**Master's Thesis**

# An Alternative Theory of Stochastic Choice: Intrapersonal Preference Aggregation


**Department of Economics**
**University of Zurich**

Supervisor: Prof. Dr. Jakub Steiner

| | |
|---|---|
| Author: | Chen Lan |
| | chen.lan@uzh.ch |
| | Im Grund 6 |
| | CH - 8600 Dübendorf |
| Matriculation Number: | 22-736-334 |
| Field of Study: | Economics |
| Submission Date: | 24.07.2024 |

**Abstract**

This paper investigates stochastic choice behavior with a novel framework where a decision-maker is modeled as comprising multiple selves with various preferences that are aggregated to determine decision-making. The aggregation process is conceptualized as a probabilistic voting procedure among these selves. We formulate theoretical models based on two specific voting rules—the Plurality rule and the Anti-Plurality rule, referred to as the MS-P Model and the MS-AP Model, respectively. The MS-P Model, paralleling the Random Preference Model, captures stochastic choice behavior of DMs with perfectly rational selves. In contrast, the MS-AP Model incorporates irrational inattention at the self-level, thus accommodating the regularity violation and context effects in the agent-level choice behavior. This result enriches the understanding of how internal conflicting interests and inattentiveness influence individual choice.

***Keywords*:** Multiple Selves, Preference Aggregation, Plurality Rule, Anti-Plurality Rule, Random Attention, Context Effects

**Acknowledgements**

I would like to express my deepest gratitude to my supervisor, Prof. Jakub Steiner, for his support and guidance. I could not have completed this thesis without his invaluable advice and constructive feedback. More than a supervisor, he has been a true mentor, inspiring and encouraging me.

I want to thank Prof. Nick Netzer for motivating the initial idea of this thesis. His courses have greatly shaped my understanding and approach to this work. I am deeply grateful to Prof. Huaxia Zeng, who has been immensely supportive throughout my academic journey. I also want to thank Prof. Shuo Liu, Prof. Sai Luo, and Prof. Qianfeng Tang for their invaluable advice and support.

I am grateful to have had the opportunity to learn from such brilliant mentors, who will be role models for my future path.

Lastly, I want to thank my parents and friends for their love, accompanying me through those up and down moments in life.

# Contents

# 1 Introduction

A large body of empirical evidence indicates that individual choice behavior involves a probabilistic nature, as decision-makers (DMs) are observed to select differently when repeatedly faced with the same set of options in identical choice situations. Moreover, the choice frequencies of selected alternatives often deviate from the equally likely pattern (Tversky, 1969). This stochastic, or random, feature of choice challenges classical theories of deterministic choice, which assume unbounded rationality—DMs are presumed to have unconstrained cognitive capabilities and access to all relevant information, enabling them to make optimal choices that maximize a deterministic, fixed rational preference.

Since the 1950s, numerous theoretical works have been developed to explain and characterize stochastic choice behavior within the framework of bounded rationality (Simon, 1955), which relaxes the assumption of perfect rationality in individuals. The most prominent models can be classified into two broad categories based on the underlying motivation for choice stochasticity. The first category attributes randomness in choice to variability in utility or preference. In this view, the utility or preference of a decision-maker (DM) fluctuates randomly, and choices are made based on the utility or preference level realized at the moment. Examples of this category include the Random Utility Model and the Random Preference Model.[1] In the second category, random choice is ascribed to uncertainty in the choice rule. Here, a DM has a stable utility or preference, but, instead of choosing alternatives yielding the highest utility level or being the most preferred, she follows some sub-optimal rule that brings stochasticity into the decision-making process. Examples

[1]See Barberá and Pattanaik (1986), Becker et al. (1963), and Block and Marschak (1959). In the Random Utility Model, each option in a choice set has a certain level of utility, which is considered a random variable due to various factors like imperfect information, changing tastes, or perceptual mistakes. The probability that a DM chooses a particular option depends on the likelihood that the utility of that option exceeds the utility of any other option. On the other hand, individual preference is viewed as random ordinal rankings over alternatives instead of utility scales in the Random Preference Model. It has been shown that the two models are equivalent in terms of the rationalizability of observed stochastic choice.

include the LUCE Model,[2] the Elimination-by-Aspect Model,[3] and the Random Attention Model.[4]

While existing models typically incorporate only one source of stochasticity, our theory seeks to interpret stochastic choice behavior in a novel way that combines both factors. We propose that a DM comprises multiple selves in mind, each of whom possesses its own preference ordering over available options. This setup allows us to capture choice stochasticity arising from the first source. We further postulate that these multiple selves collectively determine decisions on behalf of the agent as if they were a group of people voting via some voting mechanism in a probabilistic manner. To be specific, each self assigns scores to alternatives according to his preference, and then the scores received by each alternative are aggregated across different selves. These aggregated scores together yield a lottery over alternatives, which is implemented as a random device governing how the DM selects stochastically among all available alternatives.

It is worth noting that if a DM has merely one consistent preference, meaning that there exists a single self, the DM's choice behavior mirrors that of the individual self. This implies that each self assigning scores to feasible alternatives essentially represents an individual with a deterministic preference making random choices among those alternatives. In this sense, intrapersonal preference aggregation models the process of aggregating the *self-level choice behavior* (how each self chooses with a fixed preference) into the *agent-level choice behavior* (how the agent chooses with multiple preferences).[5]

[2]See Luce (1959). The LUCE Model associates subjective weights to all available alternatives and computes the probability of choosing a particular alternative as the proportion of its associated weight relative to the sum of weights of all alternatives. The LUCE Model indicates, though being challenged by empirical evidence, that the relative probability of choosing between any two alternatives in a choice set remains unchanged by including or excluding other alternatives. This principle is known as the Independence from Irrelevant Alternatives.

[3]See Tversky (1972). The Elimination-by-Aspect Model describes the decision-making process as sequentially eliminating alternatives not desirable based on specific attributes or aspects that characterize alternatives until only one remains to be selected. Choice is stochastic because in the eliminating process which aspect being considered in the sequence is random.

[4]See Brady and Rehbeck (2016), Cattaneo et al. (2020), and Manzini and Mariotti (2014). We will introduce the Random Attention Model later in Section 5.

[5]To distinguish between the decision-making entities within the multi-self framework, we shall refer to a DM or an agent using the pronouns "she" while we shall refer to a self or a voter using the pronouns "he" throughout this paper.

The hypothesis of aggregating intrapersonal preferences via a voting rule enables us to understand how internal, divergent interests jointly contribute to the agent's decision. More importantly, it allows us to capture choice stochasticity rooted in the second source, as the voting procedures in our conceptualization link to the self-level choice behavior influenced by inattention and cognitive limitations, which provides insights into the primary cause of stochastic agent-level choice patterns. In this paper, we explore two particular scoring methods —the Plurality rule and the Anti-Plurality rule—and formulate corresponding models to explain stochastic choice, which we term as the *MS-P Model* (multiple selves voting by the Plurality rule) and the *MS-AP Model* (multiple selves voting by the Anti-Plurality rule), respectively.

As it turns out, the MS-P Model can be justified as a model in which a DM consists of multiple selves in mind with various preferences, and her choice is determined by aggregating the interests of all selves, each of whom makes decisions as a perfectly rational person optimizing with full knowledge of his own preference and full attention to consider all options presented in a choice problem. In contrast, in the MS-AP Model, each self has limited and random attention which reflects the psychological phenomenon that bad events are more potent than good ones, resulting in suboptimal choices with equal chances of all alternatives except the least preferred one.

Given the perfect rationality at the self-level in the MS-P Model, it is unsurprising that the MS-P Model is equivalent to the Random Preference Model, serving as a baseline model that generates stochastic choice behavior always satisfying a fundamental property that adding more options should not make any of the existing options more attractive—the so-called Regularity condition—and does not accommodate any context effects. By contrast, the MS-AP Model, involving irrationality in the self-level behavior, extends the scope of rationalizable observed stochastic choice patterns and can justify the violation of the Regularity condition, accounting for choice behavior of DMs who are susceptible to context effects such as the decoy effect and the compromise effect.

Our approach extends existing theories of stochastic choice by modeling the following aspects of bounded rationality of DMs in the decision-making process: (i) various internal tastes: instead of having a unified preference, a DM may consist of multiple selves each with different opinions on the ranking of alternatives; (ii) limited cognitive ability: when facing a choice problem, each self with a deterministic preference may suffer from limited and random attention, leading to suboptimal self-level choice behavior; (iii) preference for randomization: when aggregating the self-level choice behavior to the agent-level, the DM intentionally incorporates stochasticity into her choice by implementing the aggregation outcomes as a lottery generating the probabilities of choosing each alternative, rather than selecting with certainty the option delivering the highest aggregated score. As a consequence, our theory offers innovative interpretations of potential underlying mechanisms of real-world choice behavior, which might be missed by standard theories.

The rest of this paper proceeds as follows. In Section 2, we review related work that provides the theoretical foundation to our framework, outlining how our theory builds upon and diverges from existing literature. In Section 3, we exposit our theoretical models and their behavioral implications for stochastic choice, as well as the representation theorems for rationalization of observed choice data. Section 4 addresses how our models can account for context effects, such as the decoy and compromise effects, by contrasting the predictive capabilities of the two models. In Section 5, we explore inattentiveness in self-level decision-making and its impact on agent-level choice behavior, highlighting the relationship between attention strategies and our proposed voting rules. Finally, in Section 6, we conclude with a discussion of limitations and directions for future research.

# 2 Related Work

The notion of multiple selves presenting contradictory identities within a person, dating back at least to James (1890) in psychology, has been widely discussed by economists to interpret intrapersonal conflict. Schelling (1978, 1984) illuminates this concept by saying that "the individual is modeled as a coherent set of preferences and certain cognitive facilities". He suggests that a person may behave as if there were two types of selves, one straight and the other wayward, competing with each other. He argues that these selves do not simultaneously maximize their collective utility; instead, the farsighted self can strategically manage the myopic self to govern the individual's behavior. Formalizing Schelling's idea, several studies, albeit using different terminologies, model an individual as having dual selves and analyze the game between the long-run, foresighted self and the short-run, impulsive self to explain phenomena of temptation and self-control in the domain of intertemporal choice (Fudenberg & Levine, 2006; Gul & Pesendorfer, 2001; Thaler & Shefrin, 1981).

Motivated by these works, we postulate in our theory that multiple selves, each having his own preference, coexist in an agent's mind. Irrational individual behavior can be caused by such conflicting internal interests. However, our framework differs from the common hypothesis in intertemporal choice theory by assuming that these different selves act simultaneously and collectively to determine the agent's decision-making. Specifically, we conceptualize an agent with multiple selves as a society where the agent's decision-making is obtained by aggregating preferences of selves through a voting procedure.

The idea of interpersonal preference aggregation is certainly not new in the context of social choice. Voting mechanisms are employed to resolve situations where a group of people with conflicting interests needs to reach a collective decision. More precisely, a social choice function associated with a particular voting rule aggregates individual preferences into a social ranking of alternatives, according to which a single alternative is chosen for society. Fishburn (1972), Intriligator (1973), Nitzan (1975), and Zeckhauser (1969) first generalize the analysis of the social choice problem to a probabilistic framework where social decisions are determined with randomness. That is, the involved preferences are aggregated into a lottery that is used to select among the alternatives randomly for society.

Analogously, our theory assumes that intrapersonal preferences are aggregated in a probabilistic manner such that the probabilities of an agent selecting each alternative from a feasible set are formed by aggregating the votes received by that alternative across selves. One appeal of this assumption arises from the advantage of admitting lotteries over alternatives when the preference aggregation procedure is frequently recurred (Brandl et al., 2016; Brandt, 2017). For instance, consider a small group of colleagues repeatedly voting on restaurants for daily lunch. Then, it might be more desirable for the group to randomly pick a restaurant every day through a chance device that takes account of everybody's taste, compared to rigidly dining at the same place for good. For this reason, aggregating intrapersonal preference probabilistically fits our goal of modeling choice probabilities that can be consistently estimated from observed frequencies of alternatives being chosen in repeated choice situations.

Furthermore, this assumption captures the preference for randomization beyond unconscious randomness in choice. In our theory, the DM could select the highest-scored alternative with certainty, but she deliberately randomizes among alternatives. Evidence reveals that individuals might intentionally choose to randomize their choices rather than sticking to a single deterministic strategy (Agranov & Ortoleva, 2017; Ahn & Sarver, 2013; Cerreia-Vioglio et al., 2019; Dwenger et al., 2018). This preference for uncertainty and randomization reflects a desire to have flexibility and avoid commitment in their choices.

# 3 Theoretical Framework

In this section, we present theoretical models to capture stochastic choice behavior exhibited by an agent with multiple selves in mind voting in a probabilistic manner using either the Plurality rule or the Anti-Plurality rule. We offer axiomatic characterizations for models, addressing whether an observed choice pattern can be explained by our framework and how the preference distribution over multiple selves can be inferred from the given choice probabilities for cases where rationalization is possible.

## 3.1 Preliminaries

We use the symbols $\subseteq$, $\subset$, $\cup$, $+$, $-$, and $\times$ to denote inclusion, proper inclusion, union, disjoint union, difference, and Cartesian product of sets, respectively. The indicator function is denoted by $\mathbb{1}\{S\}$ where $\mathbb{1}\{S\} = 1$ if the statement $S$ is true, and $\mathbb{1}\{S\} = 0$ otherwise.

Let $X$ be a nonempty, finite set of mutually exclusive alternatives, referred to as the *alternative space*. We assume throughout this paper that $|X| \geqslant 3$. Elements in $X$ are typically denoted by $x, y, z, \ldots$. Let $\mathscr{B}$ be the collection of all nonempty subsets of $X$, namely, $\mathscr{B} = 2^X - \varnothing$. Each member in $\mathscr{B}$, denoted by $A, B, C, D, \ldots$, contains alternatives available to a DM, acting as a *choice situation* or *choice problem*. Stochastic choice behavior is modeled as choice probabilities of choosing among a feasible set of alternatives.

**Definition 3.1** (Stochastic Choice)**.** Let $X$ be an alternative space and $E = \{\,(x, B) \in X \times \mathscr{B} \mid x \in B\,\}$. A *stochastic choice function* is a mapping $\pi : E \to$

$[0, 1]$ such that $\sum_{x \in B} \pi(x, B) = 1$ for any $B \in \mathscr{B}$. The triple $(X, \mathscr{B}, \pi)$ is said to be a *stochastic choice structure*.

$\pi(x, B)$ specifies the probability of the alternative $x$ being chosen from the choice set $B$. Note that the above definition of the stochastic choice function integrates the case of deterministic choice where a single alternative is chosen with certainty repeatedly in identical choice problems, i.e., $\pi(x, B) \in \{0, 1\}$ for all $(x, B) \in E$.

Let $\mathscr{P} \subseteq X \times X$ be the set of all possible linear orders on $X$ (i.e., all reflexive, transitive, complete, and antisymmetric binary relations on $X$).

**Definition 3.2** (Strict Preference)**.** An element $R \in \mathscr{P}$ is called a *strict (rational) preference* on $X$. For any $R \in \mathscr{P}$, let $\succ_R$ denote the *antisymmetric part* of $R$ such that $x \succ_R y \Leftrightarrow (x, y) \in R \wedge (y, x) \notin R$.

$R$ describes the ordinal ranking of alternatives according to the degree of preference. We assume that alternatives can be ranked completely with no cycles and no ties. The preference-based approach considers choice as outcomes derived from these preference orderings.

## 3.2 Model Formulation

We posit that a DM consists of multiple selves, each self being identified by its preference, and that selves with a particular preference take a constant portion out of all selves in the agent's mind. We refer to a self possessing the strict preference $R$ as an *$R$-type self*. To analyze the decision-making process from the preference-based perspective, we model such an agent by specifying the proportion of $R$-type selves for each conceivable $R \in \mathscr{P}$, which constitutes a *multi-self system* describing the preference distribution inside her mind.

**Definition 3.3** (Multi-Self System)**.** Let $X$ be an alternative space. A *multi-self system*, denoted as $\mu = (\mu(R))_{R \in \mathscr{P}}$, is a real vector[6] satisfying that $\mu(R) \geqslant 0$ for any $R \in \mathscr{P}$ and $\sum_{R \in \mathscr{P}} \mu(R) = 1$.

[6]We extend $\mu$ from a vector of fractions (rational numbers) to a vector of real numbers just for convenience. This simplification is acceptable because, when considering the rationalization of stochastic choice in our theory, the revealed multi-self system consists of only rational numbers, as long as choice probabilities themselves are all rational numbers (as demonstrated in Section 3.4).

Clearly, $\mu(R)$ represents the proportion of selves having the strict preference $R$, which reflects the intensity of $R$ among various preferences inside an agent's mind. Let $\Delta(\mathscr{P})$ be the set of all possible multi-self systems. In the context of social choice, a multi-self system $\mu \in \Delta(\mathscr{P})$ mirrors a society with finitely many voters, where the total number of voters is normalized to one and voters with the strict preference $R$ account for a proportion of $\mu(R)$.

We then formalize how choice probabilities are determined by aggregating preferences of multiple selves through a voting procedure. Consider a positional voting method that designates for any choice set $B$ a *scoring vector* $s = (s^1, s^2, \ldots, s^{|B|})$ with $s^1 \geqslant s^2 \geqslant \ldots \geqslant s^{|B|} \geqslant 0$. According to this rule, a voter casts $s^1$ votes to his top-ranked alternative, $s^2$ votes to his second-ranked alternative, etc. Hereafter, we consider only the normalized scoring vector, which satisfies $\sum_{k=1}^{|B|} s^k = 1$. In other words, the votes an alternative $x$ in $B$ receives from an $R$-type self who ranks $x$ as the $k$th preferred is $s^k$, denoted as $s_R(x, B)$. The overall score that $x$ in $B$ gets from an agent with the multi-self system $\mu$ is calculated as the average sum of votes assigned to $x$ in $B$ by $R$-type selves for all possible $R \in \mathscr{P}$, weighted by the proportion of $R$-type selves within the agent. Subsequently, the probability of $x$ being chosen from $B$ by the agent is determined as the aggregated scores that $x$ has received.

**Definition 3.4** (Multi-Self System Voting Probabilistically)**.** Let $X$ be an alternative space. Suppose that an agent has a multi-self system $\mu \in \Delta(\mathscr{P})$ applying a positional voting method to aggregate various preferences in a probabilistic manner. The *normalized scoring vector* for $B \in \mathscr{B}$ is denoted by $s = (s^1, s^2, \ldots, s^{|B|})$, satisfying that $s^1 \geqslant s^2 \geqslant \ldots \geqslant s^{|B|} \geqslant 0$ and $\sum_{k=1}^{|B|} s^k = 1$. For any $(x, B) \in E$, the *score* assigned by an $R$-type self to $x$ in $B$ is denoted by $s_R(x, B)$, where if $x$ is the $k$-th ranked alternative in $B$, $s_R(x, B) = s^k$. The *generated stochastic choice structure* is $(X, \mathscr{B}, \boldsymbol{\pi}_\mu)$[7] where $\forall\, (x, B) \in E$,

$$\boldsymbol{\pi}_\mu(x, B) = \sum_{R \in \mathscr{P}} \mu(R) s_R(x, B).$$

$\boldsymbol{\pi}_\mu(x, B)$ is the probability of $x$ being chosen from $B$ by such an agent. Note that for any $B \in \mathscr{B}$, $\sum_{x \in B} s_R(x, B) = \sum_{k=1}^{|B|} s_B^k = 1$ since there is no tie among

[7] To differentiate theoretical predictions from empirical observations, bold characters will be used to indicate predicted choice behavior from models, whereas normal fonts will be used for observed choice data.

alternatives according to a strict preference. Thus, $\boldsymbol{\pi}_\mu$ formulated above is indeed a valid stochastic choice function. Moreover, $\boldsymbol{\pi}_\mu$ is invariant under the multiplication of any positive number on the scoring vector, which implies that only the normalized scoring vector is relevant to stochastic choice behavior generated by a multi-self system voting in a probabilistic manner. To clarify the formulation, we offer a graphical representation of how multiple preferences associated with internal selves are aggregated into the DM's choice behavior by a voting method, as shown in figure 1.

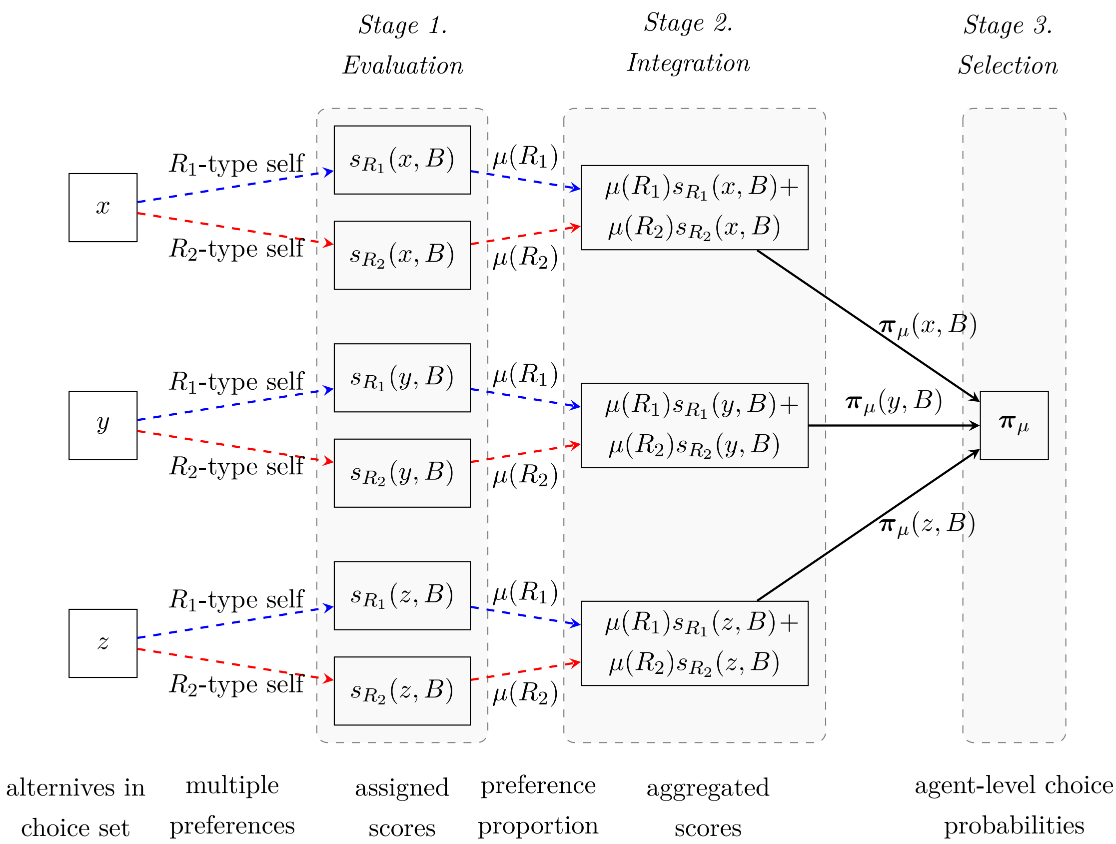


Figure 1: Illustration of intrapersonal preference aggregation in an example where a DM faces a ternary choice set with two distinct preferences in mind.

From now on, we elaborate on two particular positional voting methods—the Plurality rule and the Anti-Plurality rule—which respectively generate the MS-P Model and the MS-AP Model to explain stochastic choice behavior. Note that our framework applies to infinitely many positional voting rules, whereas the two methods proposed equip the models with low computation complexity in addition to their intuitive appeal. For those interested in alternative aggregation methods, we present the model of aggregating multiple preferences using the Borda Count rule and illustrate its main implication in Appendix B.2.

### 3.2.1 MS-P Model

Under the Plurality rule, a voter assigns one vote to his top-ranked alternative and zero to all other alternatives, i.e., the normalized scoring vector is $(1, 0, \ldots, 0)$. Put differently, we denote as $s_R^P(x, B)$ the score assigned by an $R$-type self to an arbitrary alternative $x$ in a choice situation $B$ under the Plurality rule, where

$$s_R^P(x, B) = \mathbb{1}\{x \succ_R y : \forall\, y \in B - \{x\}\}.$$

This means that $x$ receives a score of one if $x$ is the top option according to the $R$-type self's preference and zero otherwise. Formally, we define the MS-P Model exploiting Definition 3.4 as follows.

**Definition 3.5** (MS-P Model)**.** Let $X$ be an alternative space. Suppose that an agent has a multi-self system $\mu \in \Delta(\mathscr{P})$ voting in a probabilistic manner by the Plurality rule. We refer to such $\mu$ as an *MS-P Model*. The *generated stochastic choice function* $\boldsymbol{\pi}_\mu^P$ is determined as $\forall\, (x, B) \in E$,

$$\boldsymbol{\pi}_\mu^P(x, B) = \sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y : \forall\, y \in B - \{x\}\}.$$

We denote by $\Pi^{MS\text{-}P}$ the set of all possible stochastic choice functions generated by some MS-P Model $\mu \in \Delta(\mathscr{P})$.

**Definition 3.6** (MS-P Rationalization)**.** Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. We say that the stochastic choice function $\pi$ can be *rationalized by an MS-P Model* if there exists $\mu \in \Delta(\mathscr{P})$ such that $\pi(x, B) = \boldsymbol{\pi}_\mu^P(x, B)$ for all $(x, B) \in E$, denoted as $\pi \in \Pi^{MS\text{-}P}$.

### 3.2.2 MS-AP Model

Under the Anti-Plurality rule, a voter assigns one vote to all alternatives except his bottom-ranked alternative which gets zero, i.e., the normalized scoring vector is $(\frac{1}{m-1}, \ldots, \frac{1}{m-1}, 0)$ for $m \geqslant 2$ alternatives. Alternatively, we denote as $s_R^{AP}(x, B)$ the score assigned by an $R$-type self to $x$ in a choice problem $B$ under the Anti-Plurality rule, where

$$s_R^{AP}(x, B) = \begin{cases} 1 & \text{, for } |B| = 1, \\ \dfrac{1}{|B|-1} \mathbb{1}\{x \succ_R y : \exists\, y \in B - \{x\}\} & \text{, for } |B| \geqslant 2. \end{cases}$$

This means that for a singleton choice set, the single alternative always receives a score of one. For a choice set with more than one alternative, each alternative except the bottom-ranked one gets an equal share of one vote, which is $\frac{1}{|B|-1}$. Formally, we define the MS-AP Model exploiting Definition 3.4 as follows.

**Definition 3.7** (MS-AP Model)**.** Let $X$ be an alternative space. Suppose that an agent has a multi-self system $\mu \in \Delta(\mathscr{P})$ voting in a probabilistic manner by the Anti-Plurality rule. We refer to such $\mu$ as an *MS-AP Model*. The *generated stochastic choice function* $\boldsymbol{\pi}_\mu^{AP}$ is determined as $\forall\, (x, B) \in E$,

$$\boldsymbol{\pi}_\mu^{AP}(x, B) = \begin{cases} 1 & \text{, for } |B| = 1, \\ \dfrac{1}{|B|-1} \displaystyle\sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y : \exists\, y \in B - \{x\}\} & \text{, for } |B| \geqslant 2. \end{cases}$$

We denote by $\Pi^{MS\text{-}AP}$ the set of all possible stochastic choice functions generated by some MS-AP Model $\mu \in \Delta(\mathscr{P})$.

**Definition 3.8** (MS-AP Rationalization)**.** Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. We say that the stochastic choice function $\pi$ can be *rationalized by an MS-AP Model* if there exists $\mu \in \Delta(\mathscr{P})$ such that $\pi(x, B) = \boldsymbol{\pi}_\mu^{AP}(x, B)$ for all $(x, B) \in E$, denoted as $\pi \in \Pi^{MS\text{-}AP}$.

## 3.3 From Model to Choice

We now provide examples of calculating choice probabilities derived from the MS-P and MS-AP Models, followed by a discussion of their properties (see omitted proofs in Appendix A). Table 1 presents three multi-self systems for a ternary alternative space.[8] The stochastic choice functions generated by the MS-P Model and the

[8] Note that $\mu_1$ and $\mu_2$ correspond to the decoy effect stimulus, while $\mu_3$ aligns with the compromise effect stimulus, as demonstrated in Table 3. Thus, they serve as numerical cases for the discussion in Section 4.

MS-AP Model for each multi-self system are shown in Table 2. For instance, given $\mu_2$, choice probabilities are calculated as follows:

$$\begin{aligned}
\boldsymbol{\pi}^{P}_{\mu_2}(x,\{x,y,z\}) &= 0.2\cdot 1 + 0.6\cdot 1 + 0.2\cdot 0 = 0.8,\\
\boldsymbol{\pi}^{P}_{\mu_2}(y,\{x,y,z\}) &= 0.2\cdot 0 + 0.6\cdot 0 + 0.2\cdot 1 = 0.2,\\
\boldsymbol{\pi}^{P}_{\mu_2}(z,\{x,y,z\}) &= 0.2\cdot 0 + 0.6\cdot 0 + 0.2\cdot 0 = 0,\\
\boldsymbol{\pi}^{AP}_{\mu_2}(x,\{x,y,z\}) &= 0.2\cdot \frac{1}{2} + 0.6\cdot \frac{1}{2} + 0.2\cdot \frac{1}{2} = 0.5,\\
\boldsymbol{\pi}^{AP}_{\mu_2}(y,\{x,y,z\}) &= 0.2\cdot \frac{1}{2} + 0.6\cdot 0 + 0.2\cdot \frac{1}{2} = 0.2,\\
\boldsymbol{\pi}^{AP}_{\mu_2}(z,\{x,y,z\}) &= 0.2\cdot 0 + 0.6\cdot \frac{1}{2} + 0.2\cdot 0 = 0.3.
\end{aligned}$$

Table 1: Examples of multi-self systems based on a ternary alternative space $X = \{x,y,z\}$.

| $\mu_1$ | $R \in \mathscr{P}$ | $\mu_2$ | $R \in \mathscr{P}$ | $\mu_3$ | $R \in \mathscr{P}$ |
|---|---|---|---|---|---|
| | | 0.2 | $x \succ y \succ z$ | | |
| 0.2 | $x \succ z \succ y$ | 0.6 | $x \succ z \succ y$ | 0.8 | $y \succ x \succ z$ |
| 0.8 | $y \succ x \succ z$ | 0.2 | $y \succ x \succ z$ | 0.2 | $z \succ x \succ y$ |

**Proposition 3.1.** *Let $X$ be an alternative space and $\mu \in \Delta(\mathscr{P})$ be a multi-self system. Then, for any $(x,B) \in E$ with $|B| = 2$, $\boldsymbol{\pi}^{P}_{\mu}(x,B) = \boldsymbol{\pi}^{AP}_{\mu}(x,B)$.*

It turns out that the two scoring methods, Plurality and Anti-Plurality, generate identical choice probabilities on binary sets. Intuitively, this is true for any positional voting rule as each method results in the same normalized scoring vector $(1,0)$ on binary sets.

**Condition 3.1** (Regularity)**.** Let $(X,\mathscr{B},\pi)$ be a stochastic choice structure. Then, $\pi$ is said to satisfy the *Regularity* condition[9] if for any $(x,B),(x,B') \in E$ with $B \subseteq B'$, $\pi(x,B) \geqslant \pi(x,B')$.

**Proposition 3.2.** *Let $X$ be an alternative space and $\mu \in \Delta(\mathscr{P})$ be a multi-self system. Then, $\boldsymbol{\pi}^{P}_{\mu}$ satisfies the Regularity condition, while $\boldsymbol{\pi}^{AP}_{\mu}$ may or may not satisfy the Regularity condition.*

[9] Block and Marschak (1959) refer to this condition as “the effect of enlarging the feasible set”.

Table 2: Stochastic choice functions generated by the MS-P and MS-AP Model associated with multi-self systems given in Table 1.

| $(x,B) \in E$ \ $\pi$ | $\mu_1$ | | $\mu_2$ | | $\mu_3$ | |
|---|---|---|---|---|---|---|
| | $\boldsymbol{\pi}^P_{\mu_1}$ | $\boldsymbol{\pi}^{AP}_{\mu_1}$ | $\boldsymbol{\pi}^P_{\mu_2}$ | $\boldsymbol{\pi}^{AP}_{\mu_2}$ | $\boldsymbol{\pi}^P_{\mu_3}$ | $\boldsymbol{\pi}^{AP}_{\mu_3}$ |
| $(x,\{x,y,z\})$ | 0.2 | 0.5 | 0.8 | 0.5 | 0 | 0.5 |
| $(y,\{x,y,z\})$ | 0.8 | 0.4 | 0.2 | 0.2 | 0.8 | 0.4 |
| $(z,\{x,y,z\})$ | 0 | 0.1 | 0 | 0.3 | 0.2 | 0.1 |
| $(x,\{x,y\})$ | 0.2 | 0.2 | 0.8 | 0.8 | 0.2 | 0.2 |
| $(y,\{x,y\})$ | 0.8 | 0.8 | 0.2 | 0.2 | 0.8 | 0.8 |
| $(x,\{x,z\})$ | 1 | 1 | 1 | 1 | 0.8 | 0.8 |
| $(z,\{x,z\})$ | 0 | 0 | 0 | 0 | 0.2 | 0.2 |
| $(y,\{y,z\})$ | 0.8 | 0.8 | 0.4 | 0.4 | 0.8 | 0.8 |
| $(z,\{y,z\})$ | 0.2 | 0.2 | 0.6 | 0.6 | 0.2 | 0.2 |

*Notes*: Choice probabilities on singleton sets are omitted as an alternative is chosen with certainty when it is the only feasible option, irrespective of the composition of the multi-self system or the voting rule in use.

The Regularity condition indicates that enlarging the choice set should not increase the probability of the original alternatives being chosen. Though it sounds natural on the grounds of common sense and underlies several existing theories as well as the MS-P Model, the Regularity condition has been observed to be violated in experiments. Notably, choice probabilities generated by an MS-AP Model do not necessarily satisfy it. For example in Table 2, $\boldsymbol{\pi}^{AP}_{\mu_1}$ is shown to violate the Regularity condition since $\boldsymbol{\pi}^{AP}_{\mu_1}(x,\{x,y\}) = 0.2 < 0.5 = \boldsymbol{\pi}^{AP}_{\mu_1}(x,\{x,y,z\})$.

**Condition 3.2** (Stochastic Quasi-Transitivity over Triplet)**.** Let $(X,\mathscr{B},\pi)$ be a stochastic choice structure. Then, $\pi$ is said to satisfy the *Stochastic Quasi-Transitivity over Triplet* condition[10] if for any distinct $x,y,z \in X$, $\pi(x,\{x,z\}) \leqslant \pi(x,\{x,y\}) + \pi(y,\{y,z\})$.

**Proposition 3.3.** *Let $X$ be an alternative space and $\mu \in \Delta(\mathscr{P})$ be a multi-self system. Then, both $\boldsymbol{\pi}^P_\mu$ and $\boldsymbol{\pi}^{AP}_\mu$ satisfy the Stochastic Quasi-Transitivity over Triplet condition.*

[10]We adopt this notion from the argument for probabilistic judgments by Barberá and Valenciano (1983). It is also mentioned by Block and Marschak (1959) as "triangular".

The Stochastic Quasi-Transitivity over Triplet condition suggests that for any three distinct alternatives $x, y, z$ in $X$, the probability of choosing $x$ from $\{x, z\}$ should not exceed the sum of the probabilities of choosing $x$ from $\{x, y\}$ and $y$ from $\{y, z\}$. This condition serves as a testable criterion for our framework since if $\pi$ does not fulfill this condition, it cannot be explained by the theory.

## 3.4 Axiomatic Characterization

As previously mentioned, the choice probabilities generated by the MS-P Model are equivalent to those generated by the Random Preference Model when associated with the same $\mu \in \Delta(\mathscr{P})$.

**Proposition 3.4.** *Let $X$ be an alternative space and $\mu \in \Delta(\mathscr{P})$. Let $\boldsymbol{\pi}_\mu^{RP}$ be the stochastic choice function generated by the Random Preference Model where the probability distribution of preference is given by $\mu$. Then, for any $(x, B) \in E$, $\boldsymbol{\pi}_\mu^{RP}(x, B) = \sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y : \forall y \in B - \{x\}\} = \boldsymbol{\pi}_\mu^{P}(x, B)$*

The proof follows immediately from the formulations of the Random Preference Model and the MS-P Model. Intuitively, for any $R \in \mathscr{P}$, an $R$-type self in the MS-P Model votes one for the top-ranked alternative and not for others, aligning with the Random Preference Model where an agent chooses the maximal alternative with a probability of one according to the realized preference $R$. Despite that the MS-P Model incorporates multiple preferences acting simultaneously while the Random Preference Model assumes that individual preference randomly fluctuates but only one is realized at the time of choice, this difference between models regarding the underlying mechanism vanishes in the additive computation.

It is already established that a stochastic choice function can be rationalized by a Random Utility Model (or a Random Preference Model) if and only if the so-called Block-Marschak polynomials are all nonnegative, which is also known as Block-Marschak inequalities (Block & Marschak, 1959). Falmagne (1978) elaborates on the necessity and sufficiency of Block-Marschak inequalities based on utility scales for the Random Utility Model, which has been extended to a framework of preference orderings by Barberá and Pattanaik (1986). For theoretical completeness, we restate the representation theorem for the MS-P Model in our framework in Appendix B.1 but omit the detailed proof here.

We now turn to characterize the MS-AP Model. In contrast to the Plurality rule, the Anti-Plurality rule introduces more randomness and inconsistency into choice. Specifically, a degenerated stochastic choice function (i.e., only one alternative is chosen for sure from any feasible set) can never be modeled by an MS-AP Model, as choice probabilities of choosing among any set with $m \geqslant 3$ alternatives cannot exceed $\frac{1}{m-1}$ (which is at most $\frac{1}{2}$) in the MS-AP Model. On the other hand, the Anti-Plurality rule can be viewed as the reverse of the Plurality rule: instead of voting for the top-ranked alternative, each voter vetoes the bottom-ranked alternative. Reflecting this reversal between Plurality and Anti-Plurality, we present a representation theorem for the MS-AP Model (see proof in Appendix A.4).

Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. For any choice set $B \in \mathscr{B}$, let $\mathfrak{S}(B)$ be the set of all permutations of $B$. In other words, each element in $\mathfrak{S}(B)$ is an ordered $|B|$-tuple of all distinct alternatives in B. Let $\mathfrak{S} = \cup_{B\in\mathscr{B}}\mathfrak{S}(B)$. It is worth noting that there exists a one-to-one and onto mapping between $\mathscr{P}$ and $\mathfrak{S}(X)$. We denote this bijection by $\rho : \mathscr{P} \to \mathfrak{S}(X)$ where $\forall\, \sigma = (x_1, \ldots, x_{|X|}) \in \mathfrak{S}(X)$, $\rho^{-1}(\sigma) = \{\, (x_i, x_j) \in X \times X \mid i = 1, \ldots, |X| \text{ and } j = i, \ldots, |X| \,\} \in \mathscr{P}$.[11] That is, each permutation $\sigma \in \mathfrak{S}(X)$ corresponds to a strict preference over $X$ which ranks alternatives as the ordering in the permutation.

We first define from the observed choice behavior $\pi$ a function $\psi_\pi : E \to \mathbb{R}$ where $\forall\, (x, A) \in E$,

$$\psi_\pi(x, A) = \sum_{B\in\mathscr{B}:A\subseteq B} (-1)^{|B|-|A|} \left(1 - (|B| - 1)\pi(x, B)\right). \tag{1}$$

[11] We exploit an example to illustrate the notation. Consider $X = \{x, y, z\}$ and $R = \{(x, x), (x, y), (x, z), (y, y), (y, z), (z, z)\}$. We use $x \succ_R y \succ_R z$ as the shorthand notation for $R$. The permutation counterpart of $R$ can be obtained as $\rho(R) = (x, y, z) \in \mathfrak{S}(X)$. Let $\sigma = (y, z)$ and $\sigma' = (z, y)$. We have that $\mathfrak{S}(\{y, z\}) = \{\sigma, \sigma'\}$. In the following part, we use notation such as $(\sigma, x)$ and $(x, \sigma)$, which in this example refer to the permutations $(y, z, x)$ and $(x, y, z)$, respectively.

We next define from $\psi_\pi$ a function $\gamma_\pi : \mathfrak{S} \to \mathbb{R}$ where

$$\forall\, x \in X,\ \gamma_\pi(x) = \psi_\pi(x, X),$$

$$\forall\,(x,\sigma) \text{ such that } x \in X - D \text{ and } \sigma \in \mathfrak{S}(D) \text{ for some } D \in 2^X - \varnothing - X,$$

$$\gamma_\pi(x,\sigma) = \begin{cases} \dfrac{\psi_\pi(x, X-D)\cdot\gamma_\pi(\sigma)}{\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma')} & , \text{ if } \displaystyle\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') \neq 0, \\ 0 & , \text{ if } \displaystyle\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') = 0, \end{cases} \tag{2}$$

by noticing that $\mathfrak{S} = X \cup \{\,(x,\sigma) \mid \exists\, D \in 2^X - \varnothing - X : x \in X-D, \sigma \in \mathfrak{S}(D)\,\}$.

**Theorem 3.1** (MS-AP Representation)**.** *Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. $\pi \in \Pi^{MS\text{-}AP}$ if and only if $\psi_\pi(x,A) \geqslant 0$, $\forall\,(x,A) \in E$. Furthermore, if $\pi \in \Pi^{MS\text{-}AP}$, then there exists $\mu \in \Delta(\mathscr{P})$ where $\mu(R) = \gamma_\pi(\rho(R))$, $\forall\, R \in \mathscr{P}$ such that $\pi(x,B) = \boldsymbol{\pi}_\mu^{AP}(x,B)$, $\forall\,(x,B) \in E$.*

We discuss the logic of the representation theorem here, starting from the MS-AP Model and then moving to the observed choice data. For any $R \in \mathscr{P}$ and any $x \in X$, let $U_R(x) = \{x\} + \{\, y \in X \mid y \succ_R x \,\}$ denote the upper contour set of $x$ according to $R$. Consider an arbitrary multi-self system $\mu \in \Delta(\mathscr{P})$. We define a function $\boldsymbol{\psi}_\mu : E \to \mathbb{R}$ such that $\forall\,(x,A) \in E$,

$$\boldsymbol{\psi}_\mu(x,A) = \sum_{R \in \mathscr{P}} \mu(R)\mathbb{1}\{A = U_R(x)\}. \tag{3}$$

$\boldsymbol{\psi}_\mu(x,A)$ describes, in the multi-self system, the total proportion of preferences that rank all alternatives outside $A$ below $x$ and rank $x$ as the least preferred alternatives among those in $A$. In fact, $\psi_\pi$, derived from observed choice data, serves as the counterpart to $\boldsymbol{\psi}_\mu$, derived from a multi-self system, in the sense that they are constructed from stochastic choice functions in the same way, except that the former is formed from some observed stochastic choice function while the latter is formulated by some generated stochastic choice function. As a consequence, the rationalization of an observed stochastic choice function $\pi$ by some MS-AP Model $\mu$ boils down to the equivalence relation between $\psi_\pi$ and $\boldsymbol{\psi}_\mu$.

Intuitively, the polynomials $\psi_\pi$ recover information about preference orderings from observed choice data in the same way that $\boldsymbol{\psi}_\mu$ does for a multi-self system. More specifically, $\psi_\pi$ reveals the composition of preference distributions that can explain observed choice data as predicted by an MS-AP Model. In this case, $\psi_\pi(x, A)$ specifies the total proportion of preferences that rank $x$ above all alternatives in $X - A$ and below all alternatives in $A - \{x\}$. Hence, for rationalizability, the polynomials have to be all nonnegative. In addition, one feasible preference distribution can be elicited from $\psi_\pi$ recursively by moving backward from the bottom-ranked alternative to the top-ranked alternative, in the form of conditional probabilities, as defined by $\gamma_\pi$.

# 4 Accounting for Context Effects

While a rich body of theoretical works has been built to model stochastic choice, experimental studies have explored in the meantime the patterns of choice probabilities, particularly concerning how they are influenced by the relation among alternatives in the choice set, namely, the local context. Context-dependent effects of stochastic choice behavior, such as the decoy effect and the compromise effect, examine how the relative probability of selecting between two alternatives in the core set changes with the addition of a third alternative, contingent on the new option's relation to the original ones. These effects are robustly observed across various domains of choice, ranging from consumption products, gambles, and political candidates, to perceptual tasks (Müller et al., 2012; O'Curry & Pitts, 1995; Trueblood et al., 2013), calling for investigation of the underlying mechanisms. Our theory provides an approach to model and explain both the decoy effect and the compromise effect.

For the purpose of illustration, we specify the relation between binary alternatives in terms of the magnitude of their choice probabilities: for a given stochastic choice function $\pi$ and distinct $x, y \in X$, we say that (i) $x$ is *dominated* by $y$ (or $y$ *dominates* $x$) if $\pi(x, \{x, y\}) = 0$;[12] (ii) $x$ is *relatively inferior* to $y$ (or $y$ is *relatively superior* to $x$) if $0 < \pi(x, \{x, y\}) < \pi(y, \{x, y\}) < 1$; (iii) $x$ is *indifferent* to $y$ if $\pi(x, \{x, y\}) = \pi(y, \{x, y\})$. We further refer to the ratio $I_\pi(x, y) = \frac{\pi(y,\{x,y\})}{\pi(x,\{x,y\})}$[13] as the *inferiority* of $x$ against $y$. The more likely $y$ is chosen over $x$ in the binary set, the more intensely $x$ is inferior to $y$, and vice versa. Clearly, (i) $\pi(y, \{x, y\}) = \frac{I_\pi(x,y)}{I_\pi(x,y)+1}$; (ii) $I_\pi(x, y) = 0 \Leftrightarrow x$ dominates $y$, $I_\pi(x, y) = \infty \Leftrightarrow x$ is dominated by $y$,

[12] In contrast to relative inferior relation, we can also say that $x$ is *completely inferior* to $y$, meaning that $x$ is dominated by $y$.

[13] The quotient is well defined on the extended real line with the convention $\frac{1}{0} = \infty$.

and $I_\pi(x,y) = 1 \Leftrightarrow x$ is indifferent to $y$; (iii) $1 < I_\pi(x,y) < \infty \Leftrightarrow x$ is relatively inferior to $y$, and $0 < I_\pi(x,y) < 1 \Leftrightarrow x$ is relatively superior to $y$.

Table 3: An illustration for decoy effect and compromise effect within the framework of multi-self systems over the ternary alternative space $X = \{x, y, z\}$.

(a) Decoy stimulus

| $\mu_d$ | $R \in \mathscr{P}$ |
|---|---|
| $p$ | $x \succ y \succ z$ |
| $q$ | $x \succ z \succ y$ |
| $1-p-q$ | $y \succ x \succ z$ |

$(0 \leqslant p < 1,\ 0 < q < 1,\ 0 < p+q < 1)$

*Notes*: $\mu_d$ demonstrates all conceivable stimuli of the decoy effect in which the asymmetrically dominated option $z$ serves as a decoy of $x$ (i.e., $z$ is dominated by $x$ but no dominance relation occurs between $y$ and $z$), given that neither $x$ dominates $y$ nor $y$ dominates $x$.

(b) Compromise stimulus

| $\mu_c$ | $R \in \mathscr{P}$ |
|---|---|
| $w$ | $y \succ x \succ z$ |
| $1-w$ | $z \succ x \succ y$ |

$(0 < w < 1)$

*Notes*: $\mu_c$ demonstrates all conceivable stimuli of the compromise effect in which $z$ serves as the extreme option that leads $x$ to turn into a compromise (i.e., $x$ becomes a middle option between $y$ and $z$); meanwhile, $z$ does not get into dominance relation either with $x$ or with $y$.

## 4.1 Decoy Effect

The *decoy effect* (Huber et al., 1982), also known as the asymmetric dominance effect or the attraction effect, occurs when adding an asymmetrically dominated alternative (decoy) to the core set raises the relative probability of choosing the target (the alternative dominating the decoy) over the competitor (the alternative neither dominated by nor dominating the decoy), provided that neither of the two in the original binary set is dominated by the other.

Consider any multi-self system consistent with the premise for generating the decoy effect, as presented in the left panel of Table 3 with parameters $p$, $q$ satisfying $0 \leqslant p < 1$, $0 < q < 1$, and $0 < p + q < 1$. Simple algebraic work shows that $\frac{\pi_{\mu_d}(x,\{x,y\})}{\pi_{\mu_d}(y,\{x,y\})} = \frac{p+q}{1-p-q}$ [14] and that

[14] In this section, we omit the superscript (P or AP) indicating voting methods for binary choice probabilities (we do so at no cost because of Proposition 3.1).

(i) it always holds that in the MS-P Model:

$$\frac{p+q}{1-p-q} = \frac{\boldsymbol{\pi}^P_{\mu_d}(x, \{x, y, z\})}{\boldsymbol{\pi}^P_{\mu_d}(y, \{x, y, z\})} = \frac{\boldsymbol{\pi}_{\mu_d}(x, \{x, y\})}{\boldsymbol{\pi}_{\mu_d}(y, \{x, y\})};$$

(ii) it holds that in the MS-AP Model:

$$\frac{1}{1-q} = \frac{\boldsymbol{\pi}^{AP}_{\mu_d}(x, \{x, y, z\})}{\boldsymbol{\pi}^{AP}_{\mu_d}(y, \{x, y, z\})} > \frac{\boldsymbol{\pi}_{\mu_d}(x, \{x, y\})}{\boldsymbol{\pi}_{\mu_d}(y, \{x, y\})}$$

only when $\frac{1-p-q}{p+q} > 1-q$, which equivalent says $I_{\boldsymbol{\pi}_{\mu_d}}(x, y) > \frac{I_{\boldsymbol{\pi}_{\mu_d}}(z,y)}{I_{\boldsymbol{\pi}_{\mu_d}}(z,y)+1}$.

Hence, an MS-P Model is inherently incompatible with the decoy effect. By contrast, whether the decoy effect can be accommodated by an MS-AP Model is contingent on the intensity of the decoy $z$'s inferiority against the competitor $y$, compared with the inferiority of the target $x$ against $y$. Specifically, if at first the target $x$ is indifferent or relatively inferior to the competitor $y$, then the addition of the decoy $z$ will certainly increase the relative probability of $x$ being chosen against $y$ in the ternary set. However, in the circumstance where $x$ is initially relatively superior to $y$, introducing the decoy $z$ might decrease the relative chance of choosing $x$ against $y$ in the ternary set even if $z$ is just relatively inferior to $y$ while completely inferior to $x$, as long as the inferiority of $z$ against $y$ is sufficiently large.

## 4.2 Compromise Effect

The *compromise effect* (Simonson, 1989) arises when the introduction of a third extreme option to the core set increases the relative probability of choosing the compromise alternative (the one becoming an intermediate among the triplet) over the other alternative, provided that the extreme option neither dominates nor is dominated by the original two.

Likewise, we consider all multi-self systems consistent with the premise for generating the compromise effect, as shown in the right panel of Table 3 with the parameter $w$ satisfying $0 < w < 1$. We then have that $\frac{\boldsymbol{\pi}_{\mu_c}(x,\{x,y\})}{\boldsymbol{\pi}_{\mu_c}(y,\{x,y\})} = \frac{1-w}{w}$ and that it always holds that

$$0 = \frac{\boldsymbol{\pi}^P_{\mu_c}(x, \{x, y, z\})}{\boldsymbol{\pi}^P_{\mu_c}(y, \{x, y, z\})} < \frac{\boldsymbol{\pi}_{\mu_c}(x, \{x, y\})}{\boldsymbol{\pi}_{\mu_c}(y, \{x, y\})} \text{ in the MS-P Model, and}$$

$$\frac{1}{w} = \frac{\boldsymbol{\pi}^{AP}_{\mu_c}(x, \{x, y, z\})}{\boldsymbol{\pi}^{AP}_{\mu_c}(y, \{x, y, z\})} > \frac{\boldsymbol{\pi}_{\mu_c}(x, \{x, y\})}{\boldsymbol{\pi}_{\mu_c}(y, \{x, y\})} \text{ in the MS-AP Model.}$$

Therefore, the compromise effect is incompatible with any MS-P Model but is consistently predicted by an MS-AP Model with no exception. That is, in an MS-AP Model, the addition of the extreme option $z$ surely induces a positive compromise effect to the intermediate option $x$, no matter how frequently $x$ is chosen in the original binary set.

# 5 Inattentive Self

Intrapersonal preference aggregation involves combining the choice behavior of each self, who has a deterministic preference, into the overall choice behavior of the agent, who has multiple preferences. As previously noted, if the DM has only one preference in mind, i.e., the multi-self system degenerates to a single self with a deterministic preference, the DM's choice behavior aligns exactly with the choice behavior of that self. This means that each self assigning scores to feasible alternatives essentially corresponds to an individual with a certain preference choosing among these alternatives randomly. Therefore, a normalized voting rule can be seen as reflecting the behavior of a person with a deterministic preference who makes choices stochastically.

Recent works by Brady and Rehbeck (2016), Cattaneo et al. (2020), and Manzini and Mariotti (2014) on limited and random attention have advanced our understanding of the rationale behind the voting rules employed in intrapersonal preference aggregation. By recognizing that selves may be inattentive and not consider all feasible options, the Plurality rule or the Anti-Plurality rule can be interpreted as the results of specific selves' attention strategies, which further sheds light on the behavioral implications such as the regularity violation and context effects discussed in our theory.

We first introduce the Random Attention Model as illustrated by Cattaneo et al. (2020). It captures the idea that individuals do not necessarily consider all options when presented with a set of options. Instead, attention is randomly allocated across subsets of available options according to an attention strategy. The individual then chooses the maximal option from the subset that receives attention. This framework integrates attention strategies into the decision-making process and explains how randomness in attention allocation affects choice probabilities.

**Definition 5.1** (Attention Rule, Cattaneo et al. (2020))**.** Let $X$ be an alternative space and $F = \{\, (A, B) \in \mathscr{B} \times \mathscr{B} \mid A \subseteq B \,\}$. An *attention rule* or *attention strategy* is a function $\lambda : F \to [0, 1]$ such that for any $B \in \mathscr{B}$,

$$\sum_{\substack{A \in \mathscr{B}: \\ A \subseteq B, A \neq \varnothing}} \lambda(A, B) = 1.$$

We interpret $\lambda(A, B)$ as the probability that a DM pays attention to and selects from a potentially smaller set $A \subseteq B$ of alternatives when facing the choice situation $B$.

**Definition 5.2** (Random Attention Model, Cattaneo et al. (2020))**.** Let $X$ be an alternative space. Suppose that an individual has a deterministic strict preference $R \in \Delta(\mathscr{P})$ and the attention rule is specified by $\lambda$. We refer to such $(R, \lambda)$ as a *Random Attention Model*. The *generated stochastic choice function* $\boldsymbol{\pi}_R^{\lambda}$ is determined as $\forall\, (x, B) \in E$,

$$\boldsymbol{\pi}_R^{\lambda}(x, B) = \sum_{\substack{A \in \mathscr{B}: \\ x \in A \subseteq B}} \lambda(A, B) \mathbb{1}\{x \succ_R y : \forall\, y \in A - \{x\}\}.$$

$\boldsymbol{\pi}_R^{\lambda}(x, B)$ describes the probability of $x$ being chosen by an individual who has random attention when presented with the choice set $B$, as an alternative is chosen only when it is the most preferred alternative in the consideration set that receives attention.

## 5.1 Plurality Rule and Perfect Attention

When it comes to attention strategy, the baseline scenario is that a DM has perfect attention so that she precisely selects the most preferred alternative after full consideration of all available alternatives in the choice situation faced by her. Intuitively, this pattern of choice parallels the Plurality rule, where a self with a deterministic preference within a multi-self system always assigns a score of one to the top-ranked alternative and zero to all other alternatives (i.e., choosing the best alternative with certainty) according to his preference.

Formally, suppose that an $R$-type self chooses with perfect, full attention under any feasible set of alternatives. Then the attention rule for any $B \in \mathscr{B}$ is given by

$$\lambda(B, B) = 1 \text{ and } \lambda(A, B) = 0, \forall\, A \subset B \text{ and } A \neq \varnothing,$$

where the whole choice set is the only possible consideration set. The probability of the self choosing $x \in B$[15] is thus determined as

$$\boldsymbol{\pi}_R^{\lambda}(x, B) = \mathbb{1}\{x \succ_R y : \forall\, y \in B - \{x\}\},$$

which is consistent with the scores assigned by an $R$-type self under the Plurality rule.

## 5.2 Anti-Plurality Rule and Random Attention

Under the Anti-Plurality rule, all options but the bottom-ranked one are treated equally, which deviates from the optimal behavior of a fully rational self. As discussed by Cattaneo et al. (2020), such self-level choice patterns can be characterized by the Random Attention Model using specific attention strategies. Notice that a variety of attention rules could justify the self-level choice behavior associated with the Anti-Plurality rule. Here, we provide one motivating example.

We suppose that a self with a deterministic preference, when presented to a choice set, associates weights with every conceivable consideration set and suffers from random attention such that the probability of a possible subset of alternatives being considered is proportional to the associated weight, as proposed by Brady and Rehbeck (2016). Furthermore, we assume for simplicity that a self can only make pairwise comparisons due to limited attention. Formally, we construct a random attention rule as follows[16] : for any $(A, B) \in F$,

$$\lambda(A, B) = \frac{w(A, B)}{\sum_{\substack{A' \in \mathscr{B}: \\ A' \subseteq B, A' \neq \varnothing}} w(A', B)},$$

where the weight of the consideration set $A$ given the current choice set $B$, denoted

[15] We denote the choice behavior of an $R$-type self as $\boldsymbol{\pi}_R^{\lambda}$, since it deduces to the choice behavior $\boldsymbol{\pi}_\mu$ of a degenerated multi-self system $\mu$ with $\mu(R) = 1$.

[16] The constructed attention rule satisfies the monotonic assumption that for any $A \subseteq B \in \mathscr{B}, A \neq \varnothing$, and any $x \in B - A$, it holds that $\lambda(A, B - \{x\}) \geqslant \lambda(A, B)$, which serves as an identifying restriction in the Random Attention Model.

as $w(A, B)$, is determined by the number of alternatives in the choice situation that are worse than some alternatives in the subset, if $A$ contains only two options; otherwise, $w(A, B)$ is zero. That is,

$$w(A, B) = \begin{cases} 0 & \text{, for } |A| \neq 2, \\ \dfrac{1}{|\{\, y \in B \mid \exists x \in A \ s.t.\ x \succ_R y \,\}|} & \text{, for } |A| = 2. \end{cases}$$

In particular, the weights given to each binary consideration set are dependent on the present choice problem and endogenous to individual preference, representing the psychological strengths that a self associates with consideration sets. The more low-ranked options a consideration set contains according to a particular preference ordering, the higher the strength associated with the consideration set is, which reflects a general psychological bias suggesting that negative events have a greater impact than positive ones and people tend to give greater weights to negative events, known as "bad is stronger than good" (Baumeister et al., 2001; Rozin & Royzman, 2001).

It turns out that if $x$ is the bottom-ranked alternative in $B$, then $\boldsymbol{\pi}_R^\lambda(x, B) = 0$; otherwise,

$$\begin{aligned} \boldsymbol{\pi}_R^\lambda(x, B) &= \sum_{y \in B : x \succ_R y} \lambda(\{x, y\}, B) \\ &= \frac{\sum_{y \in B : x \succ_R y} w(\{x, y\}, B)}{\sum_{\substack{A' \in \mathscr{B}: \\ A' \subseteq B, A' \neq \varnothing}} w(A', B)} \\ &= \frac{\sum_{y \in B : x \succ_R y} w(\{x, y\}, B)}{\sum_{\{x', y'\} \subseteq B} w(\{x', y'\}, B)} \\ &= \frac{1}{(|B| - 1) \cdot \frac{1}{(|B|-1)} + (|B| - 2) \cdot \frac{1}{(|B|-1)-1} + \ldots + 1 \cdot \frac{1}{2-1}} \\ &= \frac{1}{|B| - 1}. \end{aligned}$$

As indicated, the choice probabilities of an individual selecting with the constructed random attention strategy align with the scores assigned by a self with the same preference under the Anti-Plurality rule.

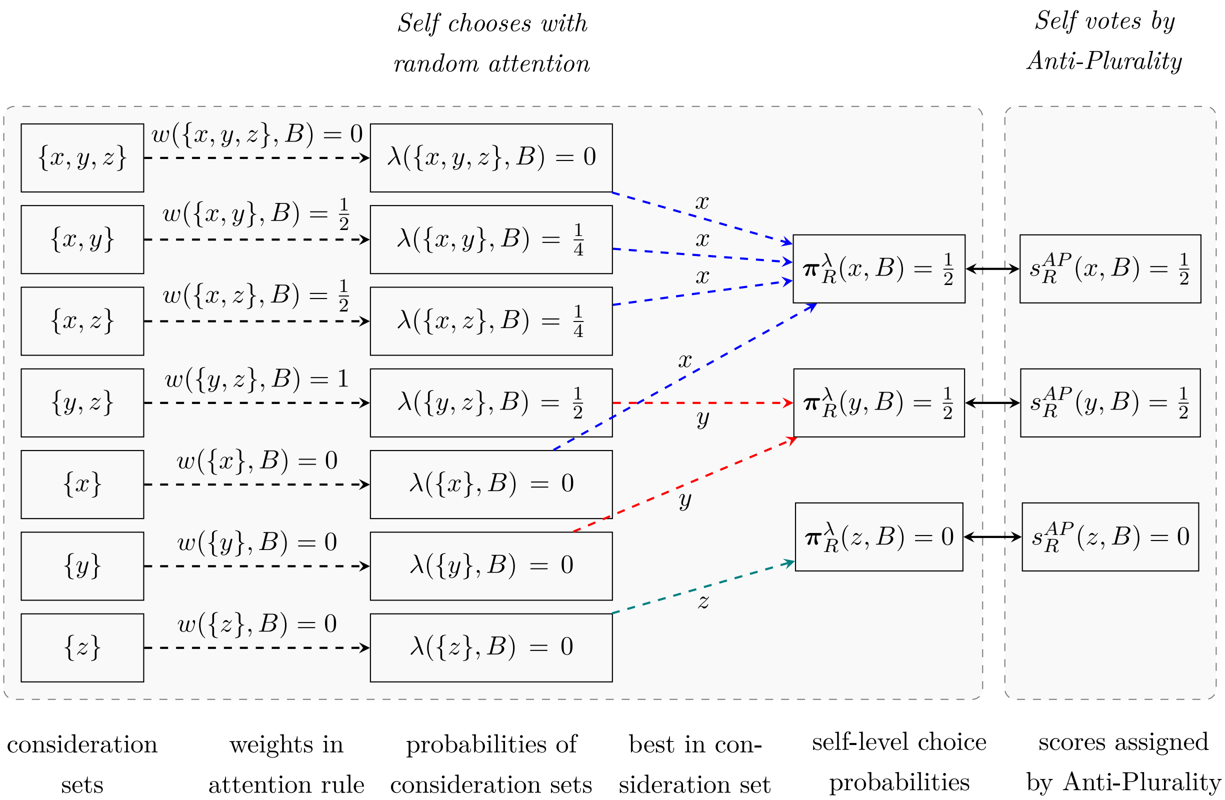


Figure 2: Affinity between random attention and Anti-Plurality rule in an example where a self faces a ternary choice set $B = \{x, y, z\}$ with a deterministic preference $x \succ y \succ z$ and follows the random attention rule as constructed.

By linking the attentiveness of selves with voting rules, we propose that observed inconsistencies, such as regularity violations in real-world decision-making, may originate from the self-level irrationalities inherent in the intrapersonal preference aggregation process.

# 6 Conclusion and Future Work

In this paper, we introduce an innovative theoretical framework that accounts for stochastic choice behavior by modeling decision-makers as having multiple coexisting selves, each with its own preference. We argue that such an agent's choice behavior emerges from aggregating the choice behavior of these selves in a probabilistic manner, which can be conceptualized as a voting procedure. We present two specific models: the MS-P Model, which utilizes the Plurality rule, and the MS-AP Model, which employs the Anti-Plurality rule. We further explore how inattentive selves, modeled through random attention strategies, can account for the self-level choice behavior associated with these rules.

The MS-P Model is shown to be equivalent to the Random Preference Model and is consistent with choice behavior adhering to the Regulation condition, which is unsurprising from the perfect rationality exhibited in the self-level choice behavior. On the other hand, the MS-AP Model, incorporating irrationality due to random attention, effectively explains violations of the Regularity condition and accounts for context effects such as the decoy and compromise effects, offering a richer explanation of observed choice behavior. This work provides insights into how internal divergences and attentional constraints can impact choice outcomes.

Despite the contributions, this paper has several limitations and evokes future research. First, our framework is based on strict preference orderings. Extending the theory to accommodate weak preferences could enhance its applicability to a broader range of decision-making scenarios. Second, the current characterization theorems depend on specific voting rules. Future research could aim to uncover how intrapersonal preferences are aggregated in practice and explore alternative voting mechanisms beyond those considered here. Third, eye movement and response time data in experimental studies could provide empirical validation and refinement of our models. Last but not least, the theory proposed is heuristic and does not account for the real brain mechanisms underlying decision-making. Future research could combine this framework with neurobiological findings to provide further explanations for the cognitive processes driving stochastic choice behavior.

# References


Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, *125*(1), 40–68.

Ahn, D. S., & Sarver, T. (2013). Preference for flexibility and random choice. *Econometrica*, *81*(1), 341–361.

Barberá, S., & Pattanaik, P. K. (1986). Falmagne and the rationalizability of stochastic choices in terms of random orderings. *Econometrica: Journal of the Econometric Society*, 707–715.

Barberá, S., & Valenciano, F. (1983). Collective probabilistic judgements. *Econometrica: Journal of the Econometric Society*, 1033–1046.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, *5*(4), 323–370.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral science*, *8*(1), 41–55.

Block, H. D., & Marschak, J. (1959). Random orderings and stochastic theories of response. *Cowles Foundation Discussion Papers*, *289*(66).

Brady, R. L., & Rehbeck, J. (2016). Menu-dependent stochastic feasibility. *Econometrica*, *84*(3), 1203–1223.

Brandl, F., Brandt, F., & Seedig, H. G. (2016). Consistent probabilistic social choice. *Econometrica*, *84*(5), 1839–1880.

Brandt, F. (2017). Rolling the dice: Recent results in probabilistic social choice. In U. Endriss (Ed.), *Trends in computational social choice* (pp. 3–26). AI Access.

Cattaneo, M. D., Ma, X., Masatlioglu, Y., & Suleymanov, E. (2020). A random attention model. *Journal of Political Economy*, *128*(7), 2796–2836.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., & Riella, G. (2019). Deliberately stochastic. *American Economic Review*, *109*(7), 2425–2445.

Cioabă, S. M., & Murty, M. R. (2009). Möbius inversion and graph colouring. In *A first course in graph theory and combinatorics* (pp. 52–71). New Delhi, India: Hindustan Book Agency.

Dwenger, N., Kübler, D., & Weizsäcker, G. (2018). Flipping a coin: Evidence from university applications. *Journal of Public Economics*, *167*, 240–250.

Falmagne, J.-C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, *18*(1), 52–72.

Fishburn, P. C. (1972). Lotteries and social choices. *Journal of Economic Theory*, *5*(2), 189–207.

Fudenberg, D., & Levine, D. K. (2006). A dual–self model of impulse control. *American economic review*, *96*(5), 1449–1476.

Gul, F., & Pesendorfer, W. (2001). Temptation and self-control. *Econometrica*, *69*(6), 1403–1435.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, *9*(1), 90–98.

Intriligator, M. D. (1973). A probabilistic model of social choice. *The Review of Economic Studies*, *40*(4), 553–560.

James, W. (1890). The consciousness of self. In *The principles of psychology* (pp. 183–197, Vol. 1). New York: Henry Holt And Company.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

Manzini, P., & Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, *82*(3), 1153–1176.

Müller, H., Kroll, E. B., & Vogt, B. (2012). Do real payments really matter? a re-examination of the compromise effect in hypothetical and binding choice settings. *Marketing Letters*, *23*, 73–92.

Nitzan, S. (1975). Social preference ordering in a probabilistic voting model. *Public Choice*, *24*(1), 93–100.

O'Curry, Y. P. S., & Pitts, R. (1995). The attraction effect and political choice in two elections. *Journal of Consumer Psychology*, *4*(1), 85–101.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and social psychology review*, *5*(4), 296–320.

Schelling, T. C. (1978). Egonomics, or the art of self-management. *The American Economic Review*, *68*(2), 290–294.

Schelling, T. C. (1984). Self–command in practice, in policy, and in a theory of rational choice. *The American Economic Review*, *74*(2), 1–11.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 99–118.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research*, *16*(2), 158–174.

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self–control. *Journal of Political Economy*, *89*(2), 392–406.

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological science*, *24*(6), 901–908.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review*, *76*(1), 31.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, *79*(4), 281.

Van Lint, J. H., & Wilson, R. M. (2001). Lattices and Möbius inversion. In *A course in combinatorics* (pp. 333–350). Cambridge University Press.

Zeckhauser, R. (1969). Majority rule with lotteries on alternatives. *The Quarterly Journal of Economics*, *83*(4), 696–703.

# A Appendix. Omitted Proofs

## A.1 Proof of Proposition 3.1

*Proof.*
Fix distinct $x, y \in X$. By Definitions 3.5 and 3.7,

$$\boldsymbol{\pi}^{P}_{\mu}(x, \{x, y\}) = \sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y\} = \boldsymbol{\pi}^{AP}_{\mu}(x, \{x, y\}).$$

□

## A.2 Proof of Proposition 3.2

*Proof.*
Fix $(x, B), (x, B') \in E$ with $B \subseteq B'$. Since for any $R \in \mathscr{P}$,

$$\mathbb{1}\{x \succ_R y : \forall y \in B - \{x\}\} \geqslant \mathbb{1}\{x \succ_R y : \forall y \in B' - \{x\}\}$$

We obtain that

$$\begin{aligned} \boldsymbol{\pi}^{P}_{\mu}(x, B) &= \sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y : \forall y \in B - \{x\}\} \\ &\geqslant \sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y : \forall y \in B' - \{x\}\} \\ &= \boldsymbol{\pi}^{P}_{\mu}(x, B'). \end{aligned}$$

□

## A.3 Proof of Proposition 3.3

*Proof.*
Fix distinct $x, y, z \in X$. By Proposition 3.1, it suffices to show that

$$\boldsymbol{\pi}^{P}_{\mu}(x, \{x, y\}) - \boldsymbol{\pi}^{P}_{\mu}(x, \{x, z\}) + \boldsymbol{\pi}^{P}_{\mu}(y, \{y, z\}) \geqslant 0$$

By Definition 3.5, we have that

$$\boldsymbol{\pi}^{P}_{\mu}(x, \{x, y\}) = \sum_{R \in \mathscr{P}} \mu(R) \left(\mathbb{1}\{x \succ_R y \succ_R z\} + \mathbb{1}\{x \succ_R z \succ_R y\} + \mathbb{1}\{z \succ_R x \succ_R y\}\right)$$

$$\boldsymbol{\pi}_\mu^P(x,\{x,z\}) = \sum_{R\in\mathscr{P}} \mu(R)\left(\mathbb{1}\{x \succ_R y \succ_R z\} + \mathbb{1}\{x \succ_R z \succ_R y\} + \mathbb{1}\{y \succ_R x \succ_R z\}\right)$$
$$\boldsymbol{\pi}_\mu^P(y,\{y,z\}) = \sum_{R\in\mathscr{P}} \mu(R)\left(\mathbb{1}\{x \succ_R y \succ_R z\} + \mathbb{1}\{y \succ_R z \succ_R x\} + \mathbb{1}\{y \succ_R x \succ_R z\}\right).$$

Since $0 \leqslant \mathbb{1}\{x \succ_R y \succ_R z\} + \mathbb{1}\{y \succ_R z \succ_R x\} + \mathbb{1}\{z \succ_R x \succ_R y\} \leqslant 1$, it follows that

$$\begin{aligned}&\boldsymbol{\pi}_\mu^P(x,\{x,y\}) - \boldsymbol{\pi}_\mu^P(x,\{x,z\}) + \boldsymbol{\pi}_\mu^P(y,\{y,z\})\\ =& \sum_{R\in\mathscr{P}} \mu(R)\left(\mathbb{1}\{x \succ_R y \succ_R z\} + \mathbb{1}\{y \succ_R z \succ_R x\} + \mathbb{1}\{z \succ_R x \succ_R y\}\right) \geqslant 0.\end{aligned}$$

□

## A.4 Proof of Theorem 3.1

*Remark.* The Möbius inversion formula is generally defined for functions over a locally finite poset (see, for instance, Cioabă and Murty (2009) and Van Lint and Wilson (2001)). Notably, $(\mathscr{B}, \subseteq)$ constitutes a locally finite poset, as the set inclusion relation $\subseteq$ is a partial order (i.e., reflexive, transitive, and antisymmetric) on $\mathscr{B}$, and for any $A \subseteq B \in \mathscr{B}$, the set $\{ C \in \mathscr{B} \mid A \subseteq C \subseteq B \}$ is finite. Let $f, g : \mathscr{B} \to \mathbb{R}$. Then, the Möbius inversion formula implies that $\forall\, A \in \mathscr{B}$,

$$f(A) = \sum_{B\in\mathscr{B}:A\subseteq B} g(B) \Leftrightarrow g(A) = \sum_{B\in\mathscr{B}:A\subseteq B} (-1)^{|B|-|A|} f(B).$$

This implication plays an important role in proving the representation theorem of the MS-AP Model.

*Proof.*
Let $X$ be a choice set and $\mu \in \Delta(\mathscr{P})$. For convenience, let $\forall\,(x, B) \in E$,

$$\nu_\pi^{AP}(x,B) = (|B|-1)\pi(x,B) \tag{4}$$
$$\boldsymbol{\nu}_\mu^{AP}(x,B) = (|B|-1)\boldsymbol{\pi}_\mu^{AP}(x,B). \tag{5}$$

We prove by first claiming important properties of $\psi_\pi$, $\gamma_\pi$, and $\boldsymbol{\psi}_\mu$ (see later for their proofs).

*Claim* A.1. Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. Then, $\psi_\pi$ satisfies the following properties:

$$\begin{cases} \psi_\pi(x,A) = \displaystyle\sum_{k=0}^{|X|-|A|}(-1)^k \left( \sum_{\substack{B\in\mathscr{B}:A\subseteq B,\\ |B|=|A|+k}} \Big(1-\nu_\pi^{AP}(x,B)\Big)\right), \quad \forall\,(x,A)\in E & (6)\\ \nu_\pi^{AP}(x,A) = 1 - \displaystyle\sum_{B\in\mathscr{B}:A\subseteq B} \psi_\pi(x,B), \quad \forall\,(x,A)\in E & (7)\\ \displaystyle\sum_{x\in X}\psi_\pi(x,\{x\}) = 1 & (8)\\ \displaystyle\sum_{x\in X-D}\psi_\pi(x,X-D) = \sum_{y\in D}\psi_\pi(y,X-D+\{y\}), \quad \forall\,D\in 2^X-\varnothing-X & (9) \end{cases}$$

*Claim* A.2. Let $(X,\mathscr{B},\pi)$ be a stochastic choice structure. If $\psi_\pi(x,A)\geqslant 0$, $\forall\,(x,A)\in E$, then $\gamma_\pi$ satisfies the following properties:

$$\begin{cases} \gamma_\pi(\sigma)\geqslant 0, \quad \forall\,\sigma\in\mathfrak{S} & (10)\\ \displaystyle\sum_{\sigma\in\mathfrak{S}(D)}\gamma_\pi(x,\sigma) = \psi_\pi(x,X-D), \quad \forall\,D\in 2^X-\varnothing-X \text{ and } \forall\,x\in X-D & (11)\\ \displaystyle\sum_{\sigma\in\mathfrak{S}(D)}\gamma_\pi(\sigma) = \sum_{x\in X-D}\psi_\pi(x,X-D), \quad \forall\,D\in 2^X-\varnothing-X & (12)\\ \gamma_\pi(\sigma) = \displaystyle\sum_{x\in X-D}\gamma_\pi(x,\sigma), \quad \forall\,D\in 2^X-\varnothing-X \text{ and } \forall\,\sigma\in\mathfrak{S}(D) & (13)\\ \gamma_\pi(\sigma_1) = \displaystyle\sum_{\sigma_2\in\mathfrak{S}(X-D)}\gamma_\pi(\sigma_2,\sigma_1), \quad \forall\,D\in 2^X-\varnothing-X \text{ and } \forall\,\sigma_1\in\mathfrak{S}(D) & (14)\\ \displaystyle\sum_{\sigma\in\mathfrak{S}(X)}\gamma_\pi(\sigma) = 1 & (15) \end{cases}$$

*Claim* A.3. Let $X$ be a choice set and $\mu\in\Delta(\mathscr{P})$. Then, $\boldsymbol{\psi}_\mu$ satisfies the following properties:

$$\begin{cases} \boldsymbol{\nu}_\mu^{AP}(x,A) = 1 - \displaystyle\sum_{B\in\mathscr{B}:A\subseteq B}\boldsymbol{\psi}_\mu(x,B), \quad \forall\,(x,A)\in E & (16)\\ \boldsymbol{\psi}_\mu(x,A) = \displaystyle\sum_{B\in\mathscr{B}:A\subseteq B}(-1)^{|B|-|A|}\Big(1-\boldsymbol{\nu}_\mu^{AP}(x,B)\Big), \quad \forall\,(x,A)\in E & (17) \end{cases}$$

*Necessity.* Suppose that $\pi$ can be rationalized by some MS-AP Model $\mu\in\Delta(\mathscr{P})$; or equivalently, $\forall\,(x,B)\in E$,

$$\nu_\pi^{AP}(x,B) = (|B|-1)\pi(x,B) = (|B|-1)\boldsymbol{\pi}_\mu^{AP}(x,B) = \boldsymbol{\nu}_\mu^{AP}(x,B).$$

We show that $\psi_\pi(x,A) \geqslant 0$, $\forall\,(x,A) \in E$.

Fix an arbitrary $(x,A) \in E$. Since by construction, $\boldsymbol{\psi}_\mu(x,A) \geqslant 0$, it suffices to show that $\psi_\pi(x,A) = \boldsymbol{\psi}_\mu(x,A)$.

It follows that,

$$\begin{aligned}
\psi_\pi(x,A) &= \sum_{B\in\mathscr{B}:A\subseteq B} (-1)^{|B|-|A|}\left(1-\nu_\pi^{AP}(x,B)\right) && \text{(by (1) and (4))}\\
&= \sum_{B\in\mathscr{B}:A\subseteq B} (-1)^{|B|-|A|}\left(1-\boldsymbol{\nu}_\mu^{AP}(x,B)\right) && \text{(by rationalizability)}\\
&= \boldsymbol{\psi}_\mu(x,A). && \text{(by (17))}
\end{aligned}$$

*Sufficiency.* Suppose that $\psi_\pi(x,A) \geqslant 0$, $\forall\,(x,A) \in E$. We show by construction that $\pi$ can be rationalized by the MS-AP Model $\mu$ where $\forall\,R \in \mathscr{P}$, $\mu(R) = \gamma_\pi(\rho(R))$.

As we can see from (10) and (15), the constructed $\mu$ is a valid multi-self system since for any $R \in \mathscr{P}$, $\mu(R) = \gamma_\pi(\rho(R)) \geqslant 0$ and $\sum_{R\in\mathscr{P}} \mu(R) = \sum_{R\in\mathscr{P}} \gamma_\pi(\rho(R)) = \sum_{\sigma\in\mathfrak{S}(X)} \gamma_\pi(\sigma) = 1$.

We then show that for any $(x,A) \in E$, $\nu_\pi^{AP}(x,A) = \boldsymbol{\nu}_\mu^{AP}(x,A)$ and therefore the MS-AP Model $\mu$ rationalizes $\pi$. Fix an arbitrary $(x,A) \in E$. It suffices to show that $\psi_\pi(x,A) = \boldsymbol{\psi}_\mu(x,A)$ since by (7) and (16), we already have that $\nu_\pi^{AP}(x,A) = 1 - \sum_{B\in\mathscr{B}:A\subseteq B} \psi_\pi(x,B)$ and $\boldsymbol{\nu}_\mu^{AP}(x,A) = 1 - \sum_{B\in\mathscr{B}:A\subseteq B} \boldsymbol{\psi}_\mu(x,B)$.

By (3) and the construction of $\mu$, we know that

$$\boldsymbol{\psi}_\mu(x,A) = \sum_{R\in\mathscr{P}} \gamma_\pi(\rho(R))\mathbb{1}\{A = U_R(x)\}. \tag{18}$$

Next, we show by cases that the equality holds as required.

If $|A| = 1$, namely, $A = \{x\}$, then

$$\begin{aligned}
\boldsymbol{\psi}_\mu(x,\{x\}) &= \sum_{\sigma\in\mathfrak{S}(X-\{x\})} \gamma_\pi(x,\sigma) && \text{(by (18))}\\
&= \psi_\pi(x,\{x\}). && \text{(by (11))}
\end{aligned}$$

If $|A| = |X|$, namely, $A = X$, then

$$\boldsymbol{\psi}_\mu(x, X) = \sum_{\sigma \in \mathfrak{S}(X - \{x\})} \gamma_\pi(\sigma, x) \qquad \text{(by (18))}$$
$$= \gamma_\pi(x) \qquad \text{(by (14))}$$
$$= \psi_\pi(x, X). \qquad \text{(by (2))}$$

If $2 \leqslant |A| \leqslant |X| - 1$, then

$$\boldsymbol{\psi}_\mu(x, A) = \sum_{\sigma \in \mathfrak{S}(X - A)} \sum_{\sigma' \in \mathfrak{S}(A - \{x\})} \gamma_\pi((\sigma', x), \sigma) \qquad \text{(by (18))}$$
$$= \sum_{\sigma \in \mathfrak{S}(X - A)} \sum_{\sigma' \in \mathfrak{S}(A - \{x\})} \gamma_\pi(\sigma', (x, \sigma))$$
$$= \sum_{\sigma \in \mathfrak{S}(X - A)} \gamma_\pi(x, \sigma) \qquad \text{(by (14))}$$
$$= \psi_\pi(x, A). \qquad \text{(by (11))}$$

$\square$

### A.4.1 Proof of Claim A.1

*Proof.*
*Proof of (6).* Fix an arbitrary $(x, A) \in E$. Simply rewriting (1) and (4) gives us:

$$\psi_\pi(x, A) = \sum_{k=0}^{|X|-|A|} \sum_{\substack{B \in \mathscr{B}: A \subseteq B, \\ |B| = |A| + k}} (-1)^k \left(1 - \nu_\pi^{AP}(x, B)\right)$$
$$= \sum_{k=0}^{|X|-|A|} (-1)^k \left( \sum_{\substack{B \in \mathscr{B}: A \subseteq B, \\ |B| = |A| + k}} \left(1 - \nu_\pi^{AP}(x, B)\right) \right).$$

*Proof of (7).* Fix an arbitrary $(x, A) \in E$. Recall from (1) and (4) that

$$\psi_\pi(x, A) = \sum_{B \in \mathscr{B}: A \subseteq B} (-1)^{|B|-|A|} \left(1 - \nu_\pi^{AP}(x, B)\right).$$

Apply the Möbius inversion formula, we have that

$$1-\nu_\pi^{AP}(x,B)=\sum_{B\in\mathscr{B}:A\subseteq B}\psi_\pi(x,B).$$

Thus, $\nu_\pi^{AP}(x,B)=1-\sum_{B\in\mathscr{B}:A\subseteq B}\psi_\pi(x,B)$.

*Proof of (8).*

$$\begin{aligned}\sum_{x\in X}\psi_\pi(x,\{x\}) &= \sum_{x\in X}\sum_{k=0}^{|X|-1}(-1)^k\left(\sum_{\substack{B\in\mathscr{B}:\{x\}\subseteq B,\\ |B|=k+1}}\Big(1-\nu_\pi^{AP}(x,B)\Big)\right) \quad \text{(by (6))}\\ &=\sum_{k=0}^{|X|-1}(-1)^k\left(\sum_{x\in X}\sum_{\substack{B\in\mathscr{B}:\{x\}\subseteq B,\\ |B|=k+1}}\Big(1-\nu_\pi^{AP}(x,B)\Big)\right).\end{aligned}$$

For any integer $k$ with $0\leqslant k\leqslant |X|-1$, since $\forall\, B\in\mathscr{B}$, $\sum_{x\in B}\pi(x,B)=1$,

$$\begin{aligned}\sum_{x\in X}\sum_{\substack{B\in\mathscr{B}:\{x\}\subseteq B,\\ |B|=k+1}}\Big(1-\nu_\pi^{AP}(x,B)\Big) &= \sum_{\substack{B\in\mathscr{B}:\\ |B|=k+1}}\sum_{x\in B}\Big(1-\nu_\pi^{AP}(x,B)\Big)\\ &=\sum_{\substack{B\in\mathscr{B}:\\ |B|=k+1}}\left(|B|-\sum_{x\in B}(|B|-1)\,\pi(x,B)\right) \quad \text{(by (4))}\\ &=\sum_{\substack{B\in\mathscr{B}:\\ |B|=k+1}}(|B|-(|B|-1))\\ &=\binom{|X|}{k+1}.\end{aligned}$$

Hence, by the binomial formula,

$$\begin{aligned}\sum_{x\in X}\psi_\pi(x,\{x\}) &= \sum_{k=0}^{|X|-1}(-1)^k\binom{|X|}{k+1}=-\sum_{k=1}^{|X|}(-1)^k\binom{|X|}{k}=1-\sum_{k=0}^{|X|}(-1)^k\binom{|X|}{k}\\ &=1-(-1+1)^{|X|}=1.\end{aligned}$$

Proof of (9). It is equivalent to show that $\forall\, A\in 2^X\text{–}\varnothing\text{–}X$,

$$\sum_{x\in A}\psi_\pi(x,A)=\sum_{y\in X-A}\psi_\pi(y,A+\{y\}).$$

Fix an arbitrary $A \in 2^X - \varnothing - X$. First, consider the LHS of this equation.

$$\begin{aligned}
\text{LHS} &= \sum_{x\in A} \psi_\pi(x, A) \\
&= \sum_{x\in A} \sum_{k=0}^{|X|-|A|} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}: A\subseteq B,\\ |B|=|A|+k}} \left(1 - \nu_\pi^{AP}(x, B)\right) \right) \quad \text{(by (6))} \\
&= \sum_{x\in A} \left(1 - \nu_\pi^{AP}(x, A)\right) + \sum_{k=1}^{|X|-|A|} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}: A\subseteq B,\\ |B|=|A|+k}} \sum_{x\in A} \left(1 - \nu_\pi^{AP}(x, B)\right) \right).
\end{aligned}$$

Since $\sum_{x\in A} \pi(x, A) = 1$, and for any $B \in \mathscr{B}$ with $A \subset B$, $\sum_{x\in A} \pi(x, B) = 1 - \sum_{y\in B-A} \pi(y, B)$, we have that

$$\begin{aligned}
\sum_{x\in A} \left(1 - \nu_\pi^{AP}(x, A)\right) &= |A| - \sum_{x\in A} (|A| - 1)\,\pi(x, A) \quad \text{(by (4))} \\
&= 1,
\end{aligned}$$

and that for any $B \in \mathscr{B}$ with $A \subset B$,

$$\begin{aligned}
\sum_{x\in A} \left(1 - \nu_\pi^{AP}(x, B)\right) &= |A| - \sum_{x\in A} (|B| - 1)\,\pi(x, B) && \text{(by (4))} \\
&= |A| - (|B| - 1) \left(1 - \sum_{y\in B-A} \pi(y, B)\right) \\
&= 1 - (|B| - |A|) + (|B| - 1) \sum_{y\in B-A} \pi(y, B) \\
&= 1 - \sum_{y\in B-A} (1 - (|B| - 1)\,\pi(y, B)) \\
&= 1 - \sum_{y\in B-A} \left(1 - \nu_\pi^{AP}(y, B)\right). && \text{(by (4))}
\end{aligned}$$

Then, substitute the two parts back into the LHS:

$$\text{LHS} = 1 + \sum_{k=1}^{|X|-|A|} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}: A\subseteq B,\\ |B|=|A|+k}} \left(1 - \sum_{y\in B-A} \left(1 - \nu_\pi^{AP}(y, B)\right)\right) \right)$$

$$
\begin{aligned}
&= 1 + \sum_{k=1}^{|X|-|A|} (-1)^k \binom{|X|-|A|}{k} - \sum_{k=1}^{|X|-|A|} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}:A\subseteq B,\\ |B|=|A|+k}} \sum_{y\in B-A} \Big(1-\nu_\pi^{AP}(y,B)\Big) \right) \\
&= (-1+1)^{|X|-|A|} + \sum_{k=1}^{|X|-|A|} (-1)^{k-1} \left( \sum_{\substack{B\in\mathscr{B}:A\subseteq B,\\ |B|=|A|+k}} \sum_{y\in B-A} \Big(1-\nu_\pi^{AP}(y,B)\Big) \right) \\
&= \sum_{k=0}^{|X|-|A|-1} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}:A\subseteq B,\\ |B|=|A|+k}} \sum_{y\in B-A} \Big(1-\nu_\pi^{AP}(y,B)\Big) \right).
\end{aligned}
$$

Then, consider the RHS of this equation.

$$
\begin{aligned}
\text{RHS} &= \sum_{y\in X-A} \psi_\pi(y, A+\{y\}) \\
&= \sum_{y\in X-A} \sum_{k=0}^{|X|-|A|-1} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}:A+\{y\}\subseteq B,\\ |B|=|A|+k+1}} \Big(1-\nu_\pi^{AP}(y,B)\Big) \right) \quad \text{(by (6))} \\
&= \sum_{k=0}^{|X|-|A|-1} (-1)^k \left( \sum_{y\in X-A} \sum_{\substack{B\in\mathscr{B}:A+\{y\}\subseteq B,\\ |B|=|A|+k+1}} \Big(1-\nu_\pi^{AP}(y,B)\Big) \right) \\
&= \sum_{k=0}^{|X|-|A|-1} (-1)^k \left( \sum_{\substack{B\in\mathscr{B}:A\subseteq B,\\ |B|=|A|+k+1}} \sum_{y\in B-A} \Big(1-\nu_\pi^{AP}(y,B)\Big) \right).
\end{aligned}
$$

Thus, we reach the conclusion that the LHS equals the RHS. □

### A.4.2 Proof of Claim A.2

*Proof.*
Suppose throughout the whole proof that $\psi_\pi(x, A) \geqslant 0$, $\forall\,(x, A) \in E$.

Note that for any $B \in \mathscr{B}$ with $|B| \geqslant 2$,

$$
\begin{aligned}
\{\,\sigma \mid \sigma \in \mathfrak{S}(B)\,\} &= \{\,(\hat\sigma, x) \mid x \in B, \hat\sigma \in \mathfrak{S}(B-\{x\})\,\} \\
&= \{\,(x, \hat\sigma) \mid x \in B, \hat\sigma \in \mathfrak{S}(B-\{x\})\,\}.
\end{aligned} \tag{19}
$$

Proof of (10). Since $\psi_\pi(x, A) \geqslant 0$, $\forall\,(x, A) \in E$, it can be easily seen from the definition of $\gamma_\pi$ in (2) that $\forall\,\sigma \in \mathfrak{S}$, $\gamma_\pi(\sigma) \geqslant 0$.

Proof of (11). Since for any $D \in 2^X - \varnothing - X$, $|D| = k$ for some integer $k$ with $1 \leqslant k \leqslant |X| - 1$, we show the statement (11) by induction.

*Base case:* For any $D \in 2^X - \varnothing - X$ with $|D| = 1$, we assume w.l.o.g. that $D = \{y\}$ for some $y \in X$. Then, the statement reduces to $\forall\,x \in X - \{y\}$, $\gamma_\pi(x, y) = \psi_\pi(x, X - \{y\})$. Fix an arbitrary $x \in X - \{y\}$.

If $\gamma_\pi(y) \neq 0$, then

$$\begin{aligned}\gamma_\pi(x, y) &= \frac{\psi_\pi(x, X - \{y\}) \cdot \gamma_\pi(y)}{\gamma_\pi(y)} \qquad \text{(by (2))}\\ &= \psi_\pi(x, X - \{y\}).\end{aligned}$$

If $\gamma_\pi(y) = 0$, then by (2), $\gamma_\pi(x, y) = 0$. We then show that $\psi_\pi(x, X - \{y\}) = 0$. It suffices to show that $\sum_{x \in X - \{y\}} \psi_\pi(x, X - \{y\}) = 0$ since for any $(x, A) \in E$, $\psi_\pi(x, A) \geqslant 0$.

It follows that

$$\begin{aligned}\sum_{x \in X - \{y\}} \psi_\pi(x, X - \{y\}) &= \psi_\pi(y, X) && \text{(by (9))}\\ &= \gamma_\pi(y) && \text{(by (2))}\\ &= 0.\end{aligned}$$

*Induction step:* Assume the induction hypothesis that for any integer $k$ ($1 \leqslant k \leqslant |X| - 2$), it holds that $\forall\,C \in 2^X - \varnothing - X$ with $|C| = k$ and $\forall\,z \in X - C$,

$$\sum_{\hat{\sigma} \in \mathfrak{S}(C)} \gamma_\pi(z, \hat{\sigma}) = \psi_\pi(z, X - C). \tag{20}$$

Now consider any $D \in 2^X - \varnothing - X$ with $|D| = k + 1$ and any $x \in X - D$.

If $\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') \neq 0$, then

$$\begin{aligned}\sum_{\sigma \in \mathfrak{S}(D)} \gamma_\pi(x, \sigma) &= \sum_{\sigma \in \mathfrak{S}(D)} \frac{\psi_\pi(x, X - D) \cdot \gamma_\pi(\sigma)}{\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma')} \qquad \text{(by (2))}\\ &= \psi_\pi(x, X - D) \frac{\sum_{\sigma \in \mathfrak{S}(D)} \gamma_\pi(\sigma)}{\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma')}\end{aligned}$$

$$= \psi_\pi(x, X{-}D).$$

If $\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') = 0$, then by (2), $\sum_{\sigma \in \mathfrak{S}(D)} \gamma_\pi(\sigma, x) = 0$. We next show that $\psi_\pi(x, X{-}D) = 0$. It suffices to show that $\sum_{x \in X-D} \psi_\pi(x, X{-}D) = 0$.

By (10), $\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') = 0$ implies that $\gamma_\pi(\sigma) = 0$ for any $\sigma \in \mathfrak{S}(D)$. Let $C = D{-}\{y\}$ for any $y \in D$. Then, by (19), we have that $\gamma_\pi(y, \hat{\sigma}) = 0$ for any $y \in D$ and $\hat{\sigma} \in \mathfrak{S}(C)$.

It follows that

$$\begin{aligned}
\sum_{x \in X-D} \psi_\pi(x, X{-}D) &= \sum_{y \in D} \psi_\pi(y, X{-}D + \{y\}) && \text{(by (9))} \\
&= \sum_{y \in D} \psi_\pi(y, X{-}C) \\
&= \sum_{y \in D} \sum_{\hat{\sigma} \in \mathfrak{S}(C)} \gamma_\pi(y, \hat{\sigma}) && \text{(by (20))} \\
&= 0.
\end{aligned}$$

*Proof of (12).* For any $D \in 2^X{-}\varnothing{-}X$ with $|D| = 1$, we assume w.l.o.g. that $D = \{y\}$ for some $y \in X$. Then, the statement reduces to $\gamma_\pi(y) = \sum_{x \in X-\{y\}} \psi_\pi(x, X{-}\{y\})$.

It follows that

$$\begin{aligned}
\gamma_\pi(y) &= \psi_\pi(y, X) && \text{(by (2))} \\
&= \sum_{x \in X-\{y\}} \psi_\pi(x, X{-}\{y\}). && \text{(by (9))}
\end{aligned}$$

Now consider an arbitrary $D \in 2^X{-}\varnothing{-}X$ with $|D| \geqslant 2$.

$$\begin{aligned}
\sum_{\sigma \in \mathfrak{S}(D)} \gamma_\pi(\sigma) &= \sum_{y \in D} \sum_{\hat{\sigma} \in \mathfrak{S}(D-\{y\})} \gamma_\pi(y, \hat{\sigma}) && \text{(by (19))} \\
&= \sum_{y \in D} \psi_\pi(y, X{-}D + \{y\}) && \text{(by (11))} \\
&= \sum_{x \in X-D} \psi_\pi(x, X{-}D). && \text{(by (9))}
\end{aligned}$$

*Proof of (13).* Consider any $D \in 2^X{-}\varnothing{-}X$ and any $\sigma \in \mathfrak{S}(D)$.

If $\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') \neq 0$, then

$$\begin{aligned}
\sum_{x \in X-D} \gamma_\pi(x, \sigma) &= \sum_{x \in X-D} \frac{\psi_\pi(x, X-D) \cdot \gamma_\pi(\sigma)}{\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma')} && \text{(by (2))} \\
&= \gamma_\pi(\sigma) \frac{\sum_{x \in X-D} \psi_\pi(x, X-D)}{\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma')} \\
&= \gamma_\pi(\sigma) \frac{\sum_{\sigma \in \mathfrak{S}(D)} \gamma_\pi(\sigma)}{\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma')} && \text{(by (12))} \\
&= \gamma_\pi(\sigma).
\end{aligned}$$

If $\sum_{\sigma' \in \mathfrak{S}(D)} \gamma_\pi(\sigma') = 0$, which by (10), implies that $\gamma_\pi(\sigma) = 0$, then by (2), $\sum_{x \in X-D} \gamma_\pi(x, \sigma) = 0 = \gamma_\pi(\sigma)$.

*Proof of (14).* Since for any $D \in 2^X - \varnothing - X$, $|D| = |X| - k$ for some integer $k$ with $1 \leqslant k \leqslant |X| - 1$, we show the statement (14) by induction.

*Base case:* For any $D \in 2^X - \varnothing - X$ with $|D| = X - 1$, we assume w.l.o.g. that $D = X - \{y\}$ for some $y \in X$. Then, the statement reduces to that for any $\sigma_1 \in \mathfrak{S}(X - \{y\})$, $\gamma_\pi(\sigma_1) = \gamma_\pi(y, \sigma_1)$, which is true directly from (13).

*Induction step:* Assume the induction hypothesis that for any integer $k$ ($1 \leqslant k \leqslant |X| - 2$), it holds that for any $C \in 2^X - \varnothing - X$ with $|C| = |X| - k$ and any $\sigma'_1 \in \mathfrak{S}(C)$,

$$\gamma_\pi(\sigma'_1) = \sum_{\sigma'_2 \in \mathfrak{S}(X-C)} \gamma_\pi(\sigma'_2, \sigma'_1). \tag{21}$$

Now consider any $D \in 2^X - \varnothing - X$ with $|D| = |X| - (k+1)$ and any $\sigma_1 \in \mathfrak{S}(D)$. We have that

$$\begin{aligned}
\gamma_\pi(\sigma_1) &= \sum_{x \in X-D} \gamma_\pi(x, \sigma_1) && \text{(by (13))} \\
&= \sum_{x \in X-D} \sum_{\sigma'_2 \in \mathfrak{S}(X-D-\{x\})} \gamma_\pi(\sigma'_2, (x, \sigma_1)) && \text{(by (21))} \\
&= \sum_{x \in X-D} \sum_{\sigma'_2 \in \mathfrak{S}(X-D-\{x\})} \gamma_\pi((\sigma'_2, x), \sigma_1)) \\
&= \sum_{\sigma_2 \in \mathfrak{S}(X-D)} \gamma_\pi(\sigma_2, \sigma_1). && \text{(by (19))}
\end{aligned}$$

*Proof of (15).*

$$\begin{aligned}
\sum_{\sigma\in\mathfrak{S}(X)} \gamma_\pi(\sigma) &= \sum_{x\in X} \sum_{\hat\sigma\in\mathfrak{S}(X-\{x\})} \gamma_\pi(x,\hat\sigma) && \text{(by (19))}\\
&= \sum_{x\in X} \psi_\pi(x,\{x\}) && \text{(by (11))}\\
&= 1. && \text{(by (8))}
\end{aligned}$$

□

### A.4.3 Proof of Claim A.3

*Proof.*
*Proof of (16).* By (4) and the definition of $\pi_\mu^{AP}$, we have that for any $(x, A) \in E$,

$$\begin{aligned}
\boldsymbol{\nu}_\mu^{AP}(x,A) &= \sum_{R\in\mathscr{P}} \mu(R)\mathbb{1}\{x \succ_R y : \exists\, y \in A{-}\{x\}\}\\
&= \sum_{R\in\mathscr{P}} \mu(R)\left(1 - \sum_{B\in\mathscr{B}:A\subseteq B} \mathbb{1}\{B = U_R(x)\}\right)\\
&= \sum_{R\in\mathscr{P}} \mu(R) - \sum_{R\in\mathscr{P}} \mu(R)\left(\sum_{B\in\mathscr{B}:A\subseteq B} \mathbb{1}\{B = U_R(x)\}\right)\\
&= 1 - \sum_{B\in\mathscr{B}:A\subseteq B}\left(\sum_{R\in\mathscr{P}} \mu(R)\mathbb{1}\{B = U_R(x)\}\right)\\
&= 1 - \sum_{B\in\mathscr{B}:A\subseteq B} \boldsymbol{\psi}_\mu(x,B). \quad \text{(by (3))}
\end{aligned}$$

*Proof of (17).* Applying the Möbius inversion formula on (16) gives us that $\forall\,(x, A) \in E$,

$$\boldsymbol{\psi}_\mu(x,A) = \sum_{B\in\mathscr{B}:A\subseteq B} (-1)^{|B|-|A|}\left(1 - \boldsymbol{\nu}_\mu^{AP}(x,B)\right).$$

□

# B Appendix. Supplements

## B.1 A Representation Theorem for MS-P Model

Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. We first define directly from $\pi$ a function $\phi_\pi : E \to \mathbb{R}$ where $\forall\,(x, A) \in E$,

$$\phi_\pi(x, A) = \sum_{B \in \mathscr{B}: A \subseteq B} (-1)^{|B|-|A|} \pi(x, B). \tag{22}$$

We then define from $\phi_\pi$ another function $\beta_\pi : \mathfrak{S} \to \mathbb{R}$ where

$$\begin{aligned}
&\forall\, x \in X,\ \beta_\pi(x) = \phi_\pi(x, X), \\
&\forall\,(\sigma, x) \text{ such that } \sigma \in \mathfrak{S}(D), x \in X - D \text{ for some } D \in 2^X - \varnothing - X, \\
&\beta_\pi(\sigma, x) = \begin{cases} \dfrac{\phi_\pi(x, X-D) \cdot \beta_\pi(\sigma)}{\sum_{\sigma' \in \mathfrak{S}(D)} \beta_\pi(\sigma')} & , \text{ if } \displaystyle\sum_{\sigma' \in \mathfrak{S}(D)} \beta_\pi(\sigma') \neq 0, \\ 0 & , \text{ if } \displaystyle\sum_{\sigma' \in \mathfrak{S}(D)} \beta_\pi(\sigma') = 0, \end{cases}
\end{aligned} \tag{23}$$

by noticing that $\mathfrak{S} = X \cup \{\,(\sigma, x) \mid \exists\, D \in 2^X - \varnothing - X : \sigma \in \mathfrak{S}(D), x \in X - D\,\}$.

**Theorem B.1** (MS-P Representation)**.** *Let $(X, \mathscr{B}, \pi)$ be a stochastic choice structure. $\pi \in \Pi^{MS\text{-}P}$ if and only if $\phi_\pi(x, A) \geqslant 0$, $\forall\,(x, A) \in E$. Furthermore, if $\pi \in \Pi^{MS\text{-}P}$, then there exists $\mu \in \Delta(\mathscr{P})$ where $\mu(R) = \beta_\pi(\rho(R))$, $\forall\, R \in \mathscr{P}$ such that $\pi(x, B) = \boldsymbol{\pi}^P_\mu(x, B)$, $\forall\,(x, B) \in E$.*

*Proof.*
The proof is omitted here and can be obtained from Barberá and Pattanaik (1986) and Falmagne (1978). □

## B.2 Aggregation by Borda Count Rule

Under the Borda Count rule, a voter assigns scores to alternatives based on their ranking: zero to the bottom-ranked alternative, one to the next-to-last alternative, and so forth, i.e., the normalized scoring vector is $(\frac{2(m-1)}{m(m-1)}, \frac{2(m-2)}{m(m-1)}, \ldots, 0)$ for $m \geqslant 2$ alternatives. Alternatively, we denote as $s^{BC}_R(x, B)$ the score assigned by an $R$-type self to an arbitrary alternative $x$ in a choice situation $B$ under the Borda

Count rule, where

$$s_R^{BC}(x,B) = \begin{cases} 1 & \text{, for } |B| = 1, \\ \dfrac{2}{|B|(|B|-1)} \displaystyle\sum_{y \in B - \{x\}} \mathbb{1}\{x \succ_R y\} & \text{, for } |B| \geqslant 2. \end{cases}$$

Formally, we define the *MS-BC Model* exploiting Definition 3.4 as follows.

**Definition B.1** (MS-BC Model)**.** Let $X$ be an alternative space. Suppose that an agent has a multi-self system $\mu \in \Delta(\mathscr{P})$ voting in a probabilistic manner by the Borda Count rule. We refer to such $\mu$ as an *MS-BC Model*. The *generated stochastic choice function* $\boldsymbol{\pi}_\mu^{BC}$ is determined as $\forall\,(x,B) \in E$,

$$\boldsymbol{\pi}_\mu^{BC}(x,B) = \begin{cases} 1 & \text{, for } |B| = 1, \\ \dfrac{2}{|B|(|B|-1)} \displaystyle\sum_{R \in \mathscr{P}} \mu(R) \left( \sum_{y \in B - \{x\}} \mathbb{1}\{x \succ_R y\} \right) & \text{, for } |B| \geqslant 2. \end{cases}$$

We denote by $\Pi^{MS\text{-}BC}$ the set of all possible stochastic choice functions generated by some MS-BC Model $\mu \in \Delta(\mathscr{P})$.

**Proposition B.1.** *Let $X$ be an alternative space and $\mu \in \Delta(\mathscr{P})$ be a multi-self system. Then, $\boldsymbol{\pi}_\mu^{BC}$ satisfies that for any $(x,B) \in E$ with $|B| \geqslant 2$, $\boldsymbol{\pi}_\mu^{BC}(x,B) = \frac{2}{|B|(|B|-1)} \sum_{y \in B - \{x\}} \boldsymbol{\pi}_\mu^{BC}(x,\{x,y\})$.*

*Proof.*
Simply applying Definition B.1, we have that for any $(x,B) \in E$ with $|B| \geqslant 2$,

$$\begin{aligned} \boldsymbol{\pi}_\mu^{BC}(x,B) &= \frac{2}{|B|(|B|-1)} \sum_{y \in B - \{x\}} \left( \sum_{R \in \mathscr{P}} \mu(R) \mathbb{1}\{x \succ_R y\} \right) \\ &= \frac{2}{|B|(|B|-1)} \sum_{y \in B - \{x\}} \boldsymbol{\pi}_\mu^{BC}(x,\{x,y\}). \end{aligned}$$

□

This proposition asserts that in the MS-BC Model, the likelihood of choosing from choice sets with more than two alternatives can be fully derived from binary choice probabilities. This restrictiveness limits the empirical applicability of the MS-BC Model, as real-world observed choice patterns are rarely expected to conform to this stringent prediction based on aggregating preferences by the Borda Count rule.