Associative Networks in Decision Making*

Jiangtao Li † Rui Tang ‡ Mu Zhang §

June 19, 2023

Abstract

We model associative networks that capture how a decision maker enlarges her consideration set through the mental association between alternatives, and demonstrate how this model serves as a toolbox to understand the impact of mental association on decision making. As a proof of concept, we characterize this model within a random attention framework and demonstrate that all the relevant parameters are uniquely identifiable. Notably, in a novel choice domain where not all perceivable alternatives are feasible, the presence of infeasible yet perceivable alternatives can influence the choice frequencies of alternatives through mental association.

Keywords: Associative network; Limited attention; Random choice; Feasibility and perceivability

JEL: D01, D91

---

*We thank Pawel Dziewulski, Junnan He, Matthew Kovach, Yusufcan Masatlioglu, Pietro Ortoleva, John Quah, Satoru Takahashi, and Junjie Zhou for helpful discussions on earlier versions of this paper. All remaining errors are our own. Rui Tang acknowledges the support of the HKUST start-up research fund.

†School of Economics, Singapore Management University, jtli@smu.edu.sg

‡Department of Economics, Hong Kong University of Science and Technology, ruitang@ust.hk

§Department of Economics, University of Michigan, muzhang@umich.edu

1 Introduction

Memory and attention are fundamental cognitive processes that are critical in decision making (Simon, 1955; Camerer, 1997; Payne, Bettman, and Johnson, 1993). The impact of these processes on decision making has been the subject of extensive research in recent years (Bordalo, Gennaioli, and Shleifer, 2020, 2022). Among the various patterns of memory, mental association is a crucial phenomenon that links the recall of one item to another based on individuals' prior experience or learning. For instance, in the context of consumer behavior, a consumer may associate Margarita with lemon-lime soda due to their similarity in color. Mental association has been widely studied in various fields, including psychology and marketing, and is shown to play a crucial role in consumer behavior (Loewenstein and O'Donoghue, 1997). In this paper, we adopt a choice-theoretical perspective to investigate the impact of mental association on decision making and develop a choice model to capture the effects of this cognitive process.

To provide a motivation for our choice model, consider the case of an athlete who just completed her training and is now standing in front of a vending machine. She initially notices a new version of Pepsi Cola. Before she makes the purchase, she recalls that a similar version of Coca Cola has recently been launched and is also available in the vending machine. After comparing the two options, she decides to purchase Coca Cola instead. The decision-making process in this scenario is influenced by the decision maker's (DM) mental association between Pepsi Cola and Coca Cola, as the initial attention towards Pepsi Cola leads to a consideration of other relevant options and ultimately results in the selection of Coca Cola.

Mental association is a cognitive process that enables DMs to expand their consideration set by linking relevant alternatives that may not have been initially considered. We capture this cognitive process through the use of associative networks, a conceptual model first introduced in cognitive psychology (Anderson and Bower, 1973; Anderson, 1996; Raaijmakers and Shiffrin, 1981) and widely applied in the marketing literature (Keller, 1993; Teichert and Schöntag, 2010; Brandt, De Mortanges, Bluemelhuber, and Van Riel, 2011; Cunha Jr, Forehand, and Angle, 2015). In an associative network, objects (nodes) are interconnected based on their semantic or conceptual relationships. When a particular node or input is activated, the network retrieves related nodes by spreading activation through the interconnected links. In our study, we employ the associative network as a descriptive model that captures how the attention given to one alternative can trigger the DM to pay attention to another alternative, abstracting away from the underlying conceptual similarities between alternatives that result in the association. More specifically, in our model, a link from alternative x to y indicates that the attention given to x can prompt the DM to further consider y.

Notably, the mental association process is not limited to feasible alternatives and may also be triggered by infeasible but perceivable options. Building on the previous example, even if Pepsi Cola is sold out in the vending machine, the athlete's attention towards it may still lead to the consideration of Coca Cola. Section 2 extends the choice domain to include such infeasible but perceivable alternatives. Specifically, we introduce the concept of **extended menus**, which comprises two distinct sets of alternatives: A, containing all feasible alternatives, and S, comprising the infeasible ones. The union of A and S comprises all **perceivable** alternatives. In the context of a vending machine, $A \cup S$ can be considered as all the drinks displayed, with A comprising those that are available and S containing those that are sold out. Although the DM cannot select any alternative from S, its presence can influence the DM's attention through association, thereby impacting the DM's choices. The primitive studied in this paper is a random choice rule, which is a function that maps every extended menu (A, S) to a distribution over A. This distribution represents the DM's choice frequencies of alternatives in A when confronted with the extended menu (A, S).

Section 3 introduces our choice model. Presented with an extended menu (A, S), the DM initially considers a random subset of alternatives in $A \cup S$, which we refer to as her initial consideration set. Following Manzini and Mariotti (2014) (henceforth MM14), we assume that the DM pays attention to each alternative in this set in a random and independent manner. The DM then proceeds to associate relevant alternatives in $A \cup S$ with alternatives in her initial consideration set, with the mental association process being captured by a directed graph over the alternatives that represents the DM's associative network. Each link of the associative network takes the form of an ordered pair of alternatives (x, y), indicating the consideration of x can prompt the DM to consider y. The association process is terminated when she cannot associate more alternatives in $A \cup S$ with her considered ones. This results in the formation of a final consideration set B. The DM then selects her most preferred alternative among the feasible ones in B, i.e., she chooses her most preferred alternative in $B \cap A$ if it is not empty. Otherwise, the default option is selected. We refer to the random choice rule induced by this choice procedure as the random consideration and association rule (RCAR).

We present five axioms to characterize RCARs. Axiom 1 specifies the baseline attention distribution of the DM, where she randomly and independently allocates attention to each alternative. Axiom 2 imposes a monotonicity condition on the random choice rule. The mental association process is captured by Axioms 3 and 4. Axiom 3 stipulates that the DM can associate more alternatives with a given alternative when there are more perceivable ones that serve as cues, while Axiom 4 asserts that if the DM can associate y with x in the presence of z but fails to do so when z is not perceivable, then she must associate z with x and y with z. Finally, Axiom 5 states that the effect of one alternative on the choice frequency of another alternative is asymmetric, as it depends on the DM's preference ordering over the two alternatives. Taken together, these axioms fully characterize our choice model. We emphasize that all parameters of our model can be uniquely identified under the new choice domain. In Section 5, we provide an additional characterization of our model in situations where all perceivable alternatives are feasible.

The remainder of the Introduction discusses some related literature. We introduce baseline notations and definitions in Section 2 and present our model of association in Section 3. Section 4 characterizes our choice model, and Section 5 studies our model in a restricted choice domain. Section 6 contains several extensions of our main model, and Section 7 concludes the paper. The Appendix collects proofs omitted from the main body of the paper.

1.1 Related Literature

Our paper contributes to the growing literature on choices with limited attention.[1] In particular, our approach is closely related to that of MM14, as both models assume that the DM allocates her attention randomly and independently. However, our model differs from the model of MM14 in that our DM has a follow-up procedure through which she continues to expand her consideration set via mental association. When considering the domain where all perceivable alternatives are feasible, our model generalizes that of MM14 and connects it with the rational choice model. Specifically, when the DM does not engage in any mental association, our model reduces to that of MM14. On the other hand, when each pair of alternatives are associated with each other, our model converges to the rational choice model: the DM always selects the best feasible alternative whenever she initially pays attention to a non-empty set of perceivable option.

Our model makes three novel contributions to the literature on limited attention. First, we examine the cognitive process of mental association, which is a fundamental mechanism in forming the DM's consideration set. We provide a concrete procedure for how this process operates. Second, our model incorporates both bottom-up attention (initial random attention) and top-down attention (mental association), which have been shown to be influential factors in decision making (Mogg, Bradley, Dixon, Fisher, Twelftree, and McWilliams, 2004; Geng and Behrmann, 2005; Jackson, Stafford, and Smith, 2009).[2] Third, we investigate the impact of infeasible but perceivable alternatives on the DM's attention. This new framework enables us to obtain novel testable implications.

More broadly, our paper contributes to the literature on random choices, which

[1]See, for instance, Masatlioglu, Nakajima, and Ozbay (2012), Brady and Rehbeck (2016), MM14, Dean, Kıbrıs, and Masatlioglu (2017), Lleras, Masatlioglu, Nakajima, and Ozbay (2017), Cattaneo, Ma, Masatlioglu, and Suleymanov (2020), Dardanoni, Manzini, Mariotti, and Tyson (2020), Cattaneo, Cheung, Ma, and Masatlioglu (2021), etc.

[2]Bottom-up attention involves the automatic processing of sensory stimuli in the environment, such as sudden loud noises or bright lights, that capture an individual's attention involuntarily. By contrast, top-down attention refers to the deliberate allocation of attention that is guided by an individual's goals, expectations, and prior knowledge. In our model, the DM's initial attention is more likely to be bottom-up, as the DM is randomly attracted to the stimuli or salient features of the options. However, the second-stage mental association is more likely to be top-down, as individuals can direct their attention towards information or options that are relevant to their self-concept (Sui, Gu, and Han, 2012).

has sought to explain the occurrence of stochastic decision making in human behavior. Various explanations have been proposed, including the possibility that the DM has random utilities, leading to stochastic choices as a result of utility maximization (Block and Marschak, 1960; Falmagne, 1978; Gul and Pesendorfer, 2006; Gul, Natenzon, and Pesendorfer, 2014)[3], and the possibility that the DM randomizes deliberately (Cerreia-Vioglio, Dillenberger, Ortoleva, and Riella, 2019; Agranov and Ortoleva, 2022). While the randomness in our DM's choice behavior is driven by random attention, our model highlights the potential for choices to be influenced by options that are not included in the feasible choice set. Our work emphasizes the importance of understanding mental associations as a potential source of random choices in cases where the DM's set of perceivable alternatives cannot be fully observed.

Our work also relates to the literature on how choices are influenced by factors beyond the choice menu. These factors can include frames (Salant and Rubinstein, 2008), the DM's reference points or status quo (Masatlioglu and Ok, 2005, 2014; Kovach and Suleymanov, 2021), and recommendations from external sources (Cheung and Masatlioglu, 2021), among others. While our approach shares some similarities with the work of Kovach and Suleymanov (2021), which examines how reference points can shape the DM's attention, our study focuses on the impact of infeasible alternatives on the DM's consideration set.

Another notable paper that has close connection to ours is Masatlioglu and Nakajima (2013) (henceforth MN13). MN13 examine an agent who initially pays attention to one alternative and then searches all alternatives connected to it. The agent then focuses on the optimal alternative among those she has considered and takes it as a new starting point to continue this process. The agent stops and selects the alternative that she currently focuses on if no better alternative is connected to it. MN13 interprets this choice procedure as either a physical searching process or a mental association process. While our paper shares similar motivation with that of MN13, the two models have different emphases: MN13 studies how the initial focal point affects the final choice of the DM through mental

[3] One of the most influential random utility models is the Luce's model (Luce, 1959), which is closely related to the widely adopted logit (Strom, 1965) and nested logit (McFadden, 1974) models in structural estimations. See also Kovach and Tserenjigmid (2022) for the behavioral foundations of these models.

association or searching, while our paper explores the role of mental association in shaping the DM's choices when the feasible choice set may not fully capture the DM's attention.

Finally, our work is conceptually connected to the literature on social networks and network games (Jackson, 2008; Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv, 2010; Goyal, 2023). In this literature, individuals are represented by the nodes on the network, and their interactions with each other are captured by the links. We adopt the concept of an associative network, where each node represents an alternative, and the links characterize the DM's mental association. Our research suggests the potential of utilizing the methods and tools developed in the network literature to investigate individual choices.

2 Preliminaries

We consider a nonempty and finite set X of alternatives, with generic alternatives denoted by x, y, z, etc. Denote by $\mathcal{M}$ the collection of all subsets of X. Elements of $\mathcal{M}$ are referred to as menus and are denoted by A, B, D, etc. When there is no confusion, we write x for the singleton menu $\{x\}$, and omit the set union mark by writing AB for $A \cup B$ and Ax for $A \cup x$.

An extended menu is a pair $(A, S) \in \mathcal{M} \times \mathcal{M}$ with $A \cap S = \emptyset$. We interpret AS as the set of perceivable alternatives, whereby A contains the feasible ones, and S includes the infeasible ones. The DM can pay attention to any alternatives in AS but can only make a choice from A or choose the default option. We denote by $\mathcal{S}$ the collection of all extended menus.

Random choice rule. A random choice rule is a map $\rho : X \times \mathcal{S} \to [0, 1]$ such that for all $(A, S) \in \mathcal{S}$, (1) $|A| \neq 0$ implies $\sum_{x \in A} \rho(x, A, S) \in (0, 1)$, and (2) $\rho(x, A, S) > 0$ implies $x \in A$. We write $\rho(x|A, S)$ for $\rho(x, A, S)$ in the rest part of the paper, and define

$$\Phi(A, S) := 1 - \sum_{x \in A} \rho(x|A, S).$$

To interpret, $\rho(x|A, S)$ is the probability that the DM chooses alternative x in extended menu (A, S), and $\Phi(A, S)$ is the probability that the DM chooses the default option (e.g., walking away from the shop, abstaining from voting).[4]

We note that condition (1) states that if A is not empty, then the DM chooses alternatives in A with a probability strictly between 0 and 1. That is, the DM has a positive probability to choose the default option. When there are no feasible alternatives, i.e., the set A is empty, we have $\Phi(\emptyset, S) = 1$. For any extended menu (A, S), we denote by $c(A, S)$ the set of alternatives in A that are chosen with positive probabilities, i.e., $c(A, S) = \{x \in A : \rho(x|A, S) > 0\}$. We refer to alternatives in $c(A, S)$ as alternatives chosen in (A, S).

Preference. A preference ordering $\succ$ is a total order defined on X. We use $\max(A; \succ)$ to denote the $\succ$-maximal alternative in A whenever A is not empty.

3 Limited Attention with Associative Network

We model a DM who forms her consideration set through mental association. The DM has a random initial consideration set and enlarges her consideration set by associating other perceivable alternatives with her considered ones, resulting in an enlarged consideration set that we refer to as her final consideration set. The DM then chooses the best feasible alternative (according to her preference ordering) in her final consideration set (and chooses the default option if each alternative in her final consideration set is not feasible). Below, we formalize this procedure.

Initial consideration set. Following MM14, we assume that each alternative has a fixed probability to be initially considered by the DM, and that the DM attributes her attention to each alternative independently. The attention probability of each alternative is given by an attention probability function $\pi : X \to (0, 1)$. For a given π, we define $\mathring{\pi} : X \to (0, 1)$ such that for all $x \in X$, $\mathring{\pi}(x) = 1 - \pi(x)$. To interpret, $\mathring{\pi}(x)$ is the probability that the DM does not pay attention to x initially. To simplify the notation, we write π_x for $\pi(x)$, $\mathring{\pi}_x$ for $\mathring{\pi}(x)$, π_A for $\prod_{x\in A} \pi(x)$, and $\mathring{\pi}_A$ for $\prod_{x\in A} \mathring{\pi}(x)$.

In a given extended menu (A, S), the DM can only pay attention to alternatives in AS. Thus, the DM's initial consideration set is a subset of AS. In particular,

[4]For recent work on allowing "not choosing" in a random choice setting, see MM14, Brady and Rehbeck (2016) and Dardanoni, Manzini, Mariotti, and Tyson (2020).

the probability for a subset B of AS to be the DM's initial consideration set is given by

$$\pi_B \mathring{\pi}_{(AS)\setminus B}.$$

The above formula assumes that the probability of the DM paying attention to each perceivable alternative remains constant regardless of whether it is feasible or not. This assumption is appropriate in situations where the feasibility of alternatives does not impact their ability to catch the DM's attention. For example, a consumer may randomly notice some products displayed in a store, whether or not they are in stock, or a food enthusiast may randomly pay attention to menu items at a restaurant, regardless of whether they are available. In Section 6, we discuss the case where this assumption is relaxed.

Associative network and the final consideration set. After forming her initial consideration set, the DM can expand it through mental association. This is facilitated by an associative network, which is a directed graph $(X, \mathcal{G})$, where $\mathcal{G} \subseteq X \times X$ is a binary relation that satisfies $\mathcal{X} \subseteq \mathcal{G}$, where $\mathcal{X} = \{(x, x) : x \in X\}$. If $(x, y) \in \mathcal{G}$, then y is **directly associated with** x, which particularly means in this paper that the attention to x will prompt the DM to further consider y. In the remaining part of the paper, we omit X and call the binary relation $\mathcal{G}$ an associative network.

The mental association process enables the DM to consider more alternatives based on what she already considers. In a given extended menu (A, S), the DM's mental association process is only relevant with the restricted associative network $\mathcal{G}_{AS}$ on AS, where $\mathcal{G}_{AS} = \{(x, y) \in \mathcal{G} : x, y \in AS\}$.

With $\mathcal{G}_{AS}$, the DM's mental association process works as follows. For each alternative $x \in AS$ that she initially considers, she includes every alternative y that satisfies $(x, y) \in \mathcal{G}_{AS}$ into her consideration set. For each such alternative y, she then expands her consideration set by including each alternative z that satisfies $(y, z) \in \mathcal{G}_{AS}$. The association process terminates when the DM cannot associate other alternatives in AS with what she already considers.

Formally, let $\bar{\mathcal{G}}_{AS} \subseteq X \times X$ be the transitive closure of $\mathcal{G}_{AS}$. That is, $\bar{\mathcal{G}}_{AS}$ is the smallest superset of $\mathcal{G}_{AS}$ satisfying the property that for all $(x, y), (y, z) \in \bar{\mathcal{G}}_{AS}$, $(x, z) \in \bar{\mathcal{G}}_{AS}$. In other words, $(x, y) \in \bar{\mathcal{G}}_{AS}$ if and only if there is a path from x to y

in the restricted graph $(AS, \mathcal{G}_{AS})$. Here, a path refers to a sequence of alternatives $x_1, ..., x_n \in AS$ where $x_1 = x$ and $x_n = y$, and for all $k \in \{1, ..., n-1\}$, $(x_k, x_{k+1}) \in \mathcal{G}_{AS}$. For a given extended menu (A, S), we say that y is **associated with** x in (A, S) if $(x, y) \in \bar{\mathcal{G}}_{AS}$.[5] To interpret, if $(x, y) \in \bar{\mathcal{G}}_{AS}$, then the consideration of x leads to the consideration of y through an association path in $\mathcal{G}_{AS}$. Thus, whenever the DM initially considers $x \in AS$, the set

$$\bar{\mathcal{G}}_{AS}(x) := \{y : (x, y) \in \bar{\mathcal{G}}_{AS}\}$$

is included in her final consideration set. When the initial consideration set of the DM is $B \subseteq AS$, her final consideration set is given by

$$\bar{\mathcal{G}}_{AS}(B) := \bigcup_{x \in B} \bar{\mathcal{G}}_{AS}(x).$$

We illustrate the mental association process described above in Figure 1.

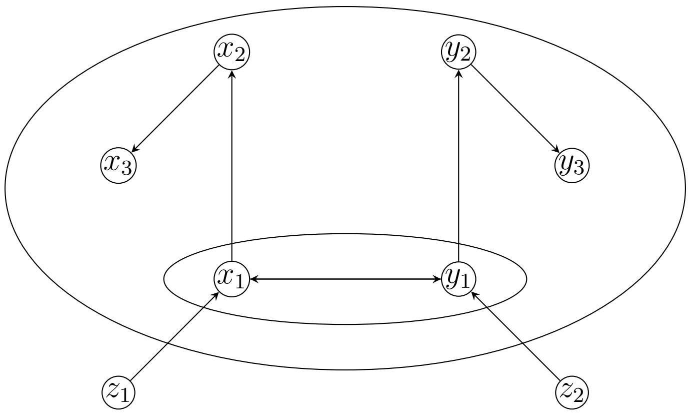

Figure 1: The extended menu contains 8 alternatives. The associative network is given by the arrows. The initial consideration set of the DM is $\{x_1, y_1\}$, and after association, her final consideration set is $\{x_1, x_2, x_3, y_1, y_2, y_3\}$. Alternatives z_1 and z_2 are not considered, as they are not associated with any considered alternative.

One implicit assumption we make is that the association process is only relevant to whether an alternative is perceivable or not, and is independent of the feasibility of the alternative. We justify this assumption by noting that in many applications, the DM first forms her consideration set and then checks the availability of each alternative. For example, in an offline store, the consumer may first consider

[5]When the extended menu (A, S) is clear, we simply say that y is associated with x.

multiple items and then check their feasibility with the seller one by one. Similarly, in an online retailing platform, the consumer may add all items she wants to compare to her shopping cart, and those that are not feasible would be automatically labeled by the platform system at the final checkout page.

Another important feature of our model is that the DM's association process depends on infeasible yet perceivable alternatives. To illustrate, consider two extended menus $(\{z\}, \{x\})$ and (z, xy),[6] and assume that y is directly associated with x and z is directly associated with y, but z is not directly associated with x. If the initial consideration set of the DM is x, then she cannot further consider z in extended menu $(\{z\}, \{x\})$, but can do so through an intermediate alternative y in extended menu (z, xy). This feature is relevant in many situations. For instance, a DM can directly associate smoothies with milk and can be cued by the picture of pineapple smoothies to further consider mango juice, but she cannot directly associate mango juice with milk without the cue of pineapple smoothies.

We note that our model can also accommodate the case where the DM is able to store the associated alternative y in her memory and use it for further association. By doing so, she can directly associate z with x even in extended menu $(\{z\}, \{x\})$. This seems to be incompatible with our model as an alternative that is not perceivable serves as the intermediate alternative in the process of the DM's mental association. In fact, our model can accommodate such case: If the DM can associate z with x through y without y being perceivable, then it is as if the DM can directly associate z with x. The two interpretations lead to the same choice behavior of the DM, and therefore, we do not distinguish between them in our choice model.

Preference and choice. After forming the final consideration set, the DM chooses the best alternative from the set if it contains at least one feasible alternative. Otherwise, she chooses the default option. Specifically, in extended menu (A, S), if the DM's final consideration set is $B \subseteq AS$, then she chooses $\max(A \cap B; \succ)$, where $\succ$ is the DM's preference.[7] The definition of our choice rule is then given as follows.

[6]We write $(\{z\}, \{x\})$ instead of (z, x) to denote the extended menu since the latter may also denote a link in the associative network, which may create some confusion.

[7]Note that if $A \cap B = \emptyset$, then $\max(A \cap B; \succ) = \emptyset$.

Definition 1. *A random choice rule ρ is a random consideration and association rule (RCAR) if there exists a tuple $(\pi, \mathcal{G}, \succ)$, where $\pi : X \to (0,1)$ is an attention probability function, $\mathcal{G}$ is an associative network, and $\succ$ is a preference ordering, such that for all $x \in X$ and $(A, S) \in \mathcal{S}$ with $x \in A$:*

$$\rho(x|A,S) = \sum_{B \subseteq AS : x = \max(\bar{\mathcal{G}}_{AS}(B);\succ)} \pi_B \mathring{\pi}_{(AS)\setminus B}.$$

In the above case, we say that $(\pi, \mathcal{G}, \succ)$ represents ρ as an RCAR.

In words, with an RCAR, the choice probability of x in extended menu (A, S) is the frequency with which there exists some initial consideration set such that x is the best alternative in the corresponding final consideration set.

4 Axioms and Representation Theorem

In this section, we first introduce a reformulation of our choice model, which highlights some of its key properties. Then we present the axioms and representation theorem, and discuss the comparative statics of our model.

4.1 Reformulation

An important feature of our model that distinguishes it from other random choice models is that not every feasible alternative is chosen with a positive probability. To see this, consider an RCAR ρ that is represented by $(\pi, \mathcal{G}, \succ)$ and an extended menu $(xy, \emptyset)$. Assume $x \succ y$ and $(y, x) \in \mathcal{G}$, meaning that x is better than y, and x is directly associated with y. In this extended menu, y can never be chosen, as whenever y is considered, the DM will further consider x, which in turn blocks the choice of y. The following proposition provides a characterization of the set of chosen alternatives in a given extended menu.

Proposition 1. *Consider an RCAR ρ that is represented by $(\pi, \mathcal{G}, \succ)$. For all $(A, S) \in \mathcal{S}$, $x \in c(A, S)$ if and only if $x = \max(\bar{\mathcal{G}}_{AS}(x) \cap A; \succ)$.*

According to Proposition 1, an alternative x is chosen in extended menu (A, S) if and only if there is no association path from x to any feasible alternative that is strictly better than it. If some feasible alternative y that is better than x is associated with x through some intermediate alternatives in AS, then the attention to x always leads to the attention to y, and thus the choice of x is blocked by y.

To proceed, we investigate the choice frequency of each chosen alternative. Consider an extended menu (A, S) and an alternative $x \in c(A, S)$. We define

$$\Gamma_{\mathcal{G}}^{x,\succ}(A, S) := \{y \in AS : x = \max(\bar{\mathcal{G}}_{AS}(y) \cap A; \succ)\}.$$

The consideration of alternatives in $\Gamma_{\mathcal{G}}^{x,\succ}(A, S)$ lead to final consideration sets in which x is $\succ$-best. Additionally, we define

$$H_{\mathcal{G}}(A, S) := \{x \in AS : \bar{\mathcal{G}}_{A,S}(x) \cap A \neq \emptyset\},$$

which is the set of alternatives in extended menu (A, S) with which some feasible alternative is associated. Clearly, we have $A \subseteq H_{\mathcal{G}}(A, S)$. With these notations, we reformulate our choice rule as follows.

Proposition 2. *Consider an RCAR ρ represented by $(\pi, \mathcal{G}, \succ)$. For all $(A, S) \in \mathcal{S}$,*

$$H_{\mathcal{G}}(A, S) = \bigcup_{x \in c(A,S)} \Gamma_{\mathcal{G}}^{x,\succ}(A, S), \tag{1}$$

$$\Phi(A, S) = \mathring{\pi}_{H_{\mathcal{G}}(A,S)}, \ \ \textit{and} \tag{2}$$

if $c(A, S) = \{x_1, ..., x_n\}$ with $x_1 \succ ... \succ x_n$, then

$$\begin{aligned} \rho(x_1|A, S) = 1 - \mathring{\pi}_{B_1}, \ \ \textit{and} \\ \forall k \in \{2, ..., n\}, \ \ \rho(x_k|A, S) = (1 - \mathring{\pi}_{B_k})\, \mathring{\pi}_{\cup_{t=1}^{k-1} B_t}, \end{aligned} \tag{3}$$

where for all $k \in \{1, ..., n\}$, $B_k = \Gamma_{\mathcal{G}}^{x_k,\succ}(A, S)$.

The interpretation of condition (1) is straightforward: Each alternative that leads to the consideration of some alternative in A must lead to the choice of some feasible alternative. Condition (2) is based on condition (1): The probability for the default option to be chosen is equal to the probability that the initially considered alternatives do not lead to the consideration of any feasible alternative.

Condition (3) is a reformulation of our choice rule: The probability for an alternative to be chosen is equal to the probability that it is finally considered through the association process while all better feasible alternatives are not finally considered.

4.2 Characterization

The first axiom stipulates the independent attention distribution for the DM.

Axiom 1—Default Independence: For all $x \in X$ and $A, B \in \mathcal{M}$ with $x \in A \cap B$:

$$\frac{\Phi(A, \emptyset)}{\Phi(A \backslash x, \emptyset)} = \frac{\Phi(B, \emptyset)}{\Phi(B \backslash x, \emptyset)}.$$

Axiom 1 follows from the I-Independence axiom of MM14.[8] When all perceivable alternatives are feasible, the DM's choice of the default option depends on whether her initial consideration set is empty or not. The effect of removing a feasible alternative on the frequency of choosing the default option is determined by the extent to which it attracts the DM's attention. Axiom 1 posits that the DM has a constant probability of attending to a particular alternative, regardless of the choice set.

Axiom 2—Feasibility Monotonicity: For all $x, y \in X$ and $(A, S) \in \mathcal{S}$ with $y \in A$ and $x \notin AS$, $\rho(y|Ax, S) \leq \rho(y|A, Sx)$, $\Phi(Ax, S) \leq \Phi(A, Sx)$, and in particular, $\Phi(Ax, S) < \Phi(A, Sx)$ implies $\Phi(A, Sx) = \Phi(A, S)$.

Axiom 2 states that if a feasible alternative x becomes infeasible but remains perceivable, the choice probabilities of other alternatives and the default option weakly increase. In particular, if the probability of choosing the default option is strictly increased after x becomes infeasible, then removing x from the extended menu will have no additional effect on the choice frequency of the default option.

The first part of Axiom 2 can be interpreted as the standard monotonicity condition of random choice rules. For the second part, consider an alternative x such that the consideration of x leads to the consideration of some other feasible

[8]The I-Independence axiom requires more than Axiom 1 does. It additionally requires that for all $x, y \in X$ and $A, B \in \mathcal{M}$ with $x, y \in A \cap B$ and $x \neq y$, $\frac{\rho(x|A \backslash y, \emptyset)}{\rho(x|A, \emptyset)} = \frac{\rho(x|B \backslash y, \emptyset)}{\rho(x|B, \emptyset)}$.

alternative in A. In this case, the feasibility of x does not affect the choice frequency of the default option: No matter whether x is feasible or not, it will lead to the choice of some feasible alternative and block the choice of the default option. On the other hand, if the consideration of x does not lead to the consideration of any other feasible alternative in A, then x blocks the choice of the default option only when it is feasible. Following this logic, when the feasibility of x has an impact on the choice frequency of the default option (i.e., $\Phi(Ax, S) < \Phi(A, Sx)$), it can be inferred that no other feasible alternative in A is associated with x. Therefore, when x is infeasible but perceivable, removing x will not further alter the choice frequency of the default option.

To state the next two axioms, we need some additional notations. For any $x \in X$ and $(A, S) \in \mathcal{S}$, we say that some alternative in A is revealed to be associated with x through S, denoted by $x \xrightarrow{S} A$, if either $\Phi(A, S) \neq \Phi(A, S\backslash x)$ or $x \in A$. If $x \in A$, then the alternative x in A is associated with x. If $\Phi(A, S) \neq \Phi(A, S\backslash x)$, by our previous analysis, there must be some alternative in A that is associated with x, since otherwise, removing x from the extended menu will not change the choice frequency of the default option.

Axiom 3—Expansion: For all $x \in X$ and $(A, S), (B, T) \in \mathcal{S}$, if $B \subseteq A$ and $BT \subseteq AS$, then $x \xrightarrow{T} B$ implies $x \xrightarrow{S} A$.

Axiom 3 states that if a DM can associate some feasible alternative with x in a given extended menu, she can also do so in a less restrictive extended menu. In other words, when there are more intermediate cues in the form of perceivable alternatives, more alternatives become associated with x.

Axiom 4—Path Connectedness: For all $x, y \in X$ and $(A, S) \in \mathcal{S}$ with $x \neq y$, if $x \xrightarrow{S} A$ and not $x \xrightarrow{S\backslash y} A\backslash y$, then $x \xrightarrow{S\backslash y} y$ and $y \xrightarrow{S\backslash x} A$.

Axiom 4 captures the key feature of the associative network: One alternative is associated with another through paths in the network. To see this, note that if the DM can associate some alternative in A with x through S but cannot do so when y is not perceivable, then y must be an intermediate alternative for this association process. Hence, the DM must associate y with x and associate some alternative in A with y.

Axiom 5—Association Asymmetry: For all $x, y \in X$ and $(A, S), (B, T) \in \mathcal{S}$ with $x \neq y$, $x \in c(A, S)$ and $y \in c(B, T)$:

$$\rho(y|A, S) \neq \rho(y|A\backslash x, S) \Rightarrow \rho(x|B, T) = \rho(x|B\backslash y, Ty) = \rho(x|B\backslash y, T).$$

Axiom 5 is similar to the I-Asymmetry axiom of MM14.[9] It says that if deleting a chosen alternative x in a given extended menu changes the choice frequency of another alternative y, then in a extended menu where y is chosen, neither the feasibility nor the perceivability of y affects the choice frequency of x.

To interpret Axiom 5, consider two alternatives x and y. If deleting x from an extended menu changes the choice frequency of y, then either the consideration of x leads to the consideration of y and the presence of x increases the likelihood of y, meaning that x cannot be better than y, or the presence of x hinders the choice of y, indicating that x is better than y. Moreover, if x is chosen in the extended menu, then the former case cannot happen and x must be better than y. Therefore, deleting y from an extended menu in which it is chosen does not alter the choice frequency of x, otherwise the same argument would imply that y is better than x, leading to a contradiction. Together, these five axioms fully characterize our random choice rule.

Theorem 1. *Axioms 1-5 are sufficient and necessary for a random choice rule ρ to be an RCAR. If ρ can be represented by both $(\pi_1, \mathcal{G}_1, \succ_1)$ and $(\pi_2, \mathcal{G}_2, \succ_2)$ as an RCAR, then $\pi_1 = \pi_2$, $\mathcal{G}_1 = \mathcal{G}_2$, and $\succ_1 = \succ_2$.*

Identification of the parameters. The attention probability π_x of a given alternative x can be identified by examining its choice frequency in extended menu $(x, \emptyset)$. Specifically, we have $\pi_x = \rho(x|x, \emptyset)$, i.e., x is chosen in $(x, \emptyset)$ when it attracts the attention of the DM.

The preference ordering $\succ$ can be identified through extended menus taking the form of $(xy, \emptyset)$. Without loss of generality, suppose that x is better than y. There are two cases to be considered: Either x is associated with y and blocks the choice of y, or x is not associated with y and thus y is chosen with probability

[9] The I-Asymmetry axiom states that for all distinct $x, y \in X$ and $A, B \in \mathcal{M}$, if $\rho(y|A, \emptyset) \neq \rho(y|A\backslash x, \emptyset)$, then $\rho(x|B, \emptyset) = \rho(x|B\backslash y, \emptyset)$.

$\mathring{\pi}_x\pi_y$. In both cases, deleting x from extended menu $(xy, \emptyset)$ strictly boosts the choice frequency of y to π_y. On the other hand, one can verify that deleting y will not boost the choice frequency of x. Therefore, the preference ranking $x \succ y$ is identified.

The associative network $\mathcal{G}$ can be identified through extended menus taking the form of $(\{x\}, \{y\})$. If $\rho(x|x, y) \neq \rho(x|x, \emptyset)$, i.e., removing an infeasible but perceivable option y changes the choice frequency of x, then x is identified to be associated with y. By our construction, the attention probability function π, the preference ordering $\succ$, and the associative network $\mathcal{G}$ are all unique.

Sketch of the sufficiency part of the proof of Theorem 1. With the identified parameters $(\pi, \mathcal{G}, \succ)$, we briefly discuss how our axioms lead to the representation of RCARs. In step 1, we show that the binary relation $\succ$ is well-defined for each distinct pair of alternatives and satisfies asymmetry and transitivity.[10] Thus, $\succ$ is indeed a preference ordering. In addition, we show that if $x \succ y$, then in any extended menu where y is chosen with a positive probability, deleting y from the extended menu will not affect the choice frequency of x. This observation follows from the Association Asymmetry axiom.

In step 2, we show that for all extended menu (A, S) and alternative $x \in S$, $\Phi(A, S) \neq \Phi(A, S \backslash x)$ if and only if some alternative in A is associated with x, i.e., $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$. With this observation, we can focus on a subset $\hat{S} \subseteq S$ in extended menu (A, S) such that $\hat{S}$ contains all the alternatives with which some alternative in A is associated. By the Feasibility Monotonicity axiom, we can show that $\Phi(A, S) = \Phi(A, \hat{S}) = \Phi(A\hat{S}, \emptyset)$.

In step 3, we show that for all extended menu (A, S), an alternative $x \in A$ is chosen with positive probability if and only if there is no better feasible alternative in (A, S) that is associated with x.

The final step is to show that each alternative is chosen with the probability specified by our model. To illustrate, we provide a simple example. Consider an

[10] The binary relation $\succ$ is asymmetric if for all $x, y \in X$, $x \succ y$ implies not $y \succ x$, and is transitive if for all $x, y, z \in X$, $x \succ y$ and $y \succ z$ imply $x \succ z$.

extended menu $(xyzw, r)$ such that

$$
\begin{aligned}
&x \succ y \succ z \succ w \succ r, \\
&\mathcal{G}_{xyzwr} = \mathcal{X} \cup \{(w, y), (r, z)\}.
\end{aligned}
$$

By step 3, w is not chosen in $(xyzw, r)$ as its consideration leads to the consideration of a better alternative y, and x, y, and z are chosen with positive probabilities. Thus we need to show

$$
\begin{aligned}
&\rho(x|xyzw, r) = \pi_x, \;\; \rho(y|xyzw, r) = \mathring{\pi}_x(1 - \mathring{\pi}_{yw}), \\
&\rho(z|xyzw, r) = \mathring{\pi}_{xyw}(1 - \mathring{\pi}_{zr}).
\end{aligned}
\tag{4}
$$

The key insight of our proof is that once we obtain the following probabilities

$$
\begin{aligned}
&\rho(x|xyzw, r) + \rho(y|xyzw, r) + \rho(z|xyzw, r), \\
&\rho(x|xyzw, r) + \rho(y|xyzw, r), \\
&\rho(x|xyzw, r),
\end{aligned}
$$

we can then derive the choice probabilities of x, y and z in $(xyzw, r)$. To obtain these probabilities, note that by step 2, we have $\Phi(xyzw, r) = \Phi(xyzwr, \emptyset)$, since z is associated with r. Hence,

$$
\begin{aligned}
1 - \mathring{\pi}_{xyzwr} &= 1 - \Phi(xyzwr, \emptyset) = 1 - \Phi(xyzw, r) \\
&= \rho(x|xyzw, r) + \rho(y|xyzw, r) + \rho(z|xyzw, r).
\end{aligned}
\tag{5}
$$

To proceed, we delete z from the extended menu. By step 1, we have $\rho(x|xyzw, r) = \rho(x|xyw, r)$ and $\rho(y|xyzw, r) = \rho(x|xyw, r)$, since z is chosen with positive probability in $(xyzw, r)$ and is worse than x and y. Again by step 3, we know that in the new extended menu (xyw, r), only x and y are chosen with positive probabilities (w is not chosen since y is associated with w). Since no alternative in xyw is associated with r, by step 2, deleting r will not affect the choice frequency of the default option, i.e., $\Phi(xyw, r) = \Phi(xyw, \emptyset)$. Hence,

$$
\begin{aligned}
1 - \mathring{\pi}_{xyw} &= 1 - \Phi(xyw, \emptyset) = 1 - \Phi(xyw, r) \\
&= \rho(x|xyw, r) + \rho(y|xyw, r) \\
&= \rho(x|xyzw, r) + \rho(y|xyzw, r).
\end{aligned} \tag{6}
$$

Next, we delete y from the extended menu. By step 1, we have $\rho(x|xyzw, r) = \rho(x|xyw, r) = \rho(x|xw, r)$. Note that now w is also chosen with a positive probability in (xw, r) since x is not associated with it. Therefore, we can continue to delete w from the extended menu while maintaining the choice probability of x, i.e., $\rho(x|xw, r) = \rho(x|x, r)$. By a similar argument, we have

$$
1 - \mathring{\pi}_x = 1 - \Phi(x, \emptyset) = 1 - \Phi(x, r) = \rho(x|x, r) = \rho(x|xyzw, r). \tag{7}
$$

Combining equations (5), (6), and (7), we obtain equation (4), and the desired choice probabilities are obtained.

Comparative Statics. We end this section by discussing some comparative statics of our model. We show that we can directly compare two DMs' associative networks without imposing any restriction on the alignment of their preferences or attention probabilities. The following proposition directly follows from the construction of $\mathcal{G}$, and its proof is omitted.

Proposition 3. *For any two RCARs ρ_1 and ρ_2 that are represented respectively by $(\pi_1, \mathcal{G}_1, \succ_1)$ and $(\pi_2, \mathcal{G}_2, \succ_2)$, with $\Phi_1(\cdot, \cdot)$ and $\Phi_2(\cdot, \cdot)$ denoting the default option's choice probabilities under the two rules respectively, the following statements are equivalent:*

(1) *For all $x \in X$ and $(A, S) \in \mathcal{S}$ with $x \notin AS$, if $\Phi_1(A, S) \neq \Phi_1(A, Sx)$, then $\Phi_2(A, S) \neq \Phi_2(A, Sx)$.*

(2) *The associative network $\mathcal{G}_1$ is a subset of $\mathcal{G}_2$.*

Condition (1) of Proposition 3 states that if DM1 associates some alternative in A with x through S, then DM2 also does so. Clearly, this happens only when DM2 conducts more mental association when making decisions, i.e., $\mathcal{G}_1 \subseteq \mathcal{G}_2$.

5 All perceivable Alternatives Are Feasible

In many applications, every perceivable alternative is feasible. For instance, in an offline book store, each book on the shelf is for sale, and in a grocery store, every item on the shelves is available for purchase. In these situations, each extended menu takes the form of $(A, \emptyset)$ for some $A \in \mathcal{M}$, and this is also the standard choice setting adopted in most choice-theoretical frameworks. In this section, we provide axioms that characterize RCARs when the DM's choices are restricted on this smaller domain.

We consider a proper subset $\mathcal{F} \subsetneq \mathcal{S}$ of extended menus where $\mathcal{F} = \{(A, \emptyset) : A \in \mathcal{M}\}$. In each extended menu in $\mathcal{F}$, every perceivable alternative is feasible. The choice data take the form of a random choice rule ρ restricted on $\mathcal{F}$. We call such a choice rule a restricted random choice rule. The first two axioms characterize alternatives that are chosen with positive probabilities.

Axiom 6—Sen's α: For all $A, B \in \mathcal{M}$, $B \subseteq A$ implies $c(A, \emptyset) \cap B \subseteq c(B, \emptyset)$.

Axiom 7—Reducibility: For all $A \in \mathcal{M}$, if for all $x \in A$, $c(A, \emptyset) \neq c(A \backslash x, \emptyset)$, then $c(A, \emptyset) = A$.

Axiom 6 states that if an alternative is selected from a larger menu, it must also be selected from any smaller menu that contains it. This axiom can be interpreted to mean that if a particular alternative, denoted as x, is not chosen in a smaller menu, then given the presence of more competitive alternatives in a larger menu, it should also remain unselected. In our context, if an alternative is not chosen, its consideration must lead to the consideration of a better alternative. Consequently, in a larger menu, the superior alternative remains to be associated with x and thus blocks the choice of x.

The contrapositive of Axiom 7 states that if not all alternatives are selected, then there exists an unselected alternative whose removal does not alter the set of chosen alternatives. To illustrate the axiom, consider an extended menu $(xyz, \emptyset)$ where only x is chosen. As both y and z are unselected, their consideration must lead to the consideration of a superior alternative in this extended menu, which has to be x. If the removal of y results in a change in the set of chosen alternatives such that z becomes chosen, then x must be associated with z through y, and x

must be directly associated with y. In this scenario, the deletion of z does not alter the association relation between x and y, and thus does not affect the set of chosen alternatives. In summary, Axiom 7 establishes the existence of an unselected alternative (if not all alternatives are chosen) whose removal does not impact the association relation among the remaining alternatives, thereby preserving the set of chosen alternatives.

We note that both Axioms 6 and 7 are weakening of the Weak Axiom of Revealed Preference (WARP).[11] This axiom characterizes the rational choice model for which there is a complete and transitive binary relation $\succsim^*$ over X such that for each $(A, \emptyset) \in \mathcal{F}$, $c(A, \emptyset)$ contains all alternatives in A that maximize $\succsim^*$.[12] It can be shown that WARP implies the Sen's α axiom and a stronger version of the Reducibility axiom which states that for all $A \in \mathcal{M}$, if $x \notin c(A, \emptyset)$, then $c(A \backslash x, \emptyset) = c(A, \emptyset)$. In fact, for any complete and binary relation $\succsim^*$, we can find an RACR ρ such that its support c satisfies that for all $A \in \mathcal{M}$, $c(A, \emptyset) = \max(A; \succsim^*)$.[13] Thus, our random choice rule generalizes the rational choice model in terms of the support of choices.

For an alternative $x \in X$ and a menu A, we say that x is independent with A, denoted by $x \vdash A$, if $x \notin A$ and for all $y \in A$, $\rho(y|A, \emptyset) = \rho(y|Ax, \emptyset)$. Note that according to this definition, for all $x \in X$, we have $x \vdash \emptyset$.

Axiom 8—Weak I-Independence: For all $x \in X$ and $A, B \in \mathcal{M}$, if $x \vdash A$ and $x \vdash B$, then $x \vdash A \cup B$.

Axiom 8 posits that if x is independent with both A and B, then it is also independent with their union. It is noteworthy that this axiom is a less stringent version of the I-Independence axiom of MM14. According to the I-Independence axiom, if x does not affect the frequency of selecting alternative y in a particular menu, then it should not impact the frequency of choosing y in any menu. In

[11] WARP states that for all $x, y \in X$ and $A, B \in \mathcal{M}$ with $x, y \in A \cap B$, if $x \in c(A, \emptyset)$ and $y \in c(B, \emptyset)$, then $x \in c(B, \emptyset)$.

[12] A binary relation $\succsim^*$ is complete if for all $x, y \in X$, either $x \succsim^* y$ or $y \succsim^* x$.

[13] To see this, consider a complete and transitive binary relation $\succsim^*$. We construct the RCAR as follows. Let ρ be represented by $(\pi, \mathcal{G}, \succ)$, where π is arbitrary, $\succ$ extends the asymmetric part $\succ^*$ of $\succsim^*$, and $(x, y) \in \mathcal{G}$ if and only if $x = y$ or $y \succ^* x$. One can verify the choice support c of this RCAR coincides with the $\succsim^*$-maximal alternatives in each extended menu in $\mathcal{F}$.

contrast, Axiom 8 necessitates that x has no impact on the choice frequency of **every** alternative in the given menu.

For any two alternatives x and y, we say that x weakly dominates y, denoted by $x \trianglerighteq y$, if there is a menu A such that $y \in A$ and $c(A, \emptyset) = x$.

Axiom 9—Dominance Asymmetry: For all $x, y, z \in X$ and $A, B \in \mathcal{M}$ with $x \neq z$, $x \trianglerighteq y$, $y \in c(A, \emptyset)$, and $z \in c(B, \emptyset)$:

$$\rho(z|A, \emptyset) \neq \rho(z|A \backslash y, \emptyset) \Rightarrow \rho(x|B, \emptyset) = \rho(x|B \backslash z, \emptyset).$$

Axiom 9 can be interpreted similarly as Axiom 5. Based on our discussion of Axiom 5, Axiom 9 states that if the choice of z is hindered by a chosen alternative y that is weakly dominated by x, then z cannot hinder the choice of x provided with that z is chosen with a non-trivial probability. The new axioms characterize RCARs in the restricted choice domain.

Theorem 2. *Axioms 1 and 6-9 are sufficient and necessary for a restricted random choice rule ρ to be an RCAR restricted on $\mathcal{F}$. If ρ is represented by both $(\pi_1, \mathcal{G}_1, \succ_1)$ and $(\pi_2, \mathcal{G}_2, \succ_2)$ as an RCAR on $\mathcal{F}$, then $\pi_1 = \pi_2$ and $\succ_1 = \succ_2$.*

Identification of the associative network. Recall that in Section 4.2, how we identify the DM's preference ordering and attention probability function relies only on extended menus of the form $(A, \emptyset)$. Hence, even in the restricted choice domain, we can still uniquely identify the DM's preference ordering $\succ$ and attention probability function π. What differs in the restricted choice domain is the identification of the DM's associative network. In what follows, we provide two identification strategies.

Example 1. Consider two distinct alternatives x and y, and assume $c(xy, \emptyset) = y$. We claim that any tuple $(\pi, \mathcal{G}, \succ)$ that represents ρ as an RCAC on $\mathcal{F}$ necessarily satisfies $(x, y) \in \mathcal{G}$: Since x is not chosen in xy, the consideration of x must prompt the consideration of some better alternative, which has to be y. $\square$

The argument in Example 1 leads to a primitive association relation identified through binary choices. The following example generalizes this observation.

Example 2. Consider four alternatives x, y, z, and w, and assume that

(1) $c(xy, \emptyset) = c(xyzw, \emptyset) = x$, and

(2) $c(xyw, \emptyset) = xw$.

The observation $c(xyzw, \emptyset) = x$ implies that x is better than the other three alternatives. By $c(\{xy, \emptyset\}) = x$ and $c(xyw, \emptyset) = xw$, one can infer that x is associated with y, and neither y nor x is associated with w, since otherwise the attention to w would prompt the consideration of x and thus blocks the choice of w. Now, by adding z to menu $(xyw, \emptyset)$, w becomes unchosen. Hence, z must be associated with w, i.e., for any tuple $(\pi, \mathcal{G}, \succ)$ that represents ρ as an RCAC on $\mathcal{F}$, it necessarily satisfies that $(w, z) \in \mathcal{G}$. $\square$

The above examples suggest the following identification of the associative network. For a given restricted choice rule ρ, define

$$\mathcal{G}^c := \mathcal{X} \cup \{(x, y) \in X^2 \backslash \mathcal{X} : c(xy, \emptyset) = y\} \cup$$
$$\{(x, y) \in X^2 : \exists A \subseteq X \text{ and } z \neq x \text{ with } c(A, \emptyset) = c(Axy, \emptyset) = z, \text{and } c(Ax, \emptyset) = xz\}.$$

Note that the definition of $\mathcal{G}^c$ only relies on the support c of the choice rule ρ.

An alternative way of identifying the associative network relies on the choice frequencies of the alternatives. For a given restricted choice rule ρ, define

$$\mathcal{G}^\rho := \mathcal{X} \cup \{(x, y) \in X^2 : \exists A \in \mathcal{M} \text{ such that } x \vdash A \text{ and } x \notin c(Axy, \emptyset)\}.$$

In Lemma 8, we demonstrate that if the restricted choice rule ρ is a RCAR on $\mathcal{F}$, then $x \vdash A$ implies that x is chosen in $(Ax, \emptyset)$. Thus, according to the definition of $\mathcal{G}^\rho$, if $x \neq y$, then $(x, y) \in \mathcal{G}^\rho$ if there is a menu A such that x is independent with A, and adding y to Ax results in a change from selecting x to not selecting it. The definition of $\mathcal{G}^\rho$ is based on a similar idea as that of $\mathcal{G}^c$: Since x is independent with A, either (1) A is empty or (2) x is inferior to all selected alternatives in A and no alternative in A is associated with x. When y is added to the menu, x becomes unselected, and this can only be attributed to the fact that the consideration of x prompts the consideration of a superior alternative through the intermediate alternative y.

Our next proposition states that $\mathcal{G}^\rho$ and $\mathcal{G}^c$ are the same if ρ is an RCAR on $\mathcal{F}$, and that both of them are the minimal associative network that represents the choice rule as an RCAR on $\mathcal{F}$.

Proposition 4. *For any random choice rule ρ restricted on $\mathcal{F}$ that is an RCAR on $\mathcal{F}$, the following statements are true.*

(1) *The two associative networks $\mathcal{G}^\rho$ and $\mathcal{G}^c$ are identical.*

(2) *If ρ is represented by $(\pi, \mathcal{G}, \succ)$ as an RCAR on $\mathcal{F}$, then $\mathcal{G}^c \subseteq \mathcal{G}$, and the choice rule is also represented by $(\pi, \mathcal{G}^c, \succ)$ as an RCAR on $\mathcal{F}$.*

To end this section, we show that there could be multiple associative networks representing the same restricted random choice rule as an RCAR on $\mathcal{F}$.

Example 3. Consider $X = \{x, y\}$ and a restricted choice rule ρ such that $\rho(x|x, \emptyset) = \frac{1}{2}$, $\rho(y|y, \emptyset) = \frac{1}{2}$, $\rho(x|xy, \emptyset) = 0$, and $\rho(y|xy, \emptyset) = \frac{3}{4}$. Clearly, it can be revealed that y is directly associated with x, and that y is better than x. However, whether x is directly associated with y does affect the choice frequencies of the alternatives as the presence of y always blocks the choice of x. That is, the DM's associative network can be either $\mathcal{G}^1 = \mathcal{X} \cup (x, y)$ or $\mathcal{G}^2 = \mathcal{X} \cup (x, y) \cup (y, x)$. $\square$

6 Extensions

The focus of this section is to explore extensions of our model.

Feasibility and the attention probability. Our model assumes that the attention probability of a perceivable alternative remains fixed, regardless of its feasibility. However, this assumption may not hold in certain contexts. For example, in some online shopping platforms, products that are sold out are explicitly labeled as "out of stock" on the display page. It is conceivable that such labeling may result in excessive attention from the consumer. Conversely, it is also plausible that the consumer may pay less attention to these alternatives since they are known to be unavailable. Therefore, a natural extension of our model is to incorporate the feasibility of alternatives as a factor that influences the attention probability assigned to them.

Formally, let $\pi^f : X \to (0, 1)$ be the attention probability function for feasible alternatives, and $\pi^n : X \to (0, 1)$ be the attention probability function for infeasible but perceivable alternatives.[14] Let $\mathcal{G}$ be the associative network and $\succ$ be the DM's preference ordering. For each extended menu (A, S) with $A \neq \emptyset$, the choice frequency of $x \in A$ is given by

$$\rho(x|A, S) = \sum_{B \subseteq AS:\ x=\max(\bar{\mathcal{G}}_{AS}(B), \succ)} \pi^f_{B\cap A}\pi^n_{B\cap S}\mathring{\pi}^f_{A\setminus B}\mathring{\pi}^n_{S\setminus B}.$$

We argue that all relevant parameters of the general model above can be uniquely identified. First, the preference relation $\succ$ and associative network $\mathcal{G}$ can be similarly identified as our main model. Second, the attention probability for feasible alternatives π^f can be identified through the choice frequency of each singleton menu, i.e., for every $x \in X$, $\pi^f_x = \rho(x|x, \emptyset)$. Finally, the attention probability for infeasible but perceivable alternatives π^n can be identified through the extent to which they boost the choice frequencies of other alternatives. To see this, consider alternative x and assume that there exists a distinct alternative y such that $(x, y) \in \mathcal{G}$. Clearly, we have $\rho(y|y, x) = 1 - (1 - \pi^n_x)(1 - \pi^f_y)$, which implies

$$\pi^n_x = 1 - \frac{1 - \rho(y|y, x)}{1 - \pi^f_y}.$$

We note that π^n cannot be fully identified: For a given alternative x, if every other alternative is not associated with it, then we are unable to identify π^n_x. Nevertheless, in such case, since the consideration of x does not prompt the consideration of any other alternative, the value of π^n_x does not affect the DM's choice frequencies.

Random associative network. Another natural extension of our model is to consider random networks (Galeotti and Rogers, 2015). Formally, let $\mathscr{G}$ be the set of all possible associative networks over X. A random associative network is a probability distribution μ over $\mathscr{G}$. For a given preference ordering $\succ$ and attention probability function π, the DM's choice frequency of x in extended menu (A, S) under the random associative network is given by

$$\sum_{\mathcal{G}\in\mathscr{G}} \mu(\mathcal{G}) \left(\sum_{B\subseteq AS: x=\max(\bar{\mathcal{G}}_{AS}(B); \succ)} \pi_B \mathring{\pi}_{(AS)\setminus B} \right).$$

To interpret, the DM's associative network is formed at the ex ante stage according

[14] The definitions of π^f_A, π^n_A, $\mathring{\pi}^f_A$ and $\mathring{\pi}^n_A$ are similar to those of π_A and $\mathring{\pi}_A$ in Section 3.

to the distribution μ. Then the DM conducts mental association following the formed associative network when making decisions.

A random associative network μ is said to be a link-independent associative network if there is a function $\theta : X^2 \to [0,1]$ such that:
(1) for all $x \in X$, $\theta(x,x) = 1$,
(2) for all $\mathcal{G} \in \mathscr{G}$, $\mu(\mathcal{G}) = \left(\prod_{(x,y)\in\mathcal{G}} \theta(x,y)\right)\left(\prod_{(x,y)\in X^2\setminus\mathcal{G}}(1-\theta(x,y))\right)$.
To interpret, $\theta(x,y)$ is the probability that the DM can associate y with x, and condition (2) indicates that the DM forms each association link independently. In particular, when the network is undirected and the value of $\theta(x,y)$ is the same for all pairs of x and y, the link-independent associative network coincides with the well-known Erdős-Rényi random graph.[15]

For a link-independent associative network, the preference ordering $\succ$ and the attention probability function π can be identified similarly as our baseline model. The function θ can also be uniquely identified. To see this, consider two distinction alternatives x and y. We have

$$\rho(y|y,x) = 1 - (1-\pi_y)(1-\pi_x\theta(x,y)),$$

i.e., the probability for y being unselected in extended menu (y,x) is equal to the probability that y is not initially attended to and not considered through the initial consideration of x. Thus, the probability for y being associated with x is given by

$$\theta(x,y) = \frac{1}{\pi_x} - \frac{1-\rho(y|y,x)}{\pi_x - \pi_x\pi_y}.$$

However, for a general random associative network μ, the identification may not be unique. We demonstrate this through the following simple example.

Example 4. Let $X = \{x,y\}$ be the space of alternatives. Consider the following random choice rule ρ.

$$\rho(y|xy,\emptyset) = \frac{1}{8},\ \rho(x|xy,\emptyset) = \rho(x|x,y) = \rho(y|y,x) = \frac{5}{8},\ \rho(x|x,\emptyset) = \rho(y|y,\emptyset) = \frac{1}{2}.$$

Based on the choice rule, it can be revealed that DM's preference is $x \succ y$, and her attention probability function is given by $\pi_x = \pi_y = \frac{1}{2}$. However, the choice

[15] See Jackson (2008) and Goyal (2023) for a more detailed discussion of random networks.

rule can be represented by more than one random associative networks: Any random associative network μ can represent the choice rule if the probability of y being associated with x and that of x being associated with y are both equal to $\frac{1}{2}$ under μ. Thus, we can consider four (deterministic) associative networks $\mathcal{G}_0 = \mathcal{X}$, $\mathcal{G}_1 = \mathcal{X} \cup \{(x, y)\}$, $\mathcal{G}_2 = \mathcal{X} \cup \{(y, x)\}$, and $\mathcal{G}_3 = \mathcal{X} \cup \{(x, y), (y, x)\}$, and the two random associative networks μ_1 and μ_2, with $\mu_1(\mathcal{G}_0) = \mu_1(\mathcal{G}_3) = \frac{1}{2}$ and $\mu_2(\mathcal{G}_1) = \mu_2(\mathcal{G}_2) = \frac{1}{2}$, can both represent ρ. $\square$

7 Conclusion

In this paper, we present a novel choice model of mental association, which underscores the significance of understanding mental associations as a cognitive procedure for consideration set formation. Our study has several potential applications in various fields. One such application is in the design of marketing and advertising strategies. By understanding consumers' mental associations, firms can tailor their marketing messages to shape consumers' perceptions of their products or services. For example, firms can create positive mental associations between their brand and other popular products in the market, thereby increasing the likelihood of their product being considered by the consumer.

Another potential application of our model is in the analysis of consumer behavior. Our model suggests that the DM's choice frequencies can be influenced by options that are not included in the feasible choice set. This has important implications for understanding how consumers make decisions in markets where there are many similar products or services, but not all of them are always available. By considering mental association procedures, researchers can develop more accurate models of consumer behavior. Similarly, our model can also be adopted by policymakers to design more effective public policies. For instance, policymakers need to be aware of negative associations of their policies, which may cue individuals to be aware of socially worse choices.

Overall, our study provides new insights into the role of mental associations in decision making and has important implications for a wide range of fields, including marketing, consumer behavior, and public policy. The potential applications of our model are vast, and we hope that our study inspires more research in this area.

8 Appendix

Proof of Proposition 1. If $x \in c(A,S)$, then there exists $B \subseteq AS$ such that $x = \max(\bar{\mathcal{G}}_{AS}(B) \cap A; \succ)$. Since $x \in \bar{\mathcal{G}}_{AS}(B)$, we have $\bar{\mathcal{G}}_{AS}(x) \subseteq \bar{\mathcal{G}}_{AS}(B)$, and thus $x = \max(\bar{\mathcal{G}}_{AS}(x) \cap A; \succ)$. Inversely, if $x = \max(\bar{\mathcal{G}}_{AS}(x) \cap A; \succ)$, then x is chosen when the initial consideration set is x. Thus $x \in c(A,S)$. $\square$

Proof of Proposition 2. To see condition (1), consider an extended menu (A,S). The right-hand-side set is a subset of the left-hand-side set. To show the inverse direction, consider $x \in AS$ such that $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$. Let $y = \max(\bar{\mathcal{G}}_{AS}(x) \cap A; \succ)$. Since $y \in \bar{\mathcal{G}}_{AS}(x)$, we have $\bar{\mathcal{G}}_{AS}(y) \subseteq \bar{\mathcal{G}}_{AS}(x)$. Thus $y = \max(\bar{\mathcal{G}}_{AS}(y) \cap A; \succ)$ and $x \in \Gamma_{\mathcal{G}}^{y,\succ}(A,S)$. By Proposition 1, $y \in c(A,S)$, and condition (1) holds.

For equation (2), note that for all initial consideration set $B \subseteq AS$, we have $\bar{\mathcal{G}}_{AS}(B) \cap A \neq \emptyset$ if and only if there exists $x \in B$ such that $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$. Thus the default option is chosen if and only if any alternative x that satisfies $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$ is not initially considered. This leads to equation (2).

For equation (3), without loss of generality, we show that the equation holds for each $k \in \{2, ..., n\}$. When the initial consideration set is B, x_k is chosen if and only if $\bar{\mathcal{G}}_{AS}(B) \cap \{x_1, ..., x_{k-1}\} = \emptyset$ and $x_k \in \bar{\mathcal{G}}_{AS}(B)$. That is, $B \cap \Gamma_{\mathcal{G}}^{x_k,\succ}(A,S) \neq \emptyset$ and for all $m \leq k-1$, $B \cap \Gamma_{\mathcal{G}}^{x_m,\succ}(A,S) = \emptyset$. This leads to equation (3). $\square$

Proof of Theorem 1. **(Necessity)** Consider a choice rule ρ that is represented by $(\pi, \mathcal{G}, \succ)$ as an RCAR. We show that Axioms 1-5 hold for ρ. For Axiom 1, note that by Proposition 2, for all $A \in \mathcal{M}$, $\Phi(A,\emptyset) = \mathring{\pi}_A$. Thus for all $x \in A$, $\frac{\Phi(A,\emptyset)}{\Phi(A \setminus x,\emptyset)} = \mathring{\pi}_x$, which is independent with the menu A.

For the first half of Axiom 2, consider $x \in X$ and $(A,S) \in \mathcal{S}$ with $x \notin AS$. By Proposition 2, we have $\Phi(Ax,S) = \mathring{\pi}_{H_{\mathcal{G}}(Ax,S)}$ and $\Phi(A,Sx) = \mathring{\pi}_{H_{\mathcal{G}}(A,Sx)}$. Since $H_{\mathcal{G}}(A,Sx) \subseteq H_{\mathcal{G}}(Ax,S)$, we have $\Phi(Ax,S) \leq \Phi(A,Sx)$. For every $y \in A$, let $B_y = \Gamma_{\mathcal{G}}^{y,\succ}(Ax,S)$, $D_y = \cup_{z \in Ax: z \succ y} B_z$, $\hat{B}_y = \Gamma_{\mathcal{G}}^{y,\succ}(A,Sx)$, and $\hat{D}_y = \cup_{z \in A: z \succ y} \hat{B}_z$. For every $y \in A$, we have $\rho(y|Ax,S) = \mathring{\pi}_{D_y}(1-\mathring{\pi}_{B_y})$ and $\rho(y|A,Sx) = \mathring{\pi}_{\hat{D}_y}(1-\mathring{\pi}_{\hat{B}_y})$. Note that if $y \succ x$, then $B_y = \hat{B}_y$. Thus for every $y \in A$ with $y \succ x$, we have $D_y = \hat{D}_y$. If $x \succ y$, then $B_y \subseteq \hat{B}_y$, and since $D_y = H_{\mathcal{G}}(Ax,S) \setminus (\cup_{z \in A: y \succ z \text{ or } z=y} B_y)$

and $\hat{D}_y = H_{\mathcal{G}}(A, Sx) \backslash \left(\cup_{z \in A: y \succ z \text{ or } z=y} \hat{B}_y \right)$, we have $\hat{D}_y \subseteq D_y$. It follows that for all $y \in A$, $\rho(y|Ax, S) \leq \rho(y|A, Sx)$.

For the second half of Axiom 2, note that $\Phi(Ax, S) < \Phi(A, Sx)$ implies $H_{\mathcal{G}}(A, Sx) \subsetneq H_{\mathcal{G}}(Ax, S)$. Thus there exists $y \in ASx$ such that $\bar{\mathcal{G}}_{ASx}(y) \cap A = \emptyset$ and $\bar{\mathcal{G}}_{ASx}(y) \cap Ax \neq \emptyset$, i.e., $\bar{\mathcal{G}}_{ASx}(y) \cap Ax = x$. Since $x \in \bar{\mathcal{G}}_{ASx}(y)$, we have $\bar{\mathcal{G}}_{ASx}(x) \subseteq \bar{\mathcal{G}}_{ASx}(y)$, and thus $\bar{\mathcal{G}}_{ASx}(x) \cap A = \emptyset$. Thus $\{y \in ASx : \bar{\mathcal{G}}_{ASx}(y) \cap A \neq \emptyset\} = \{y \in AS : \bar{\mathcal{G}}_{AS}(y) \cap A \neq \emptyset\}$. By Proposition 2, we have $\Phi(A, Sx) = \Phi(A, S)$.

For Axiom 3, consider two extended menus (A, S) and (B, T) with $B \subseteq A$ and $BT \subseteq AS$. Since $x \xrightarrow{T} B$, either (1) $x \in B$ or (2) $\Phi(B, T) \neq \Phi(B, T\backslash x)$. Case (1) directly implies $x \in A$ and thus $x \xrightarrow{S} A$. Case (2) implies $\bar{\mathcal{G}}_{BT}(x) \cap B \neq \emptyset$, and thus $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$. We have either $x \in A$ or $\Phi(A, S) \neq \Phi(A, S\backslash x)$, both of which imply $x \xrightarrow{S} A$.

For Axiom 4, consider $x, y \in X$ and $(A, S) \in \mathcal{S}$ such that $x \neq y$, $x \xrightarrow{S} A$, and not $x \xrightarrow{S\backslash y} A\backslash y$. It follows that $x \in S$, $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$, and $\bar{\mathcal{G}}_{(AS)\backslash y}(x) \cap A = \emptyset$. Since $\bar{\mathcal{G}}_{AS}(x) \cap A \neq \emptyset$, there is a sequence $(x_k)_{k=1}^n$ of mutually distinct alternatives in AS such that $x_1 = x$, $\{x_1, ..., x_n\} \cap A = x_n$, and for all $k \in \{1, ..., n-1\}$, $(x_k, x_{k+1}) \in \mathcal{G}$. However, since $\bar{\mathcal{G}}_{(AS)\backslash y}(x) \cap A = \emptyset$, for all such sequence, there exists $k \in \{2, ..., n\}$ such that $x_k = y$. Therefore, we have $x \xrightarrow{S\backslash y} y$ and $y \xrightarrow{S\backslash x} A$.

For Axiom 5, consider $x \in c(A, S)$ and $y \in c(B, T)$ that satisfy the primitive conditions of the axiom. We first show that $x \succ y$. Since $x \in c(A, S)$, for all $z \in A$ such that $z \succ x$, $\Gamma_{\mathcal{G}}^{z,\succ}(A, S) = \Gamma_{\mathcal{G}}^{z,\succ}(A\backslash x, S)$. Therefore, by Proposition 2, for all $z \in A$ such that $z \succ x$, $\rho(z|A, S) = \rho(z|A\backslash x, S)$. Since $\rho(y|A, S) \neq \rho(y|A\backslash x, S)$, we conclude that $x \succ y$. Since $x \succ y$ and $y \in c(B, T)$, a similar argument implies $\rho(x|B, T) = \rho(x|B\backslash y, Ty) = \rho(x|B\backslash y, T)$.

(Sufficiency) Throughout the proof, we assume that Axioms 1-5 hold. We first identify the attention probability function π. For each alternative $x \in X$, define $\pi_x = \rho(x|x, \emptyset) \in (0, 1)$. By Axiom 1, for all $A \in \mathcal{M}$, $\Phi(A, \emptyset) = \mathring{\pi}_A$. Next, define

$$\mathcal{G} = \mathcal{X} \cup \{(x, y) \in X^2 : x \neq y \text{ and } \rho(y|y, \emptyset) \neq \rho(y|y, x)\}$$

as the associative network. Define the binary relation $\succ$ such that for any two distinct alternatives x and y, $x \succ y$ if and only if $\rho(y|y, \emptyset) > \rho(y|xy, \emptyset)$. Note that if

$x \succ y$, then $x \in c(xy, \emptyset)$, since otherwise $\rho(y|xy, \emptyset) = 1 - \Phi(xy, \emptyset) > 1 - \Phi(y, \emptyset) = \rho(y|y, \emptyset)$, which contradicts to the definition of $\succ$. If $x \succ y$, then by Axiom 5, we have for all $(A, S) \in \mathcal{S}$ with $y \in c(A, S)$, $\rho(x|A, S) = \rho(x|A\backslash y, Sy) = \rho(x|A\backslash y, S)$.

Lemma 1. *The binary relation $\succ$ is transitive and asymmetric, and satisfies that for all $x, y \in X$ with $x \neq y$, either $x \succ y$ or $y \succ x$.*

Proof of Lemma 1. Consider two distinct alternatives x and y. We first show that either $x \succ y$ or $y \succ x$. To see this, note that $\rho(x|xy, \emptyset) + \rho(y|xy, \emptyset) = 1 - \mathring{\pi}_{xy} = \pi_x + \pi_y - \pi_{xy} < \pi_x + \pi_y = \rho(x|x, \emptyset) + \rho(y|y, \emptyset)$. Thus, either $\rho(x|xy, \emptyset) < \rho(x|x, \emptyset)$ or $\rho(y|xy, \emptyset) < \rho(y|y, \emptyset)$, i.e., either $x \succ y$ or $y \succ x$.

Next, we show that $\succ$ is asymmetric, i.e., if $x \succ y$, then not $y \succ x$. If $x \succ y$, then either $\rho(y|xy, \emptyset) = 0$, which implies $\rho(x|xy, \emptyset) = \pi_x + \pi_y - \pi_{xy} > \pi_x = \rho(x|x, \emptyset)$, or $\rho(y|xy, \emptyset) > 0$, which by Axiom 5 implies $\rho(x|xy, \emptyset) = \rho(x|x, \emptyset)$. Thus $y \not\succ x$.

Finally, we show that $\succ$ is transitive. Consider three mutually distinct alternatives x, y and z with $x \succ y$ and $y \succ z$. We show $x \succ z$. Suppose to the contrary that $z \succ x$. We consider the following three representative cases and show that all of them lead to contradictions.

Case 1: $c(xyz, \emptyset) = xyz$. By Axiom 5, we have $\rho(x|xyz, \emptyset) = \rho(x|xz, \emptyset) > 0$. Since $z \succ x$, by Axiom 5, we have $\rho(z|xz, \emptyset) = \rho(z|z, \emptyset) = \pi_z$. Thus $\rho(x|xyz, \emptyset) = \rho(x|xz, \emptyset) = \mathring{\pi}_z \pi_x$. Similarly, we have $\rho(y|xyz, \emptyset) = \mathring{\pi}_x \pi_y$ and $\rho(z|xyz, \emptyset) = \mathring{\pi}_y \pi_z$. This leads to a contradiction since

$$\begin{aligned}&\rho(x|xyz, \emptyset) + \rho(y|xyz, \emptyset) + \rho(z|xyz, \emptyset) = \pi_x + \pi_y + \pi_z - \pi_{xy} - \pi_{xz} - \pi_{yz}\\ <&\pi_x + \pi_y + \pi_z - \pi_{xy} - \pi_{xz} - \pi_{yz} + \pi_{xyz} = 1 - \mathring{\pi}_{xyz} = 1 - \Phi(xyz, \emptyset).\end{aligned}$$

Case 2: $c(xyz, \emptyset) = xy$. By Axiom 2, we have $\rho(x|xyz, \emptyset) \leq \rho(x|xy, z)$, $\rho(y|xyz, \emptyset) \leq \rho(y|xy, z)$, and $\Phi(xyz, \emptyset) \leq \Phi(xy, z)$. Since $\rho(z|xyz, \emptyset) = 0$, we have $\rho(x|xyz, \emptyset) = \rho(x|xy, z)$, $\rho(y|xyz, \emptyset) = \rho(y|xy, z)$, and $\Phi(xyz, \emptyset) = \Phi(xy, z)$. Since $x \succ y$, $y \in c(xyz, \emptyset)$ and $y \in c(xy, z)$, by Axiom 5, we have $\rho(x|x, yz) = \rho(x|xy, z) = \rho(x|xyz, \emptyset) = \rho(x|xz, \emptyset)$. Since $z \succ x$ and $x \in c(xz, \emptyset)$, we have $\rho(x|xz, \emptyset) = \mathring{\pi}_z \pi_x$. Thus $\rho(x|x, yz) = \mathring{\pi}_z \pi_x$ and $\Phi(x, yz) = 1 - \mathring{\pi}_z \pi_x > \mathring{\pi}_x$. However, note that by Axiom 2, either $\Phi(x, yz) = \Phi(x)$, or $\Phi(x, yz) = \Phi(xy)$, or $\Phi(x, zy) = \Phi(xz)$, or $\Phi(x, yz) = \Phi(xyz)$, all of which contradict to $\Phi(x, yz) > \mathring{\pi}_x$.

Case 3: $c(xyz,\emptyset)=x$. Since $\rho(y|xyz,\emptyset)=0$, by Axiom 2, we have $\rho(x|xz,y)=\rho(x|xyz,\emptyset)>0$ and $\rho(z|xz,y)=\rho(z|xyz,\emptyset)=0$. Since $z\succ x$ and $x\in c(xz,y)$, by Axiom 5, we have $\rho(z|z,xy)=\rho(z|xz,y)=0$, which contradicts to the definition of ρ that for all extended menu (A,S), $\sum_{\hat{x}\in A}\rho(\hat{x}|A,S)>0$. $\square$

Lemma 2. *For all $x\in X$ and $(A,S)\in\mathcal{S}$ with $x\in S$, if $\Phi(A,S)\neq\Phi(A,S\backslash x)$, then $\bar{\mathcal{G}}_{AS}(x)\cap A\neq\emptyset$.*

Proof of Lemma 2. Consider $x\in X$ and $(A,S)\in\mathcal{S}$ such that $\Phi(A,S)\neq\Phi(A,S\backslash x)$, i.e., $x\xrightarrow{S}A$. Note that $A\neq\emptyset$ since otherwise $\Phi(A,S)=\Phi(A,S\backslash x)=1$. Also, note that $x\xrightarrow{S}\emptyset$ is not true. Thus, by Axiom 4 and a simple induction, we can find some $y\in A$ such that $x\xrightarrow{S}y$. It suffices to show $y\in\bar{\mathcal{G}}_{Sy}(x)$, and we show this by induction.

First, note that if $|S|=1$, then $S=x$. Thus $x\xrightarrow{S}y$ implies $\Phi(y,x)\neq\Phi(y,\emptyset)$, which further implies $\rho(y|y,x)\neq\rho(y|y,\emptyset)$, i.e., $(x,y)\in\mathcal{G}$, and we are done. Next, suppose that when $|S|\leq n$, $x\xrightarrow{S}y$ implies $y\in\bar{\mathcal{G}}_{Sy}(x)$. We show that when $|S|=n+1$, $x\xrightarrow{S}y$ also implies $y\in\bar{\mathcal{G}}_{Sy}(x)$. To see this, note that if there is $z\in S\backslash x$ such that $x\xrightarrow{S\backslash z}y$, then by the induction hypothesis, $y\in\bar{\mathcal{G}}_{Sy}(x)$. Otherwise, for all $z\in S\backslash x$, $x\xrightarrow{S\backslash z}y$ does not hold. By Axiom 4, we have for all $z\in S\backslash x$, $x\xrightarrow{S\backslash z}z$ and $z\xrightarrow{S\backslash x}y$. By our induction hypothesis, for all $z\in S\backslash x$, $y\in\bar{\mathcal{G}}_{(Sy)\backslash x}(z)$ and $z\in\bar{\mathcal{G}}_{S}(x)$. By the definition of $\bar{\mathcal{G}}$, we have $y\in\bar{\mathcal{G}}_{Sy}(x)$. $\square$

Lemma 3. *For all $x\in X$ and $(A,S)\in\mathcal{S}$ with $x\in S$, if $\bar{\mathcal{G}}_{AS}(x)\cap A\neq\emptyset$, then $\Phi(A,S)\neq\Phi(A,S\backslash x)$.*

Proof of Lemma 3. Suppose $x\in S$ and $\bar{\mathcal{G}}_{AS}(x)\cap A\neq\emptyset$. Then we can find a sequence $(x_k)_{k=1}^{n}$ such that $x_1=x$, $x_n\in A$, and for all $k\in\{1,...,n-1\}$, $(x_k,x_{k+1})\in\mathcal{G}$ and $x_k\in S$. We want to show $\Phi(x_n,\{x_1,...,x_{n-1}\})\neq\Phi(x_n,\{x_2,...,x_{n-1}\})$, then by Axiom 3, we have $\Phi(A,S)\neq\Phi(A,S\backslash x)$. For notation simplicity, we use $B_{k,m}$, where $n\geq m\geq k\geq 1$, to denote the set $\{x_k,x_{k+1},...,x_m\}$.

First, we show $\Phi(x_n,B_{1,n-1})=\Phi(B_{1,n},\emptyset)$. To see this, note that $(x_{n-1},x_n)\in\mathcal{G}$ implies $\Phi(x_n,x_{n-1})\neq\Phi(x_n,\emptyset)$. By Axiom 3, $\Phi(x_n,B_{1,n-1})\neq\Phi(x_n,B_{1,n-2})$. By Axiom 2, we have $\Phi(x_n,B_{1,n-1})=\Phi(x_{n-1}x_n,B_{1,n-2})$. Thus, a simple induction establishes that $\Phi(x_n,B_{1,n-1})=\Phi(B_{1,n},\emptyset)$. By the same argument, we can inductively show $\Phi(x_n,B_{2,n-1})=\Phi(B_{2,n},\emptyset)$. Since $(x_1,x_2)\in\mathcal{G}$, by Axiom 3, we have $\Phi(B_{2,n}, \emptyset) \neq \Phi(B_{2,n}, x_1)$. By Axiom 2, we have $\Phi(B_{2,n}, x_1) = \Phi(B_{1,n}, \emptyset)$. Since $\Phi(B_{1,n}, \emptyset) \neq \Phi(B_{2,n}, \emptyset)$, we have $\Phi(x_n, B_{1,n-1}) \neq \Phi(x_n, B_{2,n-1})$. $\square$

Lemma 4. *For all $(A, S) \in \mathcal{S}$, $\Phi(A, S) = \mathring{\pi}_{H_\mathcal{G}(A,S)}$.*

Proof of Lemma 4. By Lemma 3 and Axiom 2, we can shift all alternatives in $H_\mathcal{G}(A, S)$ to the feasible set without changing the choice probability of the default option, i.e., $\Phi(A, S) = \Phi(H_\mathcal{G}(A, S), S\backslash H_\mathcal{G}(A, S))$. For all $x \in S\backslash H_\mathcal{G}(A, S)$, $\bar{\mathcal{G}}_{AS}(x) \cap H_\mathcal{G}(A, S) = \emptyset$. By Lemma 2, we have $\Phi(H_\mathcal{G}(A, S), S\backslash H_\mathcal{G}(A, S)) = \Phi(H_\mathcal{G}(A, S), \emptyset)$. Thus $\Phi(A, S) = \Phi(H_\mathcal{G}(A, S), \emptyset) = \mathring{\pi}_{H_\mathcal{G}(A,S)}$. $\square$

Lemma 5. *For all $x \in X$ and $(A, S) \in \mathcal{S}$ with $x \notin AS$, if $\Phi(Ax, S) = \Phi(A, Sx)$, then $\Phi(A, Sx) < \Phi(A, S)$.*

Proof of Lemma 5. For all $x \in X$ and $(A, S) \in \mathcal{S}$ with $x \notin AS$, we have $H_\mathcal{G}(A, S) \subsetneq H_\mathcal{G}(Ax, S)$. By Lemma 4, $\Phi(Ax, S) < \Phi(A, S)$. Hence, if $\Phi(Ax, S) = \Phi(A, Sx)$, then $\Phi(A, Sx) < \Phi(A, S)$. $\square$

Lemma 6. *For all $x \in X$ and $(A, S) \in \mathcal{S}$ with $x \in A$, if there is $y \in \bar{\mathcal{G}}_{AS}(x) \cap A$ such that $y \succ x$, then $x \notin c(A, S)$.*

Proof of Lemma 6. Consider x, y, and (A, S) that satisfy the conditions stated in the lemma. It suffices to show $x \notin c(xy, (AS)\backslash(xy))$, then by Axiom 2, we have $x \notin c(A, S)$. Suppose to the contrary that $x \in c(xy, (AS)\backslash(xy))$. Since $y \succ x$, by Axiom 5, we have $\rho(y|xy, (AS)\backslash(xy)) = \rho(y|y, (AS)\backslash y)$. Thus $\Phi(xy, (AS)\backslash(xy)) < \Phi(y, (AS)\backslash y)$. By Axiom 2, we have $\Phi(y, (AS)\backslash y) = \Phi(y, (AS)\backslash(xy))$, which contradicts to Lemma 3 as $y \in \bar{\mathcal{G}}_{AS}(x)$. $\square$

Lemma 7. *For all $x \in X$ and $(A, S) \in \mathcal{S}$ with $x \in A$, if $x \notin c(A, S)$, then there is $y \in \bar{\mathcal{G}}_{AS}(x) \cap A$ such that $y \succ x$.*

Proof of Lemma 7. Consider x and (A, S) that satisfy the conditions stated in the lemma. We prove the lemma by induction. If $x \in A$ and $x \notin c(A, S)$, then $|A| \geq 2$. When $|A| = 2$, there exists $y \neq x$ such that $c(A, S) = y$ (note that $A = xy$ in this case). Since $\rho(x|xy, S) \neq \rho(x|x, S)$, by the property of the preference $\succ$ and Axiom 5, we have $y \succ x$. Since $x \notin c(xy, S)$, by Axiom 2, $\Phi(xy, S) = \Phi(y, Sx)$. By Lemma 5, we have $\Phi(y, Sx) \neq \Phi(y, S)$, i.e., $y \in \bar{\mathcal{G}}^x_{AS}$, and we are done.

Suppose that the lemma holds when $|A| \leq n$, where $n \geq 2$. We show that it also holds for $|A| = n+1$. If there exists $z \in A \backslash x$ such that $z \notin c(A, S)$, then by Axiom 2, we have $x \notin c(A \backslash z, Sz)$. Since $|A \backslash z| = n$, we are done. Consider the case where $x = A \backslash c(A, S)$. If there exists $z \in A \backslash x$ such that $x \notin c(A \backslash z, S)$, then again we are done. Thus we consider the case where $x = A \backslash c(A, S)$ and for all $y \in A \backslash x$, $x \in c(A \backslash y, S)$. By the property of the preference $\succ$ and Axiom 5, we have for all $y \in A \backslash x$, $y \succ x$. Since $x \notin c(A, S)$, by Axiom 2, we have $\Phi(A, S) = \Phi(A \backslash x, Sx)$. By Lemma 5, we have $\Phi(A \backslash x, Sx) \neq \Phi(A \backslash x, S)$. By Lemma 2, $\bar{\mathcal{G}}_{AS}(x) \cap (A \backslash x) \neq \emptyset$, i.e., there exists $y \in A$ such that $y \succ x$ and $y \in \bar{\mathcal{G}}_{AS}(x)$. □

Now, we verify that the choice probability of each alternative in a given extended menu (A, S) coincides with the RCAR represented by $(\pi, \mathcal{G}, \succ)$. To avoid triviality, assume $A \neq \emptyset$. By Lemmas 6 and 7, an alternative in A is chosen with positive probability if and only if there is no $\succ$-better alternative in A that is associated with it. Without loss of generality, let $\{x_1, ..., x_n\}$ be those alternatives in A that are chosen with positive probabilities such that $x_1 \succ x_2 \succ ... \succ x_n$. Following Proposition 2, we construct a partition $\{B_1, ..., B_{n+1}\}$ of AS such that for all $k \in \{1, ..., n\}$,

$$B_k = \Gamma_{\mathcal{G}}^{x_k, \succ}(A, S), \text{ and}$$

$$B_{n+1} = (AS) \backslash H_{\mathcal{G}}(A, S).$$

It suffices to show that for all $k \in \{1, ..., n\}$, $\rho(x_k | A, S) = \mathring{\pi}_{B_1 \ldots B_{k-1}}(1 - \mathring{\pi}_{B_k})$. By the construction, for all $k \in \{1, ..., n\}$, we have for all $y \in (B_k \cap A) \backslash x_k$, $x_k \succ y$.

By Lemma 4, we have

$$\sum_{k=1}^{n} \rho(x_k | A, S) = 1 - \Phi(A, S) = 1 - \mathring{\pi}_{B_1 \ldots B_n}. \tag{8}$$

If $n = 1$, then we are done. Consider the case where $n \geq 2$. What remains to be shown is that for all $m \in \{1, ..., n-1\}$, we have

$$\sum_{k=1}^{m} \rho(x_k | A, S) = 1 - \mathring{\pi}_{B_1 \ldots B_m}. \tag{9}$$

To show this, first note the following observation: For all $m \in \{1, ..., n-1\}$ and all $D \subseteq AS$ such that $\cup_{k=1}^{m} B_k \subseteq D$, if $D \cap A \cap (\cup_{k=m+1}^{n} B_k) \neq \emptyset$, then

$c(D\cap A, D\cap S)\cap(\cup_{k=m+1}^{n}B_k)\neq\emptyset$. This observation is true since otherwise for all $z\in D\cap A\cap(\cup_{k=m+1}^{n}B_k)$, $\bar{\mathcal{G}}_D(z)\cap(\cup_{k=1}^{m}B_k)\neq\emptyset$, indicating that $z\in\cup_{k=1}^{m}B_k$, which is a contradiction. Thus we can enumerate alternatives in $A\cap\left(\cup_{k=m+1}^{n}B_k\right)$ as $A\cap\left(\cup_{k=m+1}^{n}B_k\right)=\{z_1,...,z_l\}$ such that $z_1\in c(A,S)$, and for all $k\in\{1,...,l-1\}$, $z_{k+1}\in c(A\backslash\{z_1,...,z_k\},S)$. By Axiom 5 and the construction of $\succ$, we have for all $k\in\{1,...,l\}$ and $\hat{k}\in\{1,...,m\}$,

$$\rho(x_{\hat{k}}|A\backslash\{z_1,...,z_k\},S)=\rho(x_{\hat{k}}|A,S).$$

Thus for all $k\in\{1,...,m\}$, we have $\rho(x_k|A,S)=\rho(x_k|A\cap(\cup_{l=1}^{m}B_l),S)$. One can easily show that $\Phi(A\cap(\cup_{l=1}^{m}B_l),S)=\mathring{\pi}_{B_1...B_m}$ and $c(A\cap(\cup_{l=1}^{m}B_l),S)=\{x_1,...,x_m\}$. Therefore, equation (9) holds. $\square$

Proof of Theorem 2. **(Necessity)** By Proposition 1, Axiom 6 holds trivially. For Axiom 7, consider $A\in\mathcal{M}$ such that $c(A,\emptyset)\subsetneq A$. Without loss of generality, let $c(A,\emptyset)=\{x_1,...,x_n\}$. By Proposition 2, we can have a partition $\{B_k\}_{k=1}^{n}$ of A such that for all $k\in\{1,...,n\}$, we have $B_k=\Gamma_{\mathcal{G}}^{x_k,\succ}(A,\emptyset)$. Let x_t be the $\succ$-worst alternative in $c(A,\emptyset)$ satisfying $B_t\neq x_t$. For such x_t, let $D^1=x_t$, and for each $k\geq 2$, define inductively $D^k=\{x\in B_t:\exists y\in D^{k-1}$ such that $(x,y)\in\mathcal{G}\}$. It is easy to show that there is a minimal number $k\geq 2$ such that $D^{k-1}\subsetneq D^k=D^{k+1}=B_t$. Pick an arbitrary $z\in D^k\backslash D^{k-1}$. We have $\Gamma_{\mathcal{G}}^{x_t,\succ}(A\backslash z,\emptyset)=B_t\backslash z$, and for all $k\in\{1,...,n\}\backslash t$, $\Gamma_{\mathcal{G}}^{x_k,\succ}(A\backslash z,\emptyset)=\Gamma_{\mathcal{G}}^{x_k,\succ}(A,\emptyset)$. It follows that $c(A,\emptyset)=c(A\backslash z,\emptyset)$, and Axiom 7 holds.

For Axiom 8, note that $x\vdash A$ if and only if for all $y\in c(A,\emptyset)$, $y\succ x$, and for all $z\in A$, $(x,z)\notin\mathcal{G}$. It follows that $x\vdash A$ and $x\vdash B$ imply $x\vdash AB$. Axiom 9 can be shown by a similar proof to that for the necessity of Axiom 5 in Theorem 1.

(Sufficiency) Assume that Axioms 1 and 6-9 hold. The attention probability π and the preference ordering $\succ$ are defined in the same way as the proof of Theorem 1. For the associative network $\mathcal{G}$, let $(x,y)\in\mathcal{G}$ if and only if either (1) $x=y$, or (2) $x\neq y$ and there exists $A\in\mathcal{M}$ such that $x\vdash A$ and $x\notin c(Axy,\emptyset)$. We proceed with a sequence of lemmas.

Lemma 8. *For all $x\in X$ and $A\in\mathcal{M}$, if $x\vdash A$, then $x\in c(Ax,\emptyset)$.*

Proof of Lemma 8. Since $x \vdash A$, we have $\sum_{y\in A} \rho(y|A, \emptyset) = \sum_{y\in A} \rho(y|Ax, \emptyset)$. Since $\Phi(Ax, \emptyset) < \Phi(A, \emptyset)$, we have $\rho(x|Ax, \emptyset) \neq 0$, i.e., $x \in c(Ax, \emptyset)$. □

Lemma 9. *The binary relation $\succ$ is a strict preference ordering and satisfies that for all $x, y \in X$ and $A \in \mathcal{M}$, if $x \succ y$ and $y \in c(A, \emptyset)$, then $\rho(x|A, \emptyset) = \rho(x|A\backslash y, \emptyset)$.*

Proof of Lemma 9. The claim that $x \succ y$ and $y \in c(A, \emptyset)$ imply $\rho(x|A, \emptyset) = \rho(x|A\backslash y, \emptyset)$ follows directly from the definition of $\succ$ and Axiom 9. Similar to the proof of Lemma 1, we can easily show that $\succ$ is well-defined for each pair of distinct alternatives and satisfies asymmetry. To see that $\succ$ is transitive, suppose to the contrary that there are three mutually distinct alternatives x, y and z such that $x \succ y$, $y \succ z$, and $z \succ x$. Consider three representative cases, where in case 1, $c(xyz, \emptyset) = xyz$, in case 2, $c(xyz, \emptyset) = xy$, and in case 3, $c(xyz, \emptyset) = x$. We want to show that all the three cases lead to contradiction. The proof for case 1 is the same as that for case 1 in Lemma 1.

For case 2, since $c(xyz, \emptyset) = xy$ and $z \succ x$, we have $c(yz, \emptyset) = y$. It follows that $y \trianglerighteq z$, and thus by Axiom 9 and the definition of $z \succ x$, we have for all $A \in \mathcal{M}$ with $x \in c(A, \emptyset)$, $\rho(y|A, \emptyset) = \rho(y|A\backslash x, \emptyset)$. By Axiom 6, $x \in c(xy, \emptyset)$, and thus we have $\rho(y|xy, \emptyset) = \rho(y|y, \emptyset)$, which contradicts to the definition of $x \succ y$.

For case 3, since $c(xyz, \emptyset) = x$, we have $x \trianglerighteq z$. By the definition of $z \succ x$, we have $z \in c(xz, \emptyset)$ and $\rho(x|xz, \emptyset) \neq \rho(x|x, \emptyset)$, which contradicts to $x \trianglerighteq z$ according to Axiom 9. □

Lemma 10. *For all $x \in X$ and $A \in \mathcal{M}$ with $x \in A$, if $x \notin c(A, \emptyset)$, then there exists $y \in \bar{\mathcal{G}}_A(x) \cap c(A, \emptyset)$ such that $y \succ x$.*

Proof of Lemma 10. We prove by induction on $|A|$. First, if $|A| = 2$, then $A = xy$ and $x \notin c(xy, \emptyset)$. Then by the construction of $\mathcal{G}$, we have $(x, y) \in \mathcal{G}$. By the construction of $\succ$, we have $y \succ x$. Therefore, the lemma holds when $|A| = 2$.

Assume that the lemma holds when $|A| \leq n$, where $n \geq 2$. Suppose now $|A| = n + 1$. Since $x \notin c(A, \emptyset)$, by Axiom 7, there exists $y \in A$ such that $c(A, \emptyset) = c(A\backslash y, \emptyset)$. If $y \neq x$, then $x \notin c(A\backslash y, \emptyset)$ and by our induction hypothesis, there exists $z \in \bar{\mathcal{G}}_{A\backslash y}(x)$ such that $z \succ x$, and we are done. Hence, we focus on the case where x is the only alternative in A such that $c(A, \emptyset) = c(A\backslash x, \emptyset)$. By a similar argument, we focus on the case where for all $z \in c(A, \emptyset)$, $x \in c(A \backslash z, \emptyset)$. Thus, by Lemma 9, we have for all $z \in c(A, \emptyset)$, $z \succ x$.

To proceed, consider $(A \backslash x, \emptyset)$. We first show that if $c(A \backslash x, \emptyset) = A \backslash x$, then $|A \backslash x| = 1$, and we are done. To see this, suppose to the contrary that $|A \backslash x| \geq 2$, and let y and z be two distinct alternatives in $A \backslash x$. By Axiom 6 and the assumptions of the case we consider, we have $x \in c(A \backslash y) = A \backslash y$ and $x \in c(A \backslash z) = A \backslash z$. Since for all $\hat{x} \in A \backslash x$, we have $\hat{x} \succ x$, by Lemma 9, we have $x \vdash A \backslash (xy)$ and $x \vdash A \backslash (xz)$. It follows from Axiom 8 that $x \vdash A \backslash x$, which by Lemma 8 is a contradiction since $x \notin c(A, \emptyset)$.

Finally, suppose that $c(A \backslash x, \emptyset) \neq A \backslash x$. In this case, by Axiom 7, there exists $y \in A \backslash x$ such that $c(A \backslash (xy), \emptyset) = c(A \backslash x, \emptyset) = c(A, \emptyset)$. By Axiom 6, we have $c(A \backslash y) = x \cup c(A)$. Since $x \in c(A \backslash y, \emptyset) = x \cup c(A)$ and for all $z \in c(A)$, $z \succ x$, by Lemma 9, we have $x \vdash A \backslash (xy)$. Since $x \notin c(A, \emptyset)$, we have $(x, y) \in \mathcal{G}$. Since $y \notin c(A \backslash x, \emptyset)$, by the induction hypothesis, there exists $z \in c(A \backslash x, \emptyset)$ such that $z \succ y$ and $z \in \bar{\mathcal{G}}_{A \backslash x}(y)$. Thus, we have $z \in \bar{\mathcal{G}}_A(x)$. Moreover, since $z \in c(A \backslash x, \emptyset)$, we have $z \in c(A, \emptyset)$, and thus $z \succ x$. $\square$

Lemma 11. *For all $x \in X$ and $A \in \mathcal{M}$, if $x \in c(A, \emptyset)$, then $x = \max(\bar{\mathcal{G}}_A(x); \succ)$.*

Proof of Lemma 11. We show that if there is an alternative in $(A, \emptyset)$ that is associated with x and $\succ$-better than x, then x is not chosen. By Axiom 6, it suffices to show that for any sequence of alternatives $(x_k)_{k=1}^n$, where $n \geq 2$, if for all $k \in \{1, ..., n-1\}$, $x_n \succ x_k$ and $(x_k, x_{k+1}) \in \mathcal{G}$, then $x_1 \notin c(x_1 ... x_n, \emptyset)$. We show this by induction on n. First, consider the case where $n = 2$. We have $(x_1, x_2) \in \mathcal{G}$ and $x_2 \succ x_1$. Suppose to the contrary that $x_1 \in c(x_1 x_2, \emptyset)$, then by Lemma 9 and the construction of $\succ$, we have $c(x_1 x_2, \emptyset) = x_1 x_2$ and $\rho(x_2 | x_1 x_2, \emptyset) = \rho(x_2 | x_2, \emptyset)$, i.e., $x_1 \vdash x_2$. However, since $(x_1, x_2) \in \mathcal{G}$, by the construction of $\mathcal{G}$, we can find $A \in \mathcal{M}$ such that $x_1 \vdash A$ and $x_1 \notin c(A x_1 x_2, \emptyset)$. By Axiom 8, we have $x_1 \vdash A x_2$, and by Lemma 8, we have $x_1 \in c(A x_1 x_2, \emptyset)$, which is a contradiction. Thus, we must have $x_1 \notin c(x_1 x_2, \emptyset)$.

Next, suppose that the induction hypothesis holds for $n \leq m$ $(m \geq 2)$. We consider the case where $n = m + 1$. Since for all $k \in \{1, ..., n-1\}$, $x_n \succ x_k$, we have $c(x_2 ... x_n, \emptyset) = x_n$ by our induction hypothesis. Suppose to the contrary that $x_1 \in c(x_1 ... x_n, \emptyset)$, by Axiom 6, we have $c(x_1 ... x_n, \emptyset) = x_1 x_n$. By Lemma 9, we have $\rho(x_n|x_1...x_n, \emptyset) = \rho(x_n|x_2...x_n, \emptyset)$. Thus $x_1 \vdash x_2...x_n$. Since $(x_1, x_2) \in \mathcal{G}$, we can find $A \in \mathcal{M}$ such that $x_1 \vdash A$ and $x_1 \notin c(Ax_1x_2, \emptyset)$. By Axiom 8, we have $x_1 \vdash Ax_2...x_n$, and by Lemma 8, we have $x_1 \in c(Ax_1...x_n, \emptyset)$. It follows from Axiom 6 that $x_1 \in c(Ax_1x_2, \emptyset)$, which is a contradiction. Therefore, we have $x_1 \notin c(x_1...x_n, \emptyset)$. □

With Lemmas 9, 10 and 11, the rest of the proof is the same as that of Theorem 1, and we omit it. □

Proof of Propsition 4. Consider a choice rule ρ that is represented by $(\pi, \mathcal{G}, \succ)$. We first show $\mathcal{G}^\rho \subseteq \mathcal{G}^c$. Consider $(x, y) \in \mathcal{G}^\rho$ with $x \neq y$. By the definition of $\mathcal{G}^\rho$, there exists $A \in \mathcal{M}$ such that $x \vdash A$ and $x \notin c(Axy, \emptyset)$. If $A = \emptyset$, then $x \notin c(xy, \emptyset)$, which implies that $(x, y) \in \mathcal{G}^c$. If $A \neq \emptyset$, then we can infer that

a. $(x, y) \in \mathcal{G}$,
b. there is $y' \in A$ such that $(y, y') \in \mathcal{G}$,
c. for all $z \in c(A, \emptyset)$, $z \succ x$, and
d. for all $w \in A$, $(x, w) \notin \mathcal{G}$.

We can find a sequence of mutually distinct alternatives $(x_k)_{k=1}^n$ in Axy such that $x_1 = x$, $x_2 = y$, $x_n \in c(A, \emptyset)$, and for all $k \in \{1, ..., n-1\}$, $(x_k, x_{k+1}) \in \mathcal{G}$ and $x_n \succ x_k$. It follows that $c(\{x_3, ..., x_n\}, \emptyset) = c(\{x_1, ..., x_n\}, \emptyset) = x_n$ and $c(\{x_1, x_3, ..., x_n\}, \emptyset) = x_1x_n$. By the definition of $\mathcal{G}^c$, we have $(x_1, x_2) \in \mathcal{G}^c$, i.e., $(x, y) \in \mathcal{G}^c$.

Next, we show $\mathcal{G}^c \subseteq \mathcal{G}^\rho$. Consider $(x, y) \in \mathcal{G}^c$ with $x \neq y$. If $x \notin c(xy, \emptyset)$, then $(x, y) \in \mathcal{G}^\rho$ (since $x \vdash \emptyset$). If there exists $A \subseteq X \backslash (xy)$ and $z \in A$ such that $c(A, \emptyset) = c(Axy, \emptyset) = z$ and $c(Ax, \emptyset) = xz$, then it can be inferred that $z \succ x$, $z \succ y$, and for all $z' \in A \backslash z$, $z \succ z'$. We can also infer that $(x, y) \in \mathcal{G}$ and for all $z' \in A$, $(x, z') \notin \mathcal{G}$. It follows that $\rho(z|A, \emptyset) = \rho(z|Ax, \emptyset)$, and thus $x \vdash A$. Since $x \notin c(Axy, \emptyset)$, we have $(x, y) \in \mathcal{G}^\rho$. Therefore, $\mathcal{G}^\rho = \mathcal{G}^c$.

For statement (2), we note that by the proof of the sufficiency part of Theorem 2, $(\pi, \mathcal{G}^\rho, \succ)$ represents the choice rule as a RCAR. It is also clear by our previous analysis that $\mathcal{G}^c \subseteq \mathcal{G}$. Thus statement (2) is true. □

References

AGRANOV, M., AND P. ORTOLEVA (2022): "Revealed Preferences for Randomization: An Overview," in *AEA Papers and Proceedings*, vol. 112, pp. 426–430. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

ANDERSON, J. R. (1996): *The Architecture of Cognition*. Harvard University Press.

ANDERSON, J. R., AND G. H. BOWER (1973): *Human Associative Memory*. Winston & Sons.

BLOCK, H. D., AND J. MARSCHAK (1960): "Random Utility Models: the Theory of Demand for Attributes," *Econometrica*, 28(4), 517–548.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2020): "Memory, Attention, and Choice," *Quarterly Journal of Economics*, 135(3), 1399–1442.

——— (2022): "Salience," *Annual Review of Economics*, 14(1), 521–544.

BRADY, R. L., AND J. REHBECK (2016): "Menu-Dependent Stochastic Feasibility," *Econometrica*, 84(3), 1203–1223.

BRANDT, C., C. P. DE MORTANGES, C. BLUEMELHUBER, AND A. C. VAN RIEL (2011): "Associative Networks: A New Approach to Market Segmentation," *International Journal of Market Research*, 53(2), 187–207.

CAMERER, C. F. (1997): "Progress in Behavioral Game Theory," *Journal of Economic Perspectives*, 11(4), 167–188.

CATTANEO, M. D., P. CHEUNG, X. MA, AND Y. MASATLIOGLU (2021): "Attention Overload," *arXiv preprint arXiv:2110.10650*.

CATTANEO, M. D., X. MA, Y. MASATLIOGLU, AND E. SULEYMANOV (2020): "A Random Attention Model," *Journal of Political Economy*, 128(7), 2796–2836.

CERREIA-VIOGLIO, S., D. DILLENBERGER, P. ORTOLEVA, AND G. RIELLA (2019): "Deliberately Stochastic," *American Economic Review*, 109(7), 2425–2445.

Cheung, P., and Y. Masatlioglu (2021): "Decision Making with Recommendation"," Discussion paper, working paper.

Cunha Jr, M., M. R. Forehand, and J. W. Angle (2015): "Riding Coattails: When Co-Branding Helps versus Hurts Less-Known Brands," *Journal of Consumer Research*, 41(5), 1284–1300.

Dardanoni, V., P. Manzini, M. Mariotti, and C. J. Tyson (2020): "Inferring Cognitive Heterogeneity from Aggregate Choices," *Econometrica*, 88(3), 1269–1296.

Dean, M., Ö. Kibris, and Y. Masatlioglu (2017): "Limited Attention and Status Quo Bias," *Journal of Economic Theory*, 169, 93–127.

Falmagne, J.-C. (1978): "Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce," *Journal of Mathematical Psychology*, 18(1), 1–14.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010): "Network Games," *The review of economic studies*, 77(1), 218–244.

Galeotti, A., and B. W. Rogers (2015): "Diffusion and Protection across a Random Graph," *Network Science*, 3(3), 361–376.

Geng, J. J., and M. Behrmann (2005): "Spatial probability as an attentional cue in visual search," *Perception & Psychophysics*, 67(7), 1252–1268.

Goyal, S. (2023): *Networks: An Economics Approach*. MIT Press.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *Econometrica*, 82(5), 1873–1912.

Gul, F., and W. Pesendorfer (2006): "Random Expected Utility," *Econometrica*, 74(1), 121–146.

Jackson, J., M. Stafford, and P. Smith (2009): "The Role of Bottom-up and Top-down Processes in Risk Perception: A Study of the Fear of Crime," *British Journal of Psychology*, 100(3), 633–652.

Jackson, M. (2008): *Social and Economic Networks*. Princeton University Press.

Keller, K. L. (1993): "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57(1), 1–22.

Kovach, M., and E. Suleymanov (2021): "Reference Dependence and Random Attention," *arXiv preprint arXiv:2106.13350*.

Kovach, M., and G. Tserenjigmid (2022): "Behavioral Foundations of Nested Stochastic Choice and Nested Logit," *Journal of Political Economy*, 130(9), 2411–2461.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2017): "When More is Less: Limited Consideration," *Journal of Economic Theory*, 170, 70–85.

Loewenstein, G. F., and T. O'Donoghue (1997): "Consumers, Plasticity, and the Design of Payday Loans," *Journal of Risk and Uncertainty*, 15(1), 75–94.

Luce, R. D. (1959): *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.

Manzini, P., and M. Mariotti (2014): "Stochastic Choice and Consideration Sets," *Econometrica*, 82(3), 1153–1176.

Masatlioglu, Y., and D. Nakajima (2013): "Choice by Iterative Search," *Theoretical Economics*, 8(3), 701–728.

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed Attention," *American Economic Review*, 102(5), 2183–2205.

Masatlioglu, Y., and E. A. Ok (2005): "Rational Choice with Status Quo Bias," *Journal of economic theory*, 121(1), 1–29.

——— (2014): "A Canonical Model of Choice with Initial Endowments," *Review of Economic Studies*, 81(2), 851–883.

McFadden, D. (1974): "A Theory of the Allocation of Time," *Journal of Political Economy*, 82(2), S111–S137.

Mogg, K., B. P. Bradley, C. Dixon, S. Fisher, H. Twelftree, and A. McWilliams (2004): "Trait Anxiety, Defensiveness and Selective Processing of Threat: An Investigation Using two Measures of Attentional Bias," *Personality and Individual Differences*, 36(7), 1359–1374.

Payne, J. W., J. R. Bettman, and E. J. Johnson (1993): *The Adaptive Decision Maker*. Cambridge University Press.

Raaijmakers, J. G., and R. M. Shiffrin (1981): "Search of Associative Memory," *Psychological review*, 88(2), 93.

Salant, Y., and A. Rubinstein (2008): "(A,f): Choice with Frames," *Review of Economic Studies*, 75(4), 1287–1296.

Simon, H. A. (1955): "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 69(1), 99–118.

Strom, C. G. J. (1965): "Maximum Likelihood Estimation of the Parameters of a Linear Regression System," *Annals of Mathematical Statistics*, 36(4), 1049–1053.

Sui, J., X. Gu, and S. Han (2012): "Attentional Control and the Self: The Self-Attention Network (SAN)," *Cognitive, Affective, & Behavioral Neuroscience*, 12(4), 706–718.

Teichert, T. A., and K. Schöntag (2010): "Exploring Consumer Knowledge Structures Using Associative Network Analysis," *Psychology & Marketing*, 27(4), 369–398.