Stochastic Choice and Social Preference

Yosuke Hashidate[1] **Keisuke Yoshihara**[2]

Abstract

This paper studies stochastic choice in social contexts, and explores the difference between inequality-averse preferences and image-conscious preferences. To do so, this paper studies a class of additive perturbed utility (Fudenberg et al., 2015). We show that the general class of inequality-averse preferences correspond to the case of item-invariance, and that image-conscious preferences correspond to the case of menu-invariance. A relationship between our model and regular axioms such as *Regularity* and *Luce's IIA* is discussed.

Keywords: Inequality Aversion, Ex-Ante Fairness, Ex-Post Fairness, Preferences for Randomization, Social Image.

JEL classification: D63, D64, D91.

---

[1] Waseda Business School, Waseda University, yosukehashidate@gmail.com

[2] Graduate School of Economics, the University of Tokyo, ksk0110@gmail.com

1. Introduction

It is widely observed that people choose altruistic behavior or behave prosocially in social contexts. However, it is challenging to explain why people take such behavior. One direct explanation for such a generous behavior is that a decision maker has a *preference for fairness* (Fehr and Schmidt, 1999).[3] Recent experimental evidence such as Dana et al. (2006) has suggested an alternative explanation; that is, the decision maker would be concerned about her *social image*, so that she might feel *social pressure* to behave generously (Dillenberger and Sadowski, 2012).[4]

There is a difficulty in the study of social preferences. It is not easy to identify the motivation or incentive to altruistic/prosocial behavior from observed behavior. Axiomatically, the study of inequality aversion generally takes preferences over allocations (*items*) as primitives. In inequality-averse preferences, the decision maker compares her payoff with other passive agents' payoffs, and the difference between payoffs gives the feeling of *envy* or *guilt*. On the other hand, the study of social image takes preferences over choice sets (*menus*) as primitives. In social image concerns, the decision maker cares about how her choice behavior is perceived by other agents. Such a belief impacts on resulting behavior.

The objective of this paper is to obtain a more comparable property. This paper studies social preferences including inequality aversion and social image concerns, by using stochastic choice functions. This paper provides the axiomatic foundations for both inequality-averse preferences and image-conscious preferences (see the full paper in detail). To do so, we apply the Fudenberg et al. (2015)'s *additive perturbed utility*: stochastic choice corresponds to the maximization of the sum of expected utility and a non-linear perturbation cost function.

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

where $\rho(A)$ is the probability distribution of choice behavior from the choice set A, u is the utility function of the decision maker, and c is a strictly convex perturbation function that rewards the decision maker for deliberately randomizing.

In the study of inequality aversion, we consider the case that cost functions are menu-dependent, i.e., c_A. In inequality-averse preferences, the decision maker may concerns not only the equality of outcomes (Fehr and Schmidt, 1999), but also the equality of opportunities (Fudenberg and Levine, 2012).[5] We present a general class of inequality-averse preferences including both ex-post fairness and ex-ante fairness. Especially, ex-ante fairness is related to menus, so stochastic choice behavior allows for menu-dependence. In the study of social image concerns, we consider the case that cost functions are item-dependent, i.e., c_x. In image-conscious preferences, the decision maker concerns how choice behavior is perceived by others. This is related to which item is chosen from menus. If she chooses a "fair" allocation, then she can be perceived as "fair" person. Then, cost functions are item-dependent, and such an image-conscious preference makes the trade-off between selfishness and altruism, which leads to deliberate randomization.

The key finding is that, in inequality-averse preferences and image-conscious preferences, the incentive to randomization are totally different. In the former case, the decision maker randomizes items to obtain the equality of opportunities. On the other hand, in the latter case, the decision maker randomizes items to conceal her selfish type/image.

---

[3] The decision maker can enjoy improving welfare of other agents.

[4] In the behavioral economics literature, for example, see

[5] Ex-ante fairness is regarded as the *equality of opportunities,* i.e., the equality of equality of ex ante expected payoffs. Ex-ante expected payoffs are obtained by randomizing the items in the menus.

2. Model

We set up the model of *inequality-averse preferences* (Subsetction 2.1) and *social image concerns* (Subsection 2.2). Let $I = \{1,\cdots,n\}$ be the set of individuals. Let 1 be the decision maker, and $S = \{2,\cdots,n\}$be the set of other agents. We assume that the set of payoffs is $\mathbb{R}$. A vector $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ is called an *allocation* of payoffs among individuals, yielding payoff $x_i \in \mathbb{R}$ for each $i \in I$.

Let $X \subset \mathbb{R}^n$ be the set of allocations (*items*). A choice set, i.e., *menu* is a non-empty subset of X. Formally, let $\mathscr{A}$ be the collection of all *finite* non-empty subsets of X.[6] The elements in $\mathscr{A}$ are denoted by $A, B \in \mathscr{A}$.

2.1. Inequality Aversion: Ex-Post Fairness and Ex-Ante Fairness

First, we explain about a model of *ex-post fairness,* i.e., equality of ex post payoffs. Fehr and Schmidt (1999) introduce the seminal model of *preferences for fairness*. Let $\succsim$ on X be a binary relation of the decision maker 1.[7]

There exists a pair (α, β) where $\alpha := (\alpha_i)_{i\in S}$ is a profile with $\alpha_i \geq 0$ for each $i \in S$, and $\beta := (\beta_i)_{i\in S}$ is a profile with $\beta_i \geq 0$ such that $\succsim$ is represented by a function $u_{IA} : X \to \mathbb{R}$ defined by

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

The interpretation of the model is as follows. First, for each $i \in S$, the term $\alpha_i \max\{x_i - x_1,0\}$ captures the disutility of *envy* if $x_1 \leq x_i$, i.e., the decision maker 1's payoff is smaller than the agent i's payoff. Next, for each $i \in S$, the term $\beta_i \max\{x_1 - x_i,0\}$ captures the disutility of *guilt* if $x_1 \geq x_i$, i.e., the decision maker 1's payoff is higher than the agent i's payoff.

Second, we explain about a model of ex-ante fairness (Fudenberg and Levine, 2012). Brock et al. (2013) and Saito (2013) provide the following model of a combination of ex-ante fairness and ex-post fairness. Let $\succsim$ on $\Delta(X)$ where $\Delta(X)$ is the set of all probability distributions over X with finite support. In addition to a pair (α, β), there exists $\gamma \in [0,1]$, and we have the tuple such that $\succsim$ is represented by $U : \Delta(X) \to \mathbb{R}$ defined by

$$U(\rho) = \gamma \sum_{x\in A} u_{IA}(x)\rho(x) + (1-\gamma)u_{IA}\Big(\sum_{x\in A} x\rho(x)\Big).$$

The interpretation of the model is as follows. The new is the second term of the model. This is the randomization on the support of ρ, and the randomized items are evaluated in the way of the Fehr and Schmidt (1999)'s inequality aversion. To understand the term, consider a simple example. For simplicity, assume $n = 2$. Consider a lottery $(0.5;(1,0);0.5,(0,1))$. $(1,0)$ means that the decision maker 1 obtains the payoff 1, and the other passive agent obtains the payoff 0. This lottery means that the decision maker 1 obtains an outcome $(1,0)$ with probability 0.5, and obtains $(0,1)$ with probability 0.5. The first term except γ says that $\frac{1}{2}u_{IA}(1,0) + \frac{1}{2}u_{IA}(0,1)$, and this is a standard evaluation of expected utility theory. On the other hand, the second term except $1-\gamma$ says that $u_{IA}\Big(\frac{1}{2}(1,0) + \frac{1}{2}(0,1)\Big)$, and this captures ex-ante fairness in the sense that such a randomization can lead to $\left(\frac{1}{2}, \frac{1}{2}\right)$.

---

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

[7] See Rhode (2010) for the axiomatization of the Fehr and Schmidt (1999). In the recent development, see Hashidate (2019).

We provide a general class of *inequality-averse preference*. There are two remarkable viewpoint of our model. First, if the decision maker exhibits *ex-ante fairness*, observed choice behavior seems to be stochastic. We present the stochastic choice model with inequality-averse preferences. Second, the stochastic choice model includes both ex-post fairness and ex-ante fairness. Moreover, the model is general in the sense that we allow for menu-dependence on the incentive to randomization. For example, in the model of Brock et al. (2013)/Saito (2013), the parameter γ can be *menu-dependent*.

Definition 1. There exists a pair $\langle(\alpha,\beta),(c_A)_{A\in\mathscr{A}}\rangle$ where $\alpha := (\alpha_i)_{i\in S}$ is a profile with $\alpha_i \geq 0$ for each $i \in S$, and $\beta := (\beta_i)_{i\in S}$ is a profile with $\beta_i \geq 0$ for each $i \in S$, and $(c_A)_{A\in\mathscr{A}}$ is a profile of cost functions for each menu $A \in \mathscr{A}$ such that

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

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

In the model, the cost function depends on choice sets, i.e., *menus*. Intuitively, different menus have different implementation costs. In the menu that includes a "fair" allocation, the inequality-averse decision maker may not have an incentive to randomization. On the other hand, in the menu that does not include "fair" allocations, the inequality-averse decision maker has an incentive to randomization. In other words, the attitude toward randomization depend on menus.

As mentioned above, Brock et al. (2013) study an experimental study on *ex-ante fairness*, and Saito (2013) provides an axiomatic foundation for a combination of ex-ante fairness and ex-post fairness. Their model is a special case of our model.

Example 1. *(Saito, 2013)*

In a general class, by letting $C(\rho) := \sum_{x\in A} c_A(\rho(x))$, we have the functional form

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

To compare our model with Saito (2013), let

$$C(\rho) = \gamma\Big[u_{IA}\Big(\sum_{x\in A} x\rho(x)\Big) - \sum_{x\in A} u_{IA}(x)\rho(x)\Big]$$

with $\gamma \in [0,1]$. By arranging the terms, we obtain the following:

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

Notice that in Brock et al. (2013) and Saito (2013), the parameter γ is *menu-independent*. Our model is general in the sense that we allow for menu-dependence of the parameter γ.

2.2. Social Image: the Case of Shame

Dillenberger and Sadowski (2012) introduce the model of *shame of acting selfishly*, which is the seminal axiomatic model in *social image concerns*.[8] We describe their model (Theorem 1; p. 106) in our setting.

---

[8] There are some related axiomatic studies on social image concerns. Saito (2015) studies a general (image-conscious) utilitarian model, in which the decision maker exhibits not only shame of acting selfishly, but also pride of acting altruistically, and temptation to act selfishly. Hashidate (2019a) generalizes the Saito (2015)'s model in the case that social image concern is *reference-dependent*.

As in Dillenberger and Sadowski (2012), assume that $X = (k, +\infty) \times (k, +\infty)$ where $k \in \mathbb{R} \cup \{-\infty\}$. Let us denote the definition of "more selfish than" between two functions. Let u and φ be real-valued functions on X. For all $x \in X$ and $\triangle_1$ and $\triangle_2$, $(x_1 - \triangle_1, x_2 - \triangle_2)$ holds. We say that u is *more selfish than* φ if for any $x \in X$ and $\triangle_1$ and $\triangle_2$ such that $(x_1 - \triangle_1, x_2 - \triangle_2)$,

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

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

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

Dillenberger and Sadowski (2012) take the framework of preferences over menus, i.e., a binary relation $\succeq$ on $\mathscr{A}$ as a primitive. They elicit a personal norm ranking from $\succeq$ in the following way:

Definition 2. We say that the decision maker 1 is *susceptible to shame* if there exist menus A and B such that $A \succ A \cup B$.

Definition 3. Suppose that the decision maker 1 is susceptible to shame. We say that the decision maker deems y to be *normatively better than* x, i.e., $y \succ_n x$ if there exist $A \in \mathscr{A}$ with $x \in A$ such that $A \succ A \cup \{y\}$.

Definition 4. There exists a pair $\langle u, \varphi, g \rangle$ where $u : X \to \mathbb{R}$, $\varphi : X \to \mathbb{R}$, and $g : \mathbb{R}_+ \to \mathbb{R}$, such that

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

where $c_x(\rho(x)) = g\Big(\max_{y \in A} \varphi(y) - \varphi(x)\Big)(\rho(x))$ for each $x \in A$.

In the model the cost function depends on items, i.e., *allocations*. Intuitively, in the theory of *social image concerns*, the decision maker cares about how her decision-making is perceived by others. What is chosen from menus impacts on social image concerns. In such social contexts, different items, i.e., *allocations* have different implementation costs due to image concerns. This procedural aspect gives the decision maker the incentive to randomization.

In the stochastic choice model of social image, the utility function is generally *selfish*, and is not related to preferences for generosity such as inequality aversion. The desire for randomization stems from the reason that randomization mitigates revealing the decision maker's own type or image. Intuitively, the decision maker may be averse to be perceived by others that he is selfish. Thus, she chooses a deliberate randomization.

3. Comments

3.1. Ex-Ante Fairness and Stochastic Choice

The stochastic choice model of inequality aversion deviates from the axiom of *Regularity*.

Axiom 1. (*Regularity*): ρ satisfies *Regularity* if for any $A, B \in \mathscr{A}$ with $x \in A \subseteq B$, $\rho(x|A) \geq \rho(x|B)$.

Consider an example. Let $B = \{(0.5,0.5), (1,0)\}$. Suppose that the item (0,1) is added into the menu B. Let us denote the menu by $A = \{(0.5,0.5), (1,0), (0,1)\}$. Suppose the decision maker is inequality-averse. Then, in the menu B, the choice probability of (0.5,0.5) can be higher than that of (1,0). For example, first we observe that $\rho((0.5,0.5)|B) = 0.8$ and $\rho((1,0)|B) = 0.2$. Second, in the menu A, there may be a strong incentive to randomization for the mixing between (1,0) and (0,1). The resulting choice behavior is $\rho((0.5,0.5)|A) = 0.5$, $\rho((1,0)|A) = 0.25$, and $\rho((0,1)|A) = 0.25$.

3.2. Shame and Stochastic Choice

The stochastic choice model of social image deviates from the axiom of *Luce's Independence of Irrelevant Alternatives* (Luce's IIA). We state the axiom in the following. The axioms states that the likelihood of choosing an item x relative to y is independent of what other items are available in the menu A.

Axiom 2. (*Luce's IIA*): ρ satisfies *Luce's IIA* if for any $A, B \in \mathscr{A}$ and $x, y \in A \cap B$,

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

Consider an example. Consider three items $x = (1,0)$, $y = (0.8,0.2)$, and $z = (0,1)$. In the menu $A = \{x, y\}$, we observe that $\rho(x\,|\,\{x, y\}) > \rho(y\,|\,\{x, y\})$. However, the item z is added into the menu A. Let $B = \{x, y, z\}$. The item z is normatively better. The cost of choosing x is relatively high. Thus, as a compromise, the choice probability of choosing y can increase, i.e., $\rho(x\,|\,\{x, y, z\}) < \rho(y\,|\,\{x, y, z\})$.

Reference

Andreoni, J. and B. D. Bernheim, 2009. Social image and the 50-50 norm: a theoretical and experimental analysis of audience effects. Econometrica, 77(5), 1607-1636.

Bénabou, R., and J. Tirole, 2006. Incentives and prosocial behavior. American Economic Review, 96(5), 1652-1678.

Brock, J. M., A. Lange,, and E. Y. Ozbay, 2013. Dictating the risk: experimental evidence on giving in risky environments. American Economic Review, 103(1), 415-37.

Dana, J., D. M. Cain,, and R. M.Dawes, 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.

Dillenberger, D., and P. Sadowski, 2012. A shamed to be selfish, Theoretical Economics, 7(1), 99-124.

Fehr, E., and K. M. Schmidt, 1999. A theory of fairness, competition, and cooperation, Quarterly Journal of Economics, 817-868.

Fudenberg, D., R. Iijima,, and T. Strzalecki, 2014. Stochastic choice and revealed perturbed utility, Working Paper, Harvard University.

Fudenberg, D., R. Iijima,, and T. Strzalecki, 2015. Stochastic choice and revealed perturbed utility, Econometrica, 83(6), 2371-2409.

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

Hashidate, Y. 2019a. Social image concern and reference point formation, Working Paper.

Hashidate, Y. 2019b. A general model of inequality aversion," Working Paper.

Kawamura, T., K. Ogawa, and Y. Osaki, 2019. Controlling social image through dictating the risk: experimental evidence," Working Paper.

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, 2015. Impure altruism and impure selfishness, Journal of Economic Theory, 158, 336-370.