, E. Y. (2013): "Dictating the Risk: Experimental Evidence on Giving in Risky Environments," *American Economic Review*, 103(1), 415-37.

CAMERER, C. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton University Press.

CERREIA-VIOGLIO, S., DILLENBERGER, D., ORTOLEVA, P., AND RIELLA, G. (2019): "Deliberately Stochastic," *American Economic Review*, 109(7), 2425-45.

CHAMBERS, C. P., AND ECHENIQUE, F. (2016): *Revealed Preference Theory* (Vol. 56), Cambridge University Press.

DANA, J., CAIN, D. M., AND DAWES, R. M. (2006): " What you don't know won't hurt me: Costly (but quiet) exit in dictator games," *Organizational Behavior and Human Decision Processes*, 100(2), 193-201.

DEBREU, G. (1958): " Stochastic Choice and Cardinal Utility," *Econometrica*, 26, 440-444.

DELLAVIGNA, S., LIST, J. A., AND MALMENDIER, U. (2012): " Testing for Altruism and Social Pressure in Charitable Giving," *Quarterly Journal of Economics*, 127(1), 1-56.

DILLENBERGER, D., AND SADOWSKI, P. (2012): "Ashamed to be Selfish," *Theoretical Economics*, 7(1), 99-124.

DWENGER, N., KÜBLER, D., AND WEIZSÄCKER, G. (2018): "Flipping a Coin: Evidence from University Applications," *Journal of Public Economics*, 167, 240-250.

FEHR, E., AND SCHMIDT, K. M. (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*: 817-868.

FUDENBERG, D., IIJIMA, R., AND STRZALECKI, T. (2014): "Stochastic Choice and Revealed Perturbed Utility," Working Paper, Harvard University.

FUDENBERG, D., IIJIMA, R., AND STRZALECKI, T. (2015): "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83(6), 2371-2409.

FUDENBERG, D., AND LEVINE, D. K. (2012): "Fairness, Risk Preferences and Independence: Impossibility Theorems," *Journal of Economic Behavior and Organization*, 81(2), 606-612.

FUDENBERG, AND STRZALECKI, T. (2015): "Dynamic Logit with Choice Aversion," *Econometrica*, 83(2), 651-691.

GOEREE, J., HOLT, C., AND PALFREY, T. (2016): *Quantal Response Equilibrium: A Stochastic Theory of Games*, Princeton University Press.

GUL, F., AND PESENDORFER, W. (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403-1435.

GUL, F., AND PESENDORFER, W. (2006): "Random Expected Utility," *Econometrica*, 74(1), 121-146.

HASHIDATE, Y. (2020a): "Social Image Concern and Reference Point Formation," forthcoming in *The B.E. Journal of Theoretical Economics*.

HASHIDATE, Y. (2020b): "Axiomatization of Interdependent Preferences," Working Paper, Waseda University.

HASHIDATE, Y., AND YOSHIHARA, K. (2021): "Stochastic Expected Inequity-Averse Choice," Working Paper, Waseda University.

HOUTHAKKER, H. S. (1950): "Revealed Preference and the Utility Function," *Economica*, 17(66), 159-174.

KREPS, D. M. (1979): "A Representation Theorem for "Preference for Flexibility"," *Econometrica*, 47(3), 565-577.

MACHINA, M. J. (1985): "Stochastic Choice Functions generated from Deterministic Preferences over Lotteries," *Economic Journal*: 95(379), 575-594.

MIAO, B., AND ZHONG, S. (2018): "Probabilistic Social Preference: How Machina's Mom randomizes Her Choice," *Economic Theory*, 65(1), 1-24.

RICHTER, M. K. (1966): "Revealed Preference Theory," *Econometrica: Journal of the Econometric Society*, 34(3), 635-645.

ROHDE, K. I. (2010): "A Preference Foundation for Fehr and Schmidt's Model of Inequity Aversion," *Social Choice and Welfare*, 34(4), 537-547.

SAITO, K. (2013): "Social Preferences under Risk: Equality of Opportunity versus Equality of Outcome," *American Economic Review*, 103(7), 3084-3101.

SAITO, K. (2015a): "Impure Altruism and Impure Selfishness," *Journal of Economic Theory*: 158, 336-370.

SAITO, K. (2015b): "Preferences for Flexibility and Randomization under Uncertainty," *American Economic Review*: 105(3), 1246-1271.

SANDRONI, A., LUDWIG, S., AND KIRCHER, P. (2013): "On the Difference between Social and Private Goods," *The B.E. Journal of Theoretical Economics*: 13(1), 151-177.

THEIL, H. (1967): *Economics and Information Theory*, Elsevier/North-Holland Publishing Co., Amsterdam.