

Markov Stochastic Choice*

Kremena Valkanova[†]

September 16, 2024

Abstract

We examine the effect of item arrangement on choices using a novel decision-making model based on the Markovian exploration of choice sets. This model is inspired by experimental evidence suggesting that the decision-making process involves sequential search through rapid stochastic pairwise comparisons. Our findings show that decision-makers following a reversible process are unaffected by item rearrangements, and further demonstrate that this property can be inferred from their choice behavior. Additionally, we provide a characterization of the class of Markovian models in which the agent makes all possible pairwise comparisons with positive probability. The intersection of reversible models and those allowing all pairwise comparisons is observationally equivalent to the well-known Luce model. Finally, we characterize the class of Markovian models for which the initial fixation does not impact the final choice and show that choice data reveals the existence and composition of consideration sets.

Keywords: stochastic choice, search, bounded rationality, attention.

*I would like to thank my advisors Nick Netzer and Jakub Steiner for their invaluable guidance, support and constructive comments. I am grateful to Carlos Alós-Ferrer, Adam Dominiak, Sean Horan, Matthew Kovach, Marco LiCalzi, Fabio Maccheroni, Massimo Marinacci, Yusufcan Masatlioglu, Ariel Rubinstein, Gerelt Tserenjigmid, Kemal Yildiz, as well as seminar and conference participants at Bern, Bocconi, Virginia Tech, BRIC (2022), NASM (2021) and RUD (2020) for insightful discussions and feedback. I thank MIT and Drew Fudenberg for the hospitality during the academic year 2019/2020 and to the Swiss National Science Foundation (grant No. P1ZHP1-184169) for the generous financial support. This research has been mainly conducted while I was affiliated with the Department of Economics at the University of Zurich.

[†]E-mail: kremena.valkanova@gmail.com

1 Introduction

We frequently encounter decision-making problems that involve locating, evaluating, and selecting the best option from a set of alternatives. Such decisions arise in various contexts, including choosing a product from a supermarket shelf or a vending machine, selecting a movie on a webpage, or deciding which article to read in a newspaper. Prior research demonstrates that the arrangement and adjacency of alternatives significantly impacts the decision-making process and final choices (Chen, Chen, & Tung, 2006; Drèze, Hoch, & Purk, 1994; Huang, Juaneda, Sénécal, & Léger, 2021; Keel & Padgett, 2015; Mattis, Groot Kormelink, Masur, Moeller, & van Atteveldt, 2024). This mechanism of influencing decisions coexists with well-established attention-seeking techniques aimed at increasing the selection probability of target items¹, as well as with manipulations of the choice set composition, exploiting menu-dependent preferences.²

Predicting the effects of item arrangement on choices is crucial for choice architects, even though their objectives may vary. For instance, academics and consumer researchers, focused on eliciting unbiased preferences, would seek to minimize external influences that could distort decision-making. Alternatively, benevolent social planners may design choice environments to nudge individuals towards healthier options or to reduce cognitive load by simplifying the choice problem without influencing the decision-making process. Finally, sellers and platform designers might exploit product positioning to increase the selection frequency of certain items, thereby maximizing profits.

Without a thorough understanding of the underlying decision process, choice problem designers might disrupt consumer choices by nudging decision-makers in unforeseen or unintended directions. Moreover, identifying which types of decision-makers are more susceptible to these positioning effects is crucial for developing more effective and personalized interventions. Finding that certain decision-makers are unaffected by these factors justifies the use of varied presentation formats to simplify and guide the choice process – for example, through list-based presentations or the integration of recommender systems on online platforms.

¹See Orquin and Mueller Loose (2013) for a review.

²Adler, Schöniger, Lichters, and Sarstedt (2024) provide a recent review over this vast literature in the marketing context.

To address these challenges, we introduce and analyze a choice model we call Markov Stochastic Choice (MSC). This model allows us to examine the effects of item positioning on decision-making and captures patterns consistently observed across multiple disciplines, such as economics, marketing, psychology, and visual neuroscience. These patterns include rapid pairwise comparisons between alternatives, a tendency to compare nearby options, the final fixation often aligning with the selected item, and the inherently stochastic nature of decision-making.³

The MSC decision-making process begins with an agent randomly selecting an alternative to view according to their initial beliefs and knowledge. In each period, the agent evaluates the current alternative against a competitor, transitioning to the competitor according to a probability distribution specific to the current alternative and independent of past transitions. This approach minimizes cognitive load by not relying on complete recollection of the past exploration. Transition probabilities are assumed to be consistent across different choice sets, meaning the relative likelihood of transitioning between any two alternatives remains unaffected by other items in the menu. The decision-making process may terminate in each period with a given stopping probability, after which the agent selects the most recently considered alternative. The described procedure corresponds to a discrete-time Markov chain, where each state represents an alternative in the choice set and the transitions are captured by eye movements, or saccades. With diminishing stopping probability, i.e. when there is no time pressure, we show that the choice function of the decision-maker converges towards the limiting distribution of the Markov chain.

Various factors can influence stochastic pairwise transitions. Eye fixations towards an item might be solely driven by its salience and attention-grabbing potential regardless of its value, as suggested by Reutskaja et al. (2011) and Milosavljevic, Navalpakkam, Koch, and Rangel (2012). Alternatively, probabilistic transitions between pairs of alternatives might reflect agent's noisy preferences. The literature on stochastic choice offers several

³Notable references include Russo and Rosen (1975), Henderson and Hollingsworth (1999), Rayner (1998), Chandon, Hutchinson, Bradlow, and Young (2009), Reutskaja, Nagel, Camerer, and Rangel (2011), Wedel and Pieters (2007), Pieters (2008), Noguchi and Stewart (2014), Shimojo, Simion, Shimojo, and Scheier (2003), and Armel, Beaumel, and Rangel (2008).

explanations for this noise, including random shocks to preferences, evidence accumulation, bounded rationality, and random mistakes (see [Agranov and Ortoleva \(2017\)](#)). Our model offers a simple framework to employ widely used models of binary stochastic choice such as the Drift-Diffusion Model ([Ratcliff, 1978](#)) to study multi-alternative choice.

We first consider nudging interventions involving restrictions of certain pairwise comparisons, achieved by either placing two items too far apart or not presenting them as related items in a recommender system, while still keeping the set of reachable items the same. We find that the types of decision makers who are not influenced by such presentations of the choice problem are characterized by following a reversible Markov process, i.e. the average number of saccades from one alternative to another should be the same as in the reverse direction in the long run. Since the choices are robust to any comparability restriction, we could infer the reversibility property if we find that no rearrangement of the choice set leads to a change in the stochastic choices. What is more, the choice function itself reveals the reversibility property of the rationalizing model. This is the case when there is no cycle of alternatives for which the relative choice probabilities between each of the subsequent pairs is higher in the current choice set than in the corresponding binary sets.

Next, we identify decision makers with greater cognitive capabilities who can make all pairwise comparisons directly regardless of arrangement, i.e. their choices are rationalizable with an MSC model we call fully comparable. The corresponding choice function is such that any pair, for which the relative choice probability in the current choice set differs from that in binary sets, belongs to a cycle where all differences in relative choice probabilities between the current and binary menus have the same sign. We term the choice function between such pairs as “bounded in a cycle”. A violation of this property would imply that the decision-maker was not able to compare the pair directly.

Interestingly, when considering the intersection of reversible and fully comparable MSC models, we find that they are characterized by two well-known axioms: positivity and Independence of Irrelevant Alternatives (IIA). Thus, this special class of MSC models is observationally equivalent to the well-known Luce model ([Luce, 1959](#)). Additionally, we demonstrate that the relative transition probabilities between each pair are determined solely by the relative utility of the items and are not influenced by other factors such as salience.

Finally, we analyze the potential to influence decision-making by altering the initial fixation probability. In practical applications, increasing the likelihood of starting the decision-making process at a specific target alternative can be achieved through various techniques, such as central placement, eye-catching packaging, default or sponsored item displays, and effective advertising (Chandon et al., 2009; Orquin & Mueller Loose, 2013). Our model predicts that under time pressure, the initial fixation can significantly affect final choices. When agents have unlimited time, such nudging techniques are only effective if distinct consideration sets exist, i.e. the Markov process is reducible. We demonstrate that for a stochastic choice function to be rationalizable with an irreducible model, a cycle must exist involving the entire menu where all subsequent pairs are bounded in a cycle. Additionally, the stochastic choice function reveals the precise communicating classes, making these interventions predictable and effective.

The remainder of the paper is structured as follows: Section 2 defines the choice model. Section 3 analyzes the impact of comparability restrictions and identifies reversible and pairwise comparable models. Section 4 examines the effect of initial fixation probability on final choices and characterizes the irreducible MSC model. Section 5 reviews the relevant literature. Section 6 concludes.

2 Model

Let N be a finite set of all alternatives. A menu (or choice set) M is a non-empty subset of N . The set of all menus is denoted by \mathcal{N} and the set of all non-empty subsets of a menu M is denoted by \mathcal{M} .

2.1 Stochastic choice functions

A stochastic choice function is a mapping $p : N \times \mathcal{N} \rightarrow [0, 1]$ such that $\sum_{i \in M} p(i, M) = 1$ and $i \notin M$ implies $p(i, M) = 0$. The function $p(i, M)$ is interpreted as the probability of choosing alternative i from a menu M . We denote a row vector of choice probabilities from a menu M by $\mathbf{p}(M)$. The stochastic choice function on all menus is denoted by $\mathbf{p}(\mathcal{N})$. A stochastic choice function is positive if $p(i, M) > 0$ for all $i \in M$ and all $M \in \mathcal{N}$.

Let $\delta_{ij}(\mathbf{p}(M)) = p(i, M)p(j, \{i, j\}) - p(i, \{i, j\})p(j, M)$ be a weighted difference in choice probabilities from menu M between two alternatives i and j . A cycle of alternatives $C(M)$ on $M = \{i_1, i_2, \dots, i_{|M|}\}$ is a set of ordered pairs such that $C(M) = \{(i_1, i_2), (i_2, i_3), \dots, (i_{|M|}, i_1)\}$. The set of all cycles on M is denoted by $\mathcal{C}(M)$.

Definition 1. *A cycle of alternatives $C(M')$ is called sign-consistent on $M' \in \mathcal{M}$ w.r.t. $\mathbf{p}(M)$ if $sgn(\delta_{ij}(\mathbf{p}(M))) = sgn(\delta_{kl}(\mathbf{p}(M)))$ for all $(i, j), (k, l) \in C(M')$.*

The set of all sign-consistent cycles on all subsets of M w.r.t. $\mathbf{p}(M)$ is given by $\mathcal{C}^S(\mathbf{p}(M))$. To indicate the sign of $\delta_{ij}(\mathbf{p}(M))$, we use the notation $\mathcal{C}_0^S(\mathbf{p}(M)), \mathcal{C}_+^S(\mathbf{p}(M)), \mathcal{C}_-^S(\mathbf{p}(M))$, respectively.

2.2 General choice procedure

A decision maker evaluates the alternatives in a given menu M as follows. The agent has an initial fixation probability distribution $\boldsymbol{\pi}(M)$ over the alternatives in the menu, determining the probability of starting the exploration process at each alternative. At time period $\tau = 0$, the agent draws an alternative $i \in M$, which becomes the current best alternative.

In the next period, alternative i is compared to another available alternative j . If j wins the comparison, the decision maker transitions from i to j , making j the best alternative at period $\tau = 1$. These transitions are stochastic, with the probability of transitioning from i to j denoted by $q_{ij}(M)$. A transition probability matrix $Q(M)$ contains all pairwise transition probabilities.

The transition probabilities may be influenced by decision makers' preferences and cognitive constraints, as well as the presentation of the choice problem. These factors, in turn, are affected by items' salience, positioning, and default option status. Consequently, we assume that the transition probabilities are time-independent, i.e. revisiting alternatives does not affect the probability to make subsequent transitions. However, we allow transition probabilities between each pair of alternatives to be menu-dependent, i.e. $q_{ij}(M) \neq q_{ij}(M')$. We impose the following three assumptions on the matrices $Q(M)$ for all $M \in \mathcal{N}$.

Assumption 1 (Prolonged consideration). *For all $M \in \mathcal{N}$ and $i, j \in M$ holds $q_{ii}(M) = 1 - \sum_{j \neq i} q_{ij}(M) > 0$.*

The assumption of prolonged consideration posits that there is a positive probability that no transition occurs in a given period, regardless of the current best alternative. Furthermore, it asserts that $Q(M)$ is a right stochastic matrix.

Assumption 2 (Pairwise comparability on binary sets). *For all $i, j \in N$, $q_{ij}(\{i, j\}) = 0$ implies that $q_{ji}(\{i, j\}) > 0$.*

Pairwise comparability on binary sets means that when choosing between two alternatives, the decision maker can transition in at least one direction, i.e. the items are comparable. It implies that cognitive capacity should not limit transitions when evaluating only a pair of alternatives.

Assumption 3 (Transition ratio independence of irrelevant alternatives (TR-IIA)). *For all $M \in \mathcal{N}$ and $i, j \in M$ holds $q_{ij}(\{i, j\})q_{ji}(M) = q_{ji}(\{i, j\})q_{ij}(M)$.*

The TR-IIA assumption sets a minimal consistency requirement in decision-making across choice sets: the ratio of transition probabilities remains unaffected when the menu expands. This assumption arises when transition probability depends on the utility difference between two items and their distance – whether physical (positioning on a shelf or screen) or subjective (similarity or ease of comparison).⁴ Since these components remain unchanged by adding more items to the choice set, the transition ratio between the pair remains constant.

The decision-making procedure is repeated analogously in consecutive periods. The process is terminated with probability $\alpha \in (0, 1)$ in each period. When the process stops, the agent chooses the last considered alternative.

The choice procedure is equivalent to a finite-state discrete-time Markov chain with each state representing one alternative in menu M , initial distribution $\boldsymbol{\pi}(M)$, transition probability matrix $Q(M)$, and stopping probability α . The stochastic choice function generated by the model for each menu takes the following form:

$$(1) \quad \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \sum_{\tau=0}^{\infty} \alpha \boldsymbol{\pi}(M) (1 - \alpha)^{\tau} Q(M)^{\tau},$$

⁴See for example Russo and Rosen (1975), Reutskaja et al. (2011) and Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini (2022).

where each term in the sum represents the probability that the process stops at a given period multiplied by the probability of each alternative to be the current best at that time period.

2.3 Choice procedure without time constraints

Consider the case in which the probability that the decision process is terminated approaches zero, hence the pressure to make a decision diminishes. If the underlying Markov chain is irreducible and aperiodic⁵, hence ergodic, a limiting distribution exists and it is the unique stationary distribution. Naturally, if the stopping probability approaches zero, the generated stochastic choice function converges to the stationary distribution of the Markov chain and does not depend on the initial distribution.

Proposition 1. *A stochastic choice function $\rho(\alpha, \pi(M), Q(M))$ generated by an ergodic Markov chain converges to its stationary distribution as $\alpha \rightarrow 0$.*

Proof. See Appendix A.1. □

We denote the generated stochastic choice function as $\alpha \rightarrow 0$ by $\rho(\pi(M), Q(M))$.

Definition 2. *A stochastic choice function $\rho(\mathcal{N})$ is rationalizable with an MSC model if there exists a tuple $\langle Q(M), \pi(M) \rangle$ satisfying Assumptions 1 – 3 such that $\rho(M) = \rho(\pi(M), Q(M))$ for all $M \in \mathcal{N}$, where $\rho(\pi(M), Q(M))(I - Q(M)) = \mathbf{0}$.*

The above equality follows immediately from Proposition 1. Note that if the Markov chain is irreducible, the choice function is unique. If it is not irreducible, each communicating class has its own unique stationary distribution and hence, the choice function depends on the initial distribution $\pi(M)$.

⁵Note that the assumption of prolonged consideration made earlier ensures that the Markov chain of any MSC model is aperiodic. Recall that a communicating class of a Markov chain over M is defined as a set of states $M' \in \mathcal{M}$, which communicate, i.e. for each pair of states $i, j \in M'$ there is a sequence of states $k_1, k_2, \dots, k_h \in M'$ such that $q_{ik_1}(M)q_{k_1k_2}(M)\dots q_{k_hj}(M) > 0$. A Markov chain is irreducible when it has a single communicating class, hence the process can reach any state irrespectively of the initial state. Since the Markov chain of a MSC model is finite, irreducibility implies positive recurrence.

3 Comparability restrictions

In this section, we examine how decision makers' inability to make certain pairwise comparisons affects their final choices. Such comparability restrictions between pairs of alternatives may arise from cognitive limitations, especially when many alternatives are presented simultaneously, as in the supermarket example mentioned in the Introduction. This concept is supported by empirical evidence showing that jumps between fixations mostly occur between nearby alternatives (Chandon et al., 2009). Hence, different arrangements of the same alternatives can impact their pairwise comparability. Figure 1 illustrates such a scenario, where the decision maker cannot directly compare diagonally positioned pairs when they are displayed on a grid.



Figure 1: Limitations in a decision maker's perception can lead to varying pairwise comparability restrictions depending on the item positioning.

Additionally, a choice problem designer might restrict the visibility of the relevant items, limiting the decision maker's ability to compare certain pairs. For example, changes in the network of related products of a platform's recommendation system can influence the final purchasing decision, even if the decision maker would eventually consider the same set of items without time constraints.

We formalize the concept of comparability restrictions below. We use the notation \odot for element-wise matrix multiplication.

Definition 3. A comparability restriction of an MSC model $\langle Q(M), \pi(M) \rangle$ is a symmetric matrix $R(M)$ with elements $r_{ij}(M) \in \{0, c\}$ for $i \neq j$ and $c > 0$ such that $R(M) \odot Q(M)$ is a right stochastic matrix with the same communicating classes as $Q(M)$.

In other words, comparability restrictions prevent the decision maker from making a

direct transition between pairs of alternatives, while keeping the reachable set of alternatives unchanged and scaling the transition probabilities between non-restricted pairs with the same factor. We say that an MSC model is *robust to a comparability restriction* $R(M)$ if $\rho(\pi(M), Q(M)) = \rho(\pi(M), R(M) \odot Q(M))$.

In the following, we identify and characterize those classes of MSC models, which are robust to any comparability restrictions. This is useful for two main reasons. First, it justifies the use of different presentation formats and simplifications of the choice problem (say, presenting items as a list instead on a grid) in cases when decision makers are not affected by comparability restrictions. Second, comparability restrictions can aid decision-makers in overcoming behavioral biases such as the attraction effect as discussed in the example below.

Example 1. *The attraction effect occurs when adding an asymmetrically dominated alternative to a choice set increases the choice probability of the dominating alternative. This phenomenon has been consistently documented in a variety of choice contexts and across different species (Castillo, 2020; Huber, Payne, & Puto, 1982). The attraction effect is inconsistent with standard stochastic choice axioms such as regularity and independence of irrelevant alternatives. The MSC model can accommodate this behavioral bias and provide suggestions for its mitigation.*

More specifically, in any MSC model over $N = \{i, j, k\}$ with strictly positive transition probabilities except $q_{ik}(\{i, k\}) = q_{ik}(N) = 0$ (i.e., the target is alternative i and its decoy is alternative k) the relative choice frequency of the target vs. the competitor increases when the decoy is present. Moreover, the target is chosen more frequently from the triple than from the pair in absolute terms, $\rho(i, Q(N)) > \rho(i, Q(\{i, j\}))$, if and only if $q_{ki}(N) > q_{ji}(N)$, i.e. the decision maker is more likely to transition to the target from the decoy than from the competitor. Thus, the attraction effect should vanish if the comparability between the target and the decoy is restricted, i.e. $r_{ik}(N) = r_{ki}(N) = 0$. More details and proofs can be found in Appendix B.1.

3.1 Reversible MSC models

In this subsection we characterize the Markovian decision-making process of agents whose choices are robust to comparability restrictions. This class of MSC models satisfies a well-known property of Markov chains called *reversibility* (Levin, Peres, & Wilmer, 2017). A Markov chain is reversible if it satisfies the following detailed balance conditions

$$(2) \quad q_{ji}(M)\rho(j, Q(M)) = \rho(i, Q(M))q_{ij}(M), \forall i, j \in M.$$

The detailed balance equations postulate that the flow of probability mass is balanced for each pair of states. In our setting, detailed balance means that the average number of saccades from i to j should be the same as the saccades from j to i in the long run.⁶ We call an MSC model reversible on M if the corresponding Markov chain with states M satisfies reversibility. We say that an MSC model is reversible if it is reversible on all $M \in \mathcal{N}$.

The following theorem is our first main result. It characterizes reversible MSC models in terms of comparability restrictions and decision maker's choices from a given menu.

Theorem 1. *The following are equivalent: A stochastic choice function $\mathbf{p}(M)$ is*

- (i) *rationalizable only by reversible MSC models on M ,*
- (ii) *robust to any comparability restrictions of a rationalizing model,*
- (iii) *such that $\mathcal{C}^S(\mathbf{p}(M)) = \mathcal{C}_0^S(\mathbf{p}(M))$.*

Proof. See Appendix A.3. □

Statements (i) and (ii) of the above theorem imply that if the decision-making process is reversible, restricting some pairwise comparisons does not affect final choices. Additionally, if the process is reversible without comparability restrictions, different arrangements of alternatives, as shown in Figure 1, do not impact choice frequencies, even if the decision maker

⁶A necessary and sufficient condition for reversibility of a Markov chain is the Kolmogorov's criterion. It states that a Markov chain is reversible if and only if for all cycles $C(M') \in \mathcal{C}(M')$ and all $M' \in \mathcal{M}$ holds

$$(3) \quad \prod_{(k,l) \in C(M')} q_{kl}(M) = \prod_{(k,l) \in C(M')} q_{lk}(M).$$

can only compare nearby options. Similarly, these decision makers' choices are unaffected by the number of related products on an e-commerce website, as long as they can access the same items from any starting point.

Following Theorem 1, an analyst can use comparability restrictions to determine if the decision-making process is reversible on all $M \in \mathcal{N}$. To this end, the analyst needs to verify that the final choices do not change in response to *any* of the possible comparability restrictions. When the stochastic choice function is positive, then the analyst only needs to check the effect of comparability restrictions when the menu equals the complete set of alternatives $M = N$. This is because if stochastic choice data on N is rationalizable by a reversible MSC model, it can be shown using TR-IIA and Kolmogorov's condition that the MSC model is reversible on all subsets.

Statement (iii) offers an alternative way to identify reversible MSC models by examining the properties of the stochastic choices. It states that there is no sign-consistent cycle of alternatives for which the ratio of choice probabilities is either strictly higher or strictly lower in the larger choice set compared to the binary sets for each pair of alternatives in the cycle. The characterization result is particularly useful if an analyst wants to verify reversibility, but cannot control comparability restrictions, such as when they are due to the decision maker's unobservable cognitive limitations.

The two characterizations of reversible MSC models can complement each other to make predictions and uncover the true generating model's properties. The equivalence of Theorem 1(i) and (iii) implies that if only choice data for a particular menu is observed, comparability restrictions can simplify the choice problem without affecting the decision maker's choices. Additionally, if a stochastic choice function violates Theorem 1(iii) and is rationalizable by a non-reversible and a reversible model, comparability restrictions can help distinguish between the two. If any such restriction changes the final choices, the true generating model is non-reversible. The following example illustrates Theorem 1.

Example 2. Consider the choice function from binary sets given in Table 1 and let $\mathbf{p}(\{i, j, k, l\}) = \mathbf{p}(N) = (0.2, 0.2, 0.4, 0.2)$ ⁷. We first verify that the stochastic choice function satisfies Theo-

⁷Note that although this stochastic choice function and the ones used in Examples 2-4 all satisfy regularity, it is not a necessary condition for rationalizability with an MSC model as opposed to other models of

Table 1: Choice probabilities of alternatives i, j, k, l (rows) from binary choice sets (columns) used in Examples 2-4.

	$\{i, j\}$	$\{i, k\}$	$\{i, l\}$	$\{j, k\}$	$\{j, l\}$	$\{k, l\}$
i	0.5	0.5	0.5			
j	0.5			0.6	0.5	
k		0.5		0.4		0.4
l			0.5		0.5	0.6

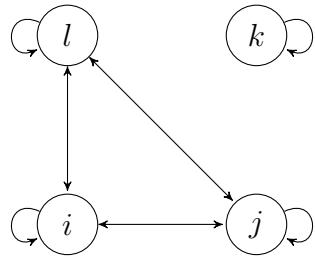


Figure 2: A reversible MSC model rationalizing the choice function given in Example 2.

rem 1(iii), i.e. we need to show $\mathcal{C}_-(\mathbf{p}(N)) = \mathcal{C}_+(\mathbf{p}(N)) = 0$ for any $M' \in \mathcal{N}$. The only pairs of alternatives whose choice ratio from N do not equal their choice ratio from the binary set are (i, k) , (j, k) , and (k, l) and their reverse pairs, i.e. $\delta_{ik}(\mathbf{p}(M)) \neq 0$, $\delta_{jk}(\mathbf{p}(M)) \neq 0$, and $\delta_{kl}(\mathbf{p}(M)) \neq 0$. Since all cycles containing these pairs contain two reverse pairs, the condition in Theorem 1(iii) is trivially satisfied.

As we prove in Appendix B.2, an MSC model that has $\pi(k, M) = 0.4$ and positive off-diagonal transition probabilities $q_{ij}(M) = q_{il}(M) = q_{ji}(M) = q_{jl}(M) = q_{lj}(M) = 0.1$ rationalizes the stochastic choice data. Figure 2 visualizes the transitions occurring with positive probability. Note that this is the greatest number of possible direct comparisons consistent with the choice data. It is evident that this model is reversible, because it holds that $q_{ij}(M)q_{jl}(M)q_{li}(M) = q_{ji}(M)q_{lj}(M)q_{il}(M)$ and in all other cycles there are transitions happening with zero probability.

Finally, we can verify that the rationalizing model is robust to comparability restrictions. stochastic choice such as random utility models.

For example, if we consider a comparability restriction with $r_{il}(M) = r_{li}(M) = 0$, the model $\langle R(M) \odot Q(M), \pi(M) \rangle$ also rationalizes the choice data.

Comparability restrictions are assumed to have a uniform effect on all transition probabilities between non-restricted pairs. This assumption is plausible if the manipulation of the choice problem does not affect the accessibility and salience of non-restricted pairs. However, choices generated by reversible models are robust to an even larger class of comparability restrictions.

Definition 4. A weak comparability restriction of an MSC model $\langle Q(M), \pi(M) \rangle$ is a symmetric matrix $R(M)$ with non-negative off-diagonal elements such that $R(M) \odot Q(M)$ is a right stochastic matrix with the same communicating classes as $Q(M)$.

In essence, weak comparability restrictions can have an asymmetric effect on non-restricted pairs. The symmetry of the matrix $R(M)$ ensures TR-IIA holds for all pairs.

Proposition 2. If an MSC model is reversible, the generated stochastic choice function $p(\mathcal{N})$ is robust to all weak comparability restrictions.

Proof. See Appendix A.4. □

In other words, agents who follow a reversible decision-making process cannot be nudged to choose certain alternatives more frequently by imposing comparability restrictions, even if these affect non-restricted pairs asymmetrically. Thus, the choices are robust across various presentations of the decision problem. For example, the choices of such decision makers are unaffected not only by reducing the number of possible comparisons, but also by reshuffling the positions of the items, leading to asymmetric changes in transition probabilities due to comparability restrictions. We conclude this subsection with the following example illustrating Proposition 2.

Example 2 (continued). Observe that an MSC model in which the previously discussed comparability restriction between i and l affects asymmetrically the likelihood to make a transition between i and j such that $r_{ij}(M) = 2$, while keeping all other off-diagonal elements of the $R(M)$ matrix equal to 1, will rationalize the same stochastic choice function.

3.2 Pairwise and fully comparable MSC models

In the previous section, we establish that choices generated by a reversible MSC model cannot be manipulated with comparability restrictions. Another type of decision maker unaffected by menu rearrangements is one whose cognitive abilities allow all pairwise comparisons. In this section, we characterize the choices of such decision makers, captured by classes of MSC models we call pairwise and fully comparable.

Definition 5. *An MSC model is pairwise comparable if it holds for all $i, j \in M$ and all $M \in \mathcal{N}$ that $q_{ij}(M) = 0$ implies $q_{ji}(M) > 0$. An MSC model is fully comparable if $q_{ij}(M) > 0$ for all $i, j \in M$ and all $M \in \mathcal{N}$.*

This definition of pairwise comparability is a generalization of Assumption 2 about pairwise comparability on binary sets to all menus. If the model is fully comparable, the agent makes all possible transitions with positive probability. Our characterization results of these classes of MSC models are based on the existence of particular sign-consistent cycles w.r.t. the stochastic choice function.

Definition 6. *The stochastic choice function $\mathbf{p}(M)$ is bounded in a cycle over a pair of alternatives $(i, j) \in M$ whenever if $\delta_{ij}(\mathbf{p}(M)) \neq 0$, $\exists C(M')$ with $(i, j) \in C(M')$ and $C(M') \in \mathcal{C}_{+/-}^S(\mathbf{p}(M))$ for $M' \in \mathcal{M}$.*

Furthermore, the characterization of fully comparable MSC models requires in addition positivity.

Theorem 2. *The following are equivalent: A stochastic choice function $\mathbf{p}(M)$ is*

- (i) rationalizable by a pairwise (fully) comparable MSC model on M ,
- (ii) (positive and) bounded in a cycle over all pairs of alternatives.

Proof. See Appendix A.5. □

Theorem 2 implies that if a choice function violates condition (ii), the decision-making process necessarily took place under comparability restrictions. What is more, the violation of the condition for an individual pair implies that no rationalizing model would assign a positive transition probability between the pair as we show with the next proposition.

Proposition 3. *A stochastic choice function $\mathbf{p}(M)$, which is not bounded in a cycle over a pair $i, j \in M$, is only rationalizable by MSC models with $q_{ij}(M) = 0$.*

Proof. See Appendix A.6. □

We illustrate Theorem 2 with the following example.

Example 3. *Consider the choice function from binary sets given in Table 1 and let $\mathbf{p}(\{i, j, k, l\}) = \mathbf{p}(N) = (0.25, 0.28, 0.2, 0.27)$. We first verify that $\mathbf{p}(N)$ satisfies the condition in Theorem 2(ii). Note that the cycle $\{(i, k), (k, j), (j, l), (l, i)\} \in \mathcal{C}_+^S(\mathbf{p}(N))$. We need to show that the remaining pairs (i, j) and (k, l) belong to a sign-consistent cycle. Since $\{(i, k), (k, j), (j, i)\} \in \mathcal{C}_+^S(\mathbf{p}(N))$ and $\{(i, k), (k, l), (l, i)\} \in \mathcal{C}_+^S(\mathbf{p}(N))$, the condition is satisfied. According to the theorem, the choice function is rationalizable by a fully comparable MSC model. Appendix B.3 contains an example of one such rationalizing model.*

3.3 Reversible and fully comparable MSC models

The classes of reversible and fully comparable MSC models are independent, because in a reversible MSC model there might be a zero transition probability between some pairs (see Example 2) and fully comparable models need not be reversible (see Example 3). We now focus on the intersection of the two classes of models. Following Theorems 1(iii) and 2(ii), the generated stochastic choice function should be such that $\delta_{ij}(\mathbf{p}(M)) = 0$, or alternatively $p(i, \{i, j\})p(j, M) = p(j, \{i, j\})p(i, M)$, $\forall i, j \in M$ and $M \in \mathcal{N}$. This property is known as Independence of irrelevant alternatives (IIA) and implies that enlarging the choice set does not have an effect on relative choice probabilities. We state the characterization result below.

Theorem 3. *The following are equivalent: A stochastic choice function $\mathbf{p}(M)$ is*

- (i) *rationalizable by a fully comparable and reversible MSC model on M ,*
- (ii) *positive and independent of irrelevant alternatives,*
- (iii) *rationalizable by an MSC model such that $\forall i, j \in M$, $\frac{q_{ij}(M)}{q_{ji}(M)} = \frac{u(j)}{u(i)}$, where $u : N \rightarrow \mathbb{R}_{++}$.*

Proof. See Appendix A.7. □

We gain several important insights from Theorem 3. First, we learn that the well-known Luce model, also referred to as multinomial logit, is observationally equivalent to a reversible and fully comparable MSC model. Recall that a stochastic choice function $\mathbf{p}(\mathcal{N})$ is a Luce rule if there exists a function $u : N \rightarrow \mathbb{R}_{++}$ such that for all $M \in \mathcal{N}$ and $i \in M$

$$p(i, M) = \frac{u(i)}{\sum_{j \in M} u(j)}.$$

In his seminal work, Luce (1959) shows that positivity and IIA characterize the Luce model. Therefore, Theorem 3 provides a procedural justification for the Luce model. Basically, stochastic choice consistent with the Luce model arises whenever the search dynamics follow a Markov process with symmetric average transitions and positive transition probabilities between all pairs of alternatives.

Therefore, choices consistent with reversible and fully comparable models can be considered most unbiased, as all pairwise comparisons are possible, determined solely by utility, and unaffected by the presentation of the choice problem. Moreover, if the choice data satisfies Theorem 1(iii), but violates IIA, we can deduce that the decision making process is reversible but the agent is unable to make all pairwise comparisons.

Finally, Theorem 3(iii) provides an interesting insight in the exploration mechanism that corresponds to a reversible and fully comparable MSC model. Specifically, the ratio of transition probabilities between each pair of alternatives equals the ratio of their utilities, where the utility function is the same as in the Luce model. This feature suggests that transitions between alternatives depend solely on the utility, not on factors like salience, positioning, or cognitive restrictions. In contrast, violations of IIA and positivity indicate that these factors do influence decisions, making the agent susceptible to choice manipulation through the presentation of the decision problem.

4 Initial fixation

Apart from comparability restrictions, the presentation of the decision problem affects the exploration and evaluation process through the first fixation, captured by the initial distribution $\pi(M)$. It is evident by the form of the generated choice function given in (1) that

our model predicts that the initial fixation can have significant and long-lasting effects on the final choices when there is time pressure, which is also in line with existing experimental evidence (Armel et al., 2008; Atalay, Bodur, & Rasoloforison, 2012; Reutskaja et al., 2011). Thus, a choice problem designer could steer the choices of such decision makers towards a target alternative by influencing the initial fixation probability.⁸ Without time constraints, the effectiveness of such interventions is more nuanced and depends on the pairwise comparability.

Following Proposition 1, when the pairwise evaluation process can eventually reach all alternatives in the menu from any starting point, with no limit on the number of transitions (indicating an irreducible Markov process⁹), final choices become independent of the initial fixation distribution, i.e. the choice function is robust to any initial fixation distributions.

We now characterize irreducible MSC models. In order to do so, we need to relax appropriately the property in Theorem 2(ii) since pairwise comparable models are a subclass of irreducible MSC models.

Theorem 4. *The following are equivalent: A stochastic choice function $\mathbf{p}(M)$ is*

- (i) *rationalizable by an irreducible MSC model on M ,*
- (ii) *robust to any initial fixation probability distributions,*
- (iii) *bounded in a cycle over all pairs of some cycle $C(M)$.*

Proof. See Appendix A.8. □

The reducibility of the decision making process implies the existence of separate communicating classes, or consideration sets. This aligns with an agent's approach of first focusing on a subcategory of alternatives and then selecting an item from that subcategory without

⁸In order to ensure the effectiveness of such an intervention our model shows that it is not only necessary that initial fixation probability of the target increases, but also that the fixation probabilities of all other competitors weakly decrease. A detailed proof is available upon request.

⁹A fully comparable MSC model is an example of an irreducible model, whereas the model depicted in Figure 2 is reducible – if the process starts at alternative k , none of the other alternatives in the menu can be reached and vice versa.

exploring the rest. Such partitions of the menu could stem from the decision maker's subjective perception of item attributes, such as taste (sweet vs. salty), brand, or ingredients. This behavior is consistent with experimental marketing literature, which identifies heuristics used by decision makers to construct consideration sets by focusing on subsets of items with specific attributes (see Hauser (2014) and references therein).

The partitioning of a menu into subsets can also be influenced externally by the presentation of the decision problem. For instance, in a menu composed of snacks, a designer can emphasize the division between salty and sweet snacks by placing them on different shelves in a supermarket. On an e-commerce platform with a recommendation system, the algorithm might categorize items into distinct subgroups and suggest different sets of related products based on the first item the user views.

If the decision making process is reducible, i.e. the underlying Markov chain has multiple communicating classes, the first fixation determines the subset of reachable alternatives. Thus, choices can be manipulated through the initial distribution, unless the probability to start in each communicating class remains constant. Increasing the likelihood of starting in a particular communicating class boosts the choice probability of all alternatives in that class while preserving their relative choice probabilities within the class. Thus, a social planner can nudge decision makers to choose a target alternative more often by directing attention to any alternative in the same communicating class as the target.

Therefore, inferring the irreducibility property of an MSC model from the choice function is advantageous in two ways. It helps evaluate the effectiveness of various attention-grabbing methods in increasing the choice probability of a target alternative, and it reveals whether the menu is perceived as fragmented or coherent.

Note that if a stochastic choice function fulfils Theorem 4(iii), this does not necessarily mean that it was generated by an irreducible model, but only that it is consistent with such a model. An analyst can infer the reducibility properties of the true model by observing changes in final choice probabilities resulting from alterations in the initial fixation probability of different communicating classes. In order to do that effectively, the analyst would ideally know which alternatives belong to the same communicating class. As we showed in Proposition 3, the stochastic choice function itself provides insights about the possible

pairwise comparisons and thus about the composition of the communicating classes. Thus, Theorem 4(ii) and (iii) can complement each other in order to infer the irreducibility property.

If choice data reveals the existence of consideration sets when applying Theorem 4, sellers can infer which products fall into the same consideration set as their own product, thus inferring their direct competitors from consumers' perspective. Furthermore, this insight aids choice problem designers in making better targeted interventions and predicting the effects of a change in the initial fixation probability. We conclude this section with the following example.

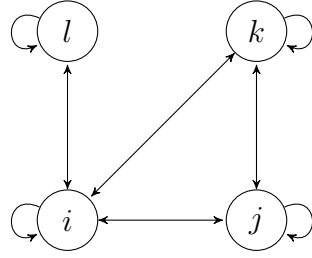


Figure 3: An irreducible MSC model rationalizing the choice function given in Example 4.

Example 4. Consider the choice function from binary sets given in Table 1. Let the choice function from the menu $M = \{i, j, k, l\}$ be $\mathbf{p}(M) = (0.24, 0.3, 0.22, 0.24)$. We first verify that $\mathbf{p}(M)$ satisfies Theorem 4(iii). Observe that $\delta_{il}(\mathbf{p}(M)) = 0$ and consider the cycle $C(M') = \{(j, i), (i, k), (k, j)\} \in \mathcal{C}_+^S(\mathbf{p}(M))$. Thus, the cycle $C(M) = \{(j, i), (i, l), (l, i), (i, k), (k, j)\}$ fulfills the requirement and the stochastic choice function is rationalizable by an irreducible MSC model such as the one contained in Appendix B.4 and depicted in Figure 3.

We can verify that the rationalizing model from Figure 3 has the highest number of possible transitions using Proposition 3. In order to show that any rationalizing model assigns zero probability to transitions between j and l , note first that $\delta_{jl}(\mathbf{p}(M)) > 0$, so we need to check whether the pair belongs to any sign-consistent cycle according to $\mathbf{p}(M)$. We consider the cycles $\{(j, k), (k, l), (l, j)\}$ and $\{(k, l), (l, j), (j, i), (i, k)\}$, because the pair cannot be part of a sign-consistent cycle with (j, l) since $\delta_{il}(\mathbf{p}(M)) = 0$. We can easily see that these are no sign-consistent cycles since $\delta_{lj}(\mathbf{p}(M)) < 0$ and $\delta_{kl}(\mathbf{p}(M)) > 0$. We can use the same argument for the pair (k, l) to prove that the transition probability should be zero.

5 Related literature

Researchers across various disciplines have recognised the advantages of incorporating Markov processes in modeling choice behavior. The literature emphasizes the tractability of these models, their superior performance over traditional frameworks in predicting choices, and their ability to generalize popular random utility-based discrete choice models, such as the multinomial logit model (Baldassi, Cerreia-Vioglio, Maccheroni, Marinacci, & Pirazzini, 2020; Blanchet, Gallego, & Goyal, 2016; Cerreia-Vioglio et al., 2022; Ragain & Ugander, 2016). We relate to this literature by employing a general Markovian model to study the effects of choice architecture and to reveal unobservable pairwise comparison patterns from choice behavior.

Furthermore, our characterization results make a technical contribution related to the problem of inverting the stationary distribution of a Markov chain studied in Morimura, Osogami, and Idé (2013) and Kumar, Tomkins, Vassilvitskii, and Vee (2015). This existing literature attempts to recover the compete transition probability matrix from the stationary distribution assuming that either the graph or the rate of reaching one state from another is (partially) known. We identify the properties of a stationary distribution that guarantee the existence of a corresponding Markov chain within the classes of reversible, pairwise comparable, and irreducible chains. Notably, we achieve this under minimal assumptions concerning the consistency of Markov chains over subsets of states.

The MSC model builds on a vast theoretical literature incorporating various behavioral limitations into models of choice behavior, in particular to search, reference dependence, and limited consideration. In much of the theoretical literature featuring sequential menu exploration, the search order is specified by the choice problem itself in a form of a list (Horan, 2010; Papi, 2012; Rubinstein & Salant, 2006; Zhang, 2017) or unobservable network of related products in a recommender system (Masatlioglu & Nakajima, 2013; Masatlioglu & Suleymanov, 2021). In some choice models such as those by Caplin and Dean (2011), Apesteguia and Ballester (2013), Dutta (2020) and Demirkan (2024) agent's ability to partition and explore the choice set is not observable and this literature is mainly concerned with the problem of rationalizability. Our MSC model adds to this literature by incorporat-

ing many of the empirically documented decision-making patterns motivating the existing models and offering a simple framework to analyse the effects of various presentation formats.

Reference points have played an important role in behavioral choice models, assumed to be static and exogenous (Masatlioglu & Ok, 2005; Salant & Rubinstein, 2008; Tversky & Kahneman, 1991), endogenous and menu-dependent (Ok, Ortoleva, & Riella, 2015; Tserenjigmid, 2019), or stochastic (Kibris, Masatlioglu, & Suleymanov, 2024; Ravid & Stevenson, 2019). The MSC model introduces a dynamic reference point represented by the current alternative, since the probability of transitioning to another option is influenced by the present state in the Markov chain. Alternatively, the initial fixation can be interpreted as a stochastic reference point. Our contribution to this literature lies in highlighting the nuanced effect that reference points have on final choices, depending on the time pressure and the existence of consideration sets.

Limited consideration in decision-making, i.e. the idea that agents do not consider all available alternatives, has also been a very influential concept in behavioral choice literature. Consideration sets have been modeled as exogenous (Demuynck & Seel, 2018; Masatlioglu, Nakajima, & Ozbay, 2012), random (Brady & Rehbeck, 2016; Manzini & Mariotti, 2014; Masatlioglu et al., 2012), or as a result of constrained optimization (Caplin, Dean, & Leahy, 2018). We contribute to this literature by identifying necessary and sufficient conditions showing the existence of such consideration sets and their composition despite the fact that the agent makes stochastic choices even if all alternatives are considered.

Finally, our paper is related to the literature on manipulating attention by different presentations of the menu and its consequences on choice behavior. Notable recent theoretical works addressing this issue are Gossner, Steiner, and Stewart (2021), which features a learning mechanism and endogenous stopping, and Kovach and Tserenjigmid (2022), which enriches the Luce (1959) model by separating the alternatives into two categories depending on their ability to attract attention. Our framework enables us to study in addition the impact of manipulating the comparability between pairs of alternatives on final choices.

6 Conclusion

Motivated by key behavioral patterns such as pairwise comparisons, attention-driven transitions, and stochastic decision-making, we introduce the Markov Stochastic Choice (MSC) model as a tool to analyse the effects of item arrangement and attention manipulation on final choices. Understanding and predicting these effects is crucial for choice problem designers interested in influencing or minimizing external effects on choices.

Our findings indicate that decision-makers may differ in their sensitivity to presentation formats. We identify decision-makers following reversible or fully comparable Markov processes as being unaffected by item positioning and show how these properties can be revealed from choice data. Additionally, we provide a behavioral foundation for the Luce model as observationally equivalent to a reversible and fully comparable MSC model without time constraints. Finally, interventions such as altering initial fixation probabilities can effectively nudge decisions, especially under time constraints or when the decision process is reducible. We are able to determine the existence and composition of consideration sets from the choice function.

The proposed model offers different directions for future research, two of which we outline here. In the current model, time pressure is reflected in the stopping probability of the decision-making process. An alternative way to model time pressure is as a fixed point in time at which the exploration of the choice set is terminated. One could then compare the effect that the modelling of the two approaches to time pressure have on final choices and use that information to make predictions.

The second direction is to extend the model to multi-attribute alternatives. This can be easily incorporated in the existing framework by letting each state of the Markov chain represent an attribute of an alternative. The multi-attribute version of the MSC model offers a natural environment to study the other context effects known in the literature, such as the compromise and similarity effects, and contribute to a better understanding of the effect of alternative- vs. attribute-based transitions on choices, which have been extensively studied in the experimental literature (see Noguchi and Stewart (2014)).

A Proofs

A.1 Proof of Proposition 1

We begin the proof by simplifying Equation (1).¹⁰

Lemma 1. *The generated stochastic choice function in (1) is equivalent to*

$$(4) \quad \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha \boldsymbol{\pi}(M)(I - (1 - \alpha)Q(M))^{-1}, \quad \forall M \in \mathcal{N}.$$

Proof. Let $\tilde{Q}(M) = (1 - \alpha)Q(M)$. Then, $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha \boldsymbol{\pi}(M) \sum_{\tau=0}^{\infty} \tilde{Q}(M)^{\tau}$. In order to simplify the expression, we need to show that the Neumann series $\sum_{\tau=0}^{\infty} \tilde{Q}(M)^{\tau}$ converges in the operator norm in order to use that $\sum_{\tau=0}^{\infty} \tilde{Q}(M)^{\tau} = (I - \tilde{Q}(M))^{-1}$ (see for example Kress (2014, Theorem 2.14)). This is the case if $\tilde{Q}(M)$ is smaller than unity in some norm. Since $Q(M)$ is a stochastic matrix $\|Q(M)\|_{\infty} = \max_i \sum_j |q_{ij}(M)| = 1$ and $\|\tilde{Q}(M)\|_{\infty} = 1 - \alpha < 1$. Hence, $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha \boldsymbol{\pi}(M)(I - \tilde{Q}(M))^{-1}$. \square

Fix a menu $M \in \mathcal{M}$. We reformulate Equation (4) in the following way:

$$\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \frac{\alpha}{1 - \alpha} \boldsymbol{\pi}(M) \frac{1}{\det\left(\frac{1}{1-\alpha}I - Q(M)\right)} \text{adj}\left(\frac{1}{1-\alpha}I - Q(M)\right),$$

where $\text{adj}(\cdot)$ denotes the adjugate matrix. Since the Markov process is ergodic, $Q(M)$ has an eigenvalue of 1. We express the determinant with the characteristic polynomial:

$$\det\left(\frac{1}{1-\alpha}I - Q(M)\right) = \prod_{l=1}^L \left(\frac{1}{1-\alpha} - \lambda_l\right) = \frac{\alpha}{1-\alpha} \prod_{l=2}^L \left(\frac{1}{1-\alpha} - \lambda_l\right),$$

where λ_l denotes the eigenvalues of $Q(M)$. Since $\lambda_l < 1$ for $l \neq 1$, the determinant is positive. We plug in the expression and simplify the resulting choice function

$$\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \boldsymbol{\pi}(M) \frac{1}{\prod_{l=2}^L \left(\frac{1}{1-\alpha} - \lambda_l\right)} \text{adj}\left(\frac{1}{1-\alpha}I - Q(M)\right).$$

¹⁰An alternative way of defining an MSC model is an absorbing Markov chain with one transient and one absorbing state corresponding to each alternative in the menu and transition probability matrix between transient states given by $Q^*(M) = (1 - \alpha)Q(M)$ and absorption probability matrix $A(M) = \alpha I$. The stationary distribution of this Markov chain is given by $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}^*(M), Q^*(M)) = \boldsymbol{\pi}^*(M)(I - Q^*(M))^{-1}A(M)$ (Grinstead & Snell, 1997). Thus, Lemma 1 is a corollary of existing results from the literature on absorbing Markov chains. Nevertheless, the proof is provided here for convenience.

Now we can let $\alpha \rightarrow 0$:

$$\rho(\pi(M), Q(M)) = \lim_{\alpha \rightarrow 0} \rho(\alpha, \pi(M), Q(M)) = \pi(M) \frac{1}{\prod_{l=2}^L (1 - \lambda_l)} \text{adj}(I - Q(M)).$$

Multiply both sides of the equation with $(I - Q(M))$ from the right and simplify:

$$\rho(\pi(M), Q(M))(I - Q(M)) = \pi(M) \frac{1}{\prod_{l=2}^L (1 - \lambda_l)} \det(I - Q(M)).$$

Since $\det(I - Q(M)) = 0$, we have $\rho(\pi(M), Q(M))(I - Q(M)) = \mathbf{0}$.

A.2 Auxiliary results

This section contains three key lemmas for proving the main theorems.

Lemma 2. *Let $\mathbf{p}(\mathcal{N})$ be a stochastic choice function rationalizable by an MSC model. For all $M \in \mathcal{N}$ and $i, j \in M$ for which $p(i, \{i, j\}) \in (0, 1)$ it holds that*

$$\frac{\delta_{ij}(\mathbf{p}(M))q_{ij}(M)}{p(j, \{i, j\})} = -\frac{\delta_{ji}(\mathbf{p}(M))q_{ji}(M)}{p(i, \{i, j\})}.$$

Proof. Take an arbitrary pair $i, j \in M$ with $p(i, \{i, j\}) \in (0, 1)$. Rationalizability of $\mathbf{p}(\mathcal{N})$ implies $p(i, \{i, j\})q_{ij}(\{i, j\}) = p(j, \{i, j\})q_{ji}(\{i, j\})$. It follows from $p(i, \{i, j\}) \in (0, 1)$ and Assumption 2 that $q_{ij}(\{i, j\}) > 0$ and $q_{ji}(\{i, j\}) > 0$. The statement is trivially satisfied for $q_{ij}(M) = q_{ji}(M) = 0$. If $q_{ij}(M) > 0$, TR-IIA implies $q_{ji}(M) > 0$ and $\frac{q_{ij}(M)}{p(j, \{i, j\})} = \frac{q_{ji}(M)}{p(i, \{i, j\})}$. Therefore,

$$\frac{\delta_{ij}(\mathbf{p}(M))q_{ij}(M)}{p(j, \{i, j\})} = (p(i, M)p(j, \{i, j\}) - p(i, \{i, j\})p(j, M)) \frac{q_{ji}(M)}{p(i, \{i, j\})},$$

and the final result follows. \square

Lemma 3. *A stochastic choice function $\mathbf{p}(\mathcal{N})$ is rationalizable with an MSC model iff*

$$(5) \quad \sum_{\{i \in M : q_{ji}(M) > 0\}} \frac{\delta_{ji}(\mathbf{p}(M))q_{ji}(M)}{p(i, \{i, j\})} - \sum_{\substack{\{i \in M : q_{ji}(M) = 0, \\ q_{ji}(\{i, j\}) = 0\}}} \frac{\delta_{ij}(\mathbf{p}(M))q_{ij}(M)}{p(j, \{i, j\})} = 0 \quad \forall j \in M, \forall M \in \mathcal{N}.$$

Proof. The rationalizability of $\mathbf{p}(\mathcal{N})$ implies that $\mathbf{p}(M)(I - Q(M)) = 0$ for all $M \in \mathcal{N}$. Therefore, it holds for all $j \in M$ and $M \in \mathcal{N}$ that

$$\sum_{i \neq j} p(j, M)q_{ji}(M) - \sum_{i \neq j} p(i, M)q_{ij}(M) = 0.$$

We now rearrange the terms depending on whether $q_{ji}(M) > 0$ or not.

$$\sum_{\{i \in M \setminus \{j\} : q_{ji}(M) > 0\}} \left(p(j, M) - p(i, M) \frac{q_{ij}(M)}{q_{ji}(M)} \right) q_{ji}(M) - \sum_{\substack{\{i \in M \setminus \{j\}: \\ q_{ji}(M) = 0\}}} p(i, M) q_{ij}(M) = 0$$

It follows from Assumptions 2 and 3 that if $q_{ji}(M) \neq 0$, then $q_{ji}(\{i, j\}) \neq 0$. Rationalizability and Assumption 2 imply that $p(i, \{i, j\}) \neq 0$. We apply Assumption 3 and the rationalizability of the choice function from binary sets to obtain:

$$\sum_{\substack{\{i \in M \setminus \{j\}: \\ q_{ji}(M) > 0\}}} \left(p(j, M) - p(i, M) \frac{p(j, \{i, j\})}{p(i, \{i, j\})} \right) q_{ji}(M) - \sum_{\substack{\{i \in M \setminus \{j\}: \\ q_{ji}(M) = 0\}}} p(i, M) q_{ij}(M) = 0,$$

Note that TR-IIA and $q_{ji}(M) = 0$ imply that either $q_{ij}(M) = 0$ or $q_{ji}(\{i, j\}) = 0$. In the latter case rationalizability guarantees that $p(i, \{i, j\}) = 0$. Hence, for those pairs holds $\delta_{ij}(\mathbf{p}(M)) = p(i, M)p(j, \{i, j\}) - p(i, \{i, j\})p(j, M) = p(i, M)$ and the final result follows. \square

Let $G^M = \{(i, j) \in M^2 : \delta_{ij}(\mathbf{p}(M)) > 0\} = \{g_b^M\}_{b \in \{1, \dots, |G^M|\}}$ be an indexed set of ordered pairs. The menu is also denoted as an indexed set $M = \{m_a : a \in \{1, \dots, |M|\}\}$.

Lemma 4. *A stochastic choice function $\mathbf{p}(\mathcal{N})$ is rationalizable with an MSC model iff*

$$(6) \quad \mathcal{D}(M)\boldsymbol{\gamma}(M) = \mathbf{0}, \quad \forall M \in \mathcal{N}.$$

where $\boldsymbol{\gamma}(M)$ is a vector of transition probabilities between the pairs in G^M , i.e. $\gamma_b(M) = q_{ij}(M)$ iff $g_b^M = (i, j)$ and $\mathcal{D}(M)$ is a matrix with dimensions $|M| \times |G^M|$ and elements

$$d_{a,b}(M) = \begin{cases} \delta_{ij}(\mathbf{p}(M))/p(j, \{i, j\}), & \text{if } m_a = i, g_b^M = (i, j) \\ -\delta_{ij}(\mathbf{p}(M))/p(j, \{i, j\}), & \text{if } m_a = j, g_b^M = (i, j), \\ 0 & \text{else.} \end{cases}$$

Proof. Take an arbitrary menu $M \in N$. Observe that the linear system (5) in Lemma 3 is trivially satisfied whenever $\mathbf{p}(M)$ is such that $\delta_{ij}(\mathbf{p}(M)) = 0$ for all pairs $i, j \in M$. If it holds for at least one pair $i, j \in M$ that $\delta_{ij}(\mathbf{p}(M)) > 0$, then $G^M \neq \emptyset$, which we assume for the remainder of the proof.

Note that all elements of the matrix $\mathcal{D}(M)$ are well-defined since $\delta_{ij}(\mathbf{p}(M)) > 0$ implies $p(j, \{i, j\}) > 0$. Further, none of the columns of the matrix is equal to $\mathbf{0}$ because each column corresponds to a pair (i, j) for which $\delta_{ij}(\mathbf{p}(M)) > 0$.

We claim that the linear system (6) is equivalent to the one in (5) and is thus necessary and sufficient for rationalizability. Consider an equation from the linear system (6) corresponding to an arbitrary matrix row a such that $m_a = j$:

$$\begin{aligned} \sum_{\{b: g_b^M = (j, i), \forall i \in M\}} d_{a,b}(M) \gamma_b(M) + \sum_{\{b: g_b^M = (i, j), \forall i \in M\}} d_{a,b}(M) \gamma_b(M) &= 0 \\ \sum_{\{i \in M: (j, i) \in G^M\}} \frac{\delta_{ji}(\mathbf{p}(M)) q_{ji}(M)}{p(i, \{i, j\})} - \sum_{\{i \in M: (i, j) \in G^M\}} \frac{\delta_{ij}(\mathbf{p}(M)) q_{ij}(M)}{p(j, \{i, j\})} &= 0 \\ \sum_{\{i \in M \setminus \{j\}: q_{ji}(M) > 0\}} \frac{\delta_{ji}(\mathbf{p}(M)) q_{ji}(M)}{p(i, \{i, j\})} - \sum_{\{i \in M \setminus \{j\}: q_{ji}(M) = 0, q_{ji}(\{i, j\}) = 0\}} \frac{\delta_{ij}(\mathbf{p}(M)) q_{ij}(M)}{p(j, \{i, j\})} &= 0, \end{aligned}$$

where we obtain the last equation by applying Lemma 2 on all pairs $(i, j) \in G^M$ for which $q_{ji}(M) \neq 0$ and use the fact that Assumption 3 and $q_{ji}(M) = 0$ imply that either $q_{ij}(M) = 0$ or $q_{ji}(\{i, j\}) = 0$. Note that applying Lemma 2 does not lead to a division by zero because if $q_{ji}(M) > 0$, then $q_{ji}(\{i, j\}) > 0$ and by rationalizability, $p(i, \{i, j\}) > 0$. \square

A.3 Proof of Theorem 1

A.3.1 Equivalence between (i) and (ii)

The necessity part of the argument follows directly from Proposition 2. In order to show the sufficiency part, fix a menu M and let $\mathbf{p}(M)$ be rationalizable by an MSC model $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ and all its comparability restrictions. We will show that $Q(M)$ satisfies the Kolmogorov criterion given in Equality (3). We consider the comparability restriction $R(M)$ with positive entries except between the pair $i, j \in M$ such that $r_{ij}(M) = r_{ji}(M) = 0$. The rationalizability of $\mathbf{p}(M)$ by $Q(M)$ and $R(M) \odot Q(M)$ implies that $\rho(\boldsymbol{\pi}(M), Q(M))(I - Q(M)) = \mathbf{0}$ and $\rho(\boldsymbol{\pi}(M), Q(M))(I - R(M) \odot Q(M)) = \mathbf{0}$. Consider the row of these linear systems corresponding to alternative j

$$\begin{aligned} \sum_{k \neq j} \rho(k, Q(M)) q_{kj}(M) - \rho(j, Q(M)) \sum_{k \neq j} q_{jk}(M) &= 0, \\ \sum_{k \neq j} \rho(k, Q(M)) r_{kj}(M) q_{kj}(M) - \rho(j, Q(M)) \sum_{k \neq j} r_{jk}(M) q_{jk}(M) &= 0. \end{aligned}$$

Using that $r_{kl}(M) = c$ except for the pair i, j we can express latter equation:

$$c \sum_{k \neq j, i} \rho(k, Q(M)) q_{kj}(M) - \rho(j, Q(M)) q_{jk}(M) = 0.$$

Thus, detailed balance given in Equation 2 is satisfied for the pair i, j in the non-restricted model $Q(M)$. We can prove analogously that detailed balance holds on all pairs such that restricting their comparability does not lead to the formation of new communicating classes. Note that detailed balance holds trivially for all pairs which are not directly comparable in $Q(M)$. Thus, all cycles $C(M')$ composed of these pairs satisfy the Kolmogorov criterion.

Finally, we consider those pairs that belong to the same communicating class according to $Q(M)$, but belong to different classes when the transition between the two is prohibited. Let M' and M'' be two disjoint subsets of M with $i \in M'$ and $j \in M''$ such that $q_{kl}(M) = 0$ for all $k \in M'$ and $l \in M''$ except $q_{ij}(M) > 0$. It easily seen that the Kolmogorov criterion is satisfied for all cycles with $(i, j) \in C(M' \cup M'')$. Therefore, the Kolmogorov criterion holds on all cycles of alternatives and the Markov process given by $Q(M)$ is reversible.

A.3.2 Equivalence between (i) and (iii)

Consider a stochastic choice function $\mathbf{p}(\mathcal{N})$ and a menu $M \in \mathcal{N}$. The condition $\mathcal{C}^S(\mathbf{p}(M)) = \mathcal{C}_0^S(\mathbf{p}(M))$ is trivially satisfied for $|M| = 2$, as is reversibility by all generating models. We assume $|M| \geq 3$ for the rest of the proof. Let $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ be a rationalizing MSC model, i.e. $\rho(\boldsymbol{\pi}(M), Q(M)) = \mathbf{p}(M)$. Note that at least one rationalizing model always exists: letting $q_{ij}(M) = 0$ for all $i, j \in M$ and $\boldsymbol{\pi}(M) = \mathbf{p}(M)$ generates the stochastic choice function $\mathbf{p}(M)$.

(iii) \implies (i): We prove the statement by contraposition, i.e. we show that if an MSC model violates reversibility on M , then $\mathbf{p}(M)$ is such that $\mathcal{C}_+^S(\mathbf{p}(M)) \neq \emptyset$.

The proof is structured in three steps. First, we show that a non-reversible MSC model on M generates stochastic choice functions with $\delta_{kl}(\mathbf{p}(M)) \neq 0$ for those pairs $k, l \in M$ that violate the detailed balance conditions given in (2). In the second step, we apply Lemma 4 from Appendix A.2 which gives necessary and sufficient conditions for rationalizability. If the Markov chain is not reversible on M , the system has a positive solution. Finally, we invoke Gordan's theorem, which gives a necessary and sufficient condition for the existence of positive solutions to a homogeneous linear system and show the relationship of the condition to the property $\mathcal{C}^S(\mathbf{p}(M)) = \mathcal{C}_0^S(\mathbf{p}(M))$.

Step 1: If an MSC model is not reversible on M , then detailed balance is violated for

at least one pair of alternatives: $\exists k, l \in M$ such that

$$(7) \quad q_{kl}(M)p(k, M) > q_{lk}(M)p(l, M).$$

Thus, $q_{kl}(M) > 0$ and $p(k, M) > 0$. We consider the following two cases: (1) $q_{lk}(\{k, l\}) = 0$ and (2) $q_{lk}(\{k, l\}) > 0$. In the former, Assumption 2 implies that $q_{kl}(\{k, l\}) > 0$ and thus, $p(k, \{k, l\}) = 0$. Since $p(k, M) > 0$, we have that $p(l, \{k, l\})p(k, M) > p(k, \{k, l\})p(l, M) = 0$, and $\delta_{kl}(\mathbf{p}(M)) > 0$. In the second case with $q_{lk}(\{k, l\}) > 0$, Assumption 3 together with $q_{kl}(M) > 0$ imply that $q_{kl}(\{k, l\}) > 0$ and $q_{lk}(M) > 0$. The generated stochastic choice function is such that $p(l, \{k, l\}) \in (0, 1)$. It follows from Inequality (7) and Assumption 3 that

$$\frac{p(l, \{k, l\})}{p(k, \{k, l\})} = \frac{q_{kl}(\{k, l\})}{q_{lk}(\{k, l\})} = \frac{q_{kl}(M)}{q_{lk}(M)} > \frac{p(l, M)}{p(k, M)}.$$

Thus, we have shown that whenever the detailed balance condition is violated for a pair of alternatives k, l on M , $\delta_{kl}(\mathbf{p}(M)) \neq 0$.

Step 2: Recall from Appendix A.2 that $G^M = \{(i, j) \in M^2 : \delta_{ij}(\mathbf{p}(M)) > 0\} = \{g_b^M\}_{b \in \{1, \dots, |G^M|\}}$. Our previous result ensures that $G^M \neq \emptyset$. We apply Lemma 4, which states that $\mathbf{p}(M)$ is such that

$$(8) \quad \mathcal{D}(M)\boldsymbol{\gamma}(M) = \mathbf{0}.$$

Step 3: As already shown, whenever the detailed balance condition is violated for a pair of alternatives k, l on M (and thus $q_{kl}(M) > 0$ and/or $q_{lk}(M) > 0$), $\delta_{kl}(\mathbf{p}(M)) \neq 0$. This means that $(k, l) \in G^M$ and $\boldsymbol{\gamma}(M) \geq \mathbf{0}$. It follows directly from Gordan's theorem¹¹ that there exists a strictly positive solution $\boldsymbol{\gamma}(M) \geq \mathbf{0}$ to the linear system (8) if and only if there does not exist a vector $\mathbf{z} = (z_1, \dots, z_{|M|}) \in \mathbb{R}^{|M|}$ such that

$$(9) \quad \mathbf{z}\mathcal{D}(M) \gg \mathbf{0}.$$

Let the menu be denoted as an indexed set $M = \{m_a : a \in \{1, \dots, |M|\}\}$ and the elements of \mathbf{z} be such that $z_a = v_k$ whenever $m_a = k$. The linear system (9) is equivalent to

$$z_a d_{a,b}(M) + z_c d_{c,b}(M) > 0, \quad \forall (k, l) \in G^M \text{ with } m_a = k, m_c = l, g_b^M = (k, l),$$

¹¹See for example Theorem 15.1(2) in Roman (2008).

which is in turn equal to

$$(10) \quad (v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} > 0, \quad \forall (k, l) \in G^M.$$

Hence, in order for Inequality (9) to hold, the vector \mathbf{z} should be such that for all pairs $(k, l) \in G^M$, we have $v_k > v_l$. Since the Markov chain is non-reversible on M and $\boldsymbol{\gamma}(M) \geq \mathbf{0}$, there does not exist a vector \mathbf{z} satisfying (10). This is the case when there exists some cycle $C(M') \in \mathcal{C}_+^S(\mathbf{p}(M))$ and $C(M') \subseteq G^M$, because then we cannot assign a value to each element of M' such that $v_k > v_l$ for all $(k, l) \in C(M')$.

(i) \implies (iii): We again prove the statement by contraposition, i.e. we show that if $\mathbf{p}(M)$ is such that $\mathcal{C}^S(\mathbf{p}(M)) \neq \mathcal{C}_0^S(\mathbf{p}(M))$, then there exists a non-reversible rationalizing MSC model. The proof is structured analogously to the sufficiency part.

Step 1: The assumption that $\mathcal{C}^S(\mathbf{p}(M)) \neq \mathcal{C}_0^S(\mathbf{p}(M))$ implies that there exists a positive sign-consistent cycle $C(M') \in \mathcal{C}_+^S(\mathbf{p}(M))$ and $C(M') \subseteq G^M$. Since $G^M \neq \emptyset$, it follows from Lemma 4 that the transition probabilities of all rationalizing models satisfy $\mathcal{D}(M)\boldsymbol{\gamma}(M) = \mathbf{0}$.

Step 2: Gordan's theorem implies that there exists a strictly positive solution $\boldsymbol{\gamma}(M) \geq \mathbf{0}$ if and only if there is no vector $\mathbf{z} \in \mathbb{R}^{|M|}$ for which

$$(11) \quad \mathbf{z}\mathcal{D}(M) \gg \mathbf{0}.$$

Letting $z_a = v_k$ whenever $m_a = k$, the system becomes equivalent to

$$(12) \quad (v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} > 0, \quad \forall (k, l) \in G^M.$$

Since it holds for all $(k, l) \in C(M')$ that $\delta_{kl}(\mathbf{p}(M)) > 0$, we cannot assign a numerical value to each element of M' such that $v_k > v_l$ for all $(k, l) \in C(M')$. Therefore, there does not exist a vector $\mathbf{z} \in \mathbb{R}^{|M|}$ that satisfies the system of inequalities (11) and there exists a rationalizing model is such that $\boldsymbol{\gamma}(M) \geq \mathbf{0}$.

Step 3: We can now use that there is a solution $\boldsymbol{\gamma}(M) \geq \mathbf{0}$ to construct a transition probability matrix $Q(M)$. In particular, we define

$$(13) \quad q'_{ij}(M) = \begin{cases} \gamma_b(M) & \text{if } g_b^M = (i, j) \in G^M, \\ \gamma_b(M) \frac{p(j, \{i, j\})}{p(i, \{i, j\})} & \text{if } g_b^M = (j, i) \in G^M \\ p(j, \{i, j\}) & \text{if } \delta_{ij}(\mathbf{p}(M)) = 0. \end{cases}$$

Note that $(j, i) \in G^M$ implies that $p(i, \{i, j\}) > 0$ since $\delta_{ji}(\mathbf{p}(M)) > 0$. Next, we let $q_{ij}(M) = \kappa q'_{ij}(M)$ for all $i \neq j$, where $\kappa > 0$ is such that $\kappa \sum_{j \neq i} q'_{ij}(M) < 1$ for all $i \in M$. Finally, we let $q_{ii}(M) = 1 - \sum_{j \neq i} q_{ij}(M)$. The constructed transition probability matrix satisfies Assumptions 1 – 3 and rationalizes $\mathbf{p}(M)$.

Finally, we will show that there exists a pair of alternatives that violates detailed balance. From the above definition of the model and the fact that $\boldsymbol{\gamma}(M) \geq \mathbf{0}$, there must exist a pair $(k, l) \in G^M$, i.e. $\delta_{kl}(\mathbf{p}(M)) = p(k, M)p(l, \{k, l\}) - p(l, M)p(k, \{k, l\}) > 0$, for which $q_{kl}(M) > 0$. Therefore, $p(l, \{k, l\}) > 0$ and $p(k, M) > 0$. We consider two cases: (1) $p(k, \{k, l\}) = 0$ and (2) $p(k, \{k, l\}) > 0$. In the first case, it follows from rationalizability and Assumptions 1 – 3 that $q_{lk}(M) = 0$. Since $q_{kl}(M)p(k, M) > 0$ and $q_{lk}(M) = 0$, detailed balance is violated on the pair k, l and the MSC model is non-reversible. In the second case with $p(k, \{k, l\}) > 0$, we have $q_{lk}(M) = q_{kl}(M) \frac{p(k, \{k, l\})}{p(l, \{k, l\})}$. Thus, $\delta_{kl}(\mathbf{p}(M)) > 0$ implies $\frac{p(l, \{k, l\})}{p(k, \{k, l\})} = \frac{q_{kl}(M)}{q_{lk}(M)} > \frac{p(l, M)}{p(k, M)}$, which is a violation of detailed balance.

A.4 Proof of Proposition 2

Consider a reversible and irreducible MSC model $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ and its generated stochastic choice function $\boldsymbol{\rho}(\boldsymbol{\pi}(M), Q(M))$ for a menu $M \in \mathcal{M}$. Let $R(M)$ be an arbitrary comparability restriction of $\langle Q(M), \boldsymbol{\pi}(M) \rangle$. We will show that $\langle R(M) \odot Q(M), \boldsymbol{\pi}(M) \rangle$ is reversible and irreducible and $\boldsymbol{\rho}(\boldsymbol{\pi}(M), Q(M))$ is its stationary distribution.

Since a comparability restriction preserves the communicating classes, $R(M) \odot Q(M)$ is irreducible and has a unique stationary distribution. The reversibility of $Q(M)$ implies that the Kolmogorov's criterion given in Equation (3) is satisfied. The symmetry of $R(M)$ implies that the Kolmogorov condition is also satisfied by $R(M) \odot Q(M)$ and is reversible.

To prove rationalizability, we need to show $\boldsymbol{\rho}(\boldsymbol{\pi}(M), Q(M))(I - R(M) \odot Q(M)) = \mathbf{0}$.

We can express the equation for all $j \in M$ as

$$\begin{aligned} \sum_{i \neq j} \rho(i, Q(M)) r_{ij}(M) q_{ij}(M) - \rho(j, Q(M)) \sum_{i \neq j} r_{ji}(M) q_{ji}(M) &= 0, \\ \sum_{i \neq j} r_{ij}(M) (\rho(i, Q(M)) q_{ij}(M) - \rho(j, Q(M)) q_{ji}(M)) &= 0, \end{aligned}$$

where the last step results from the symmetry of $R(M)$. Since $Q(M)$ must satisfy the

detailed balance condition given in Equation (2) on all pairs, one can easily see that the equality holds.

Finally, consider the case when the initial MSC model is not irreducible. Since comparability restrictions do not affect the initial distribution and preserve the communicating classes, our proof extends trivially to the case with reducible MSC models.

A.5 Proof of Theorem 2

Consider a stochastic choice function $\mathbf{p}(\mathcal{N})$ and a menu $M \in \mathcal{N}$. Since both parts of the theorem are trivially satisfied for $|M| = 2$, we assume $|M| \geq 3$ for the rest of the proof.

(i) \implies (ii): We consider a pairwise comparable MSC model on M with generated stochastic choice function $\rho(\pi(M), Q(M)) = \mathbf{p}(M)$. If it holds for all pairs $i, j \in M$ that $\delta_{ij}(\mathbf{p}(M)) = 0$, the condition in (ii) is trivially satisfied. We assume hereafter that there is at least one pair $i, j \in M$ with $\delta_{ij}(\mathbf{p}(M)) > 0$.

We apply Lemma 4 from Appendix A.2. Since $\delta_{ij}(\mathbf{p}(M)) > 0$, $(i, j) \in G^M$ by definition. The assumption that the model is pairwise comparable, i.e. there is no $k, l \in M$ for which $q_{kl}(M) = q_{lk}(M) = 0$, ensures that $\gamma(M) \gg \mathbf{0}$. We use Stiemke's lemma,¹² which states that there exists a strongly positive solution $\gamma(M)$ to the linear system in (6) if and only if there does not exist a vector $\mathbf{z} \in \mathbb{R}^{|M|}$ for which

$$(14) \quad \mathbf{z}\mathcal{D}(M) \geq \mathbf{0}.$$

Analogously to the simplification presented in the proof of Theorem 1 given in Inequality (10), this system can be written as

$$(v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} \geq 0, \forall (k, l) \in G^M \text{ and } (v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} > 0 \text{ for at least one } (k, l) \in G^M.$$

Since $\gamma(M) \gg \mathbf{0}$, such vector \mathbf{z} does not exist. Note that it must hold for $k, l \in M$ that belong to a cycle in $\mathcal{C}_+^S(\mathbf{p}(M))$ that $v_k = v_l$. Since a vector \mathbf{z} does not exist, we should not be able to find a pair $(k, l) \in G^M$ such that $v_k > v_l$. This is the case when $\mathbf{p}(M)$ is such that all pairs $(k, l) \in G^M$ belong to some sign-consistent cycle in $\mathcal{C}_+^S(\mathbf{p}(M))$, which coincides with condition (ii).

¹²See for example Theorem 15.1(1) in Roman (2008).

Finally, if the MSC model is fully comparable, the positivity assumption of the generated stochastic choice function follows directly from the Perron–Frobenius theorem.

(ii) \implies (i): Consider a stochastic choice function such that for all pairs $i, j \in M$ with $\delta_{ij}(\mathbf{p}(M)) \neq 0$, $\exists C(M')$ with $(i, j) \in C(M')$ and $C(M') \in \mathcal{C}^S(\mathbf{p}(M))$ for $M' \in \mathcal{M}$. We will show that there exists a rationalizing MSC model with $Q(M)$ that is pairwise comparable. Note that all stochastic choice functions are rationalizable by an MSC model (see the Appendix A.3.2).

Assume first that $\delta_{ij}(\mathbf{p}(M)) = 0$ for all pairs $i, j \in M$. We apply Lemma 3 from Appendix A.2 and see that Equation (5) is trivially satisfied for all $j \in M$. A pairwise comparable MSC model that rationalizes the stochastic choice function is, for example, one where $q_{ij}(M) = \kappa p(j, \{i, j\})$, where $\kappa > 0$ is such that $Q(M)$ satisfies Assumptions 1 – 3.

We assume hereafter that there is at least one pair $i, j \in M$ with $\delta_{ij}(\mathbf{p}(M)) > 0$ and thus $(i, j) \in G^M$. It follows from Lemma 4 that the transition probabilities satisfy the system

$$(15) \quad \mathcal{D}(M)\boldsymbol{\gamma}(M) = \mathbf{0}.$$

Following Stiemke's lemma, there exists a strongly positive solution $\boldsymbol{\gamma}(M) \gg \mathbf{0}$ to the system (15) if and only if there does not exist a vector $\mathbf{z} \in \mathbb{R}^{|M|}$ for which

$$(16) \quad \mathbf{z}\mathcal{D}(M) \geq \mathbf{0}.$$

Analogously to the proof of Theorem 1, the above linear system is equivalent to

$$(v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} \geq 0, \forall (k, l) \in G^M \text{ and } (v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} > 0 \text{ for at least one } (k, l) \in G^M.$$

Since it holds for all alternatives $(k, l) \in G^M$ that $\exists C(M')$ with $(k, l) \in C(M')$ and $C(M') \in \mathcal{C}^S(\mathbf{p}(M))$ for some $M' \in \mathcal{M}$, we have that $v_k = v_l$ and there is no pair $(k, l) \in G^M$ for which $(v_k - v_l) \frac{\delta_{kl}(\mathbf{p}(M))}{p(l, \{k, l\})} > 0$. Thus, there exists a rationalizing model with $\boldsymbol{\gamma}(M) \gg \mathbf{0}$.

We can use the solution to the system (15) to construct a transition probability matrix $Q(M)$ as follows:

$$(17) \quad q'_{ij}(M) = \begin{cases} \gamma_b(M) & \text{if } g_b^M = (i, j) \in G^M, \\ \gamma_b(M) \frac{p(j, \{i, j\})}{p(i, \{i, j\})} & \text{if } g_b^M = (j, i) \in G^M \\ p(j, \{i, j\}) & \text{if } \delta_{ij}(\mathbf{p}(M)) = 0. \end{cases}$$

Note that $(j, i) \in G^M$ implies that $p(i, \{i, j\}) > 0$ since $\delta_{ji}(\mathbf{p}(M)) > 0$. Next, we let $q_{ij}(M) = \kappa q'_{ij}(M)$ for all $i \neq j$, where $\kappa > 0$ is such that $\kappa \sum_{j \neq i} q'_{ij}(M) < 1$ for all $i \in M$. Finally, we let $q_{ii}(M) = 1 - \sum_{j \neq i} q_{ij}(M)$. This model satisfies Assumptions 1 – 3, rationalizes $\mathbf{p}(M)$, and is pairwise comparable.

Finally, we show that assuming positivity in addition, ensures the rationalizability with a fully comparable model. Since $\boldsymbol{\gamma}(M) \gg \mathbf{0}$, it holds for all $(i, j) \in G^M$ that $q_{ij}(M) > 0$. It is easy to see that the positivity of the stochastic choice function ensures that all other transition probabilities are positive as well, hence the constructed model is fully comparable.

A.6 Proof of Proposition 3

Since $\mathbf{p}(M)$ is not bounded in a cycle over the pair i, j , it holds $\delta_{ji}(\mathbf{p}(M)) \neq 0$ and $\#C(M') \geq 2$ with $(i, j) \in C(M')$ and $C(M') \in \mathcal{C}_{+/-}^S(\mathbf{p}(M))$ for $M' \in \mathcal{M}$. We assume by contradiction that there is a rationalizing model $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ with $q_{ij}(M) > 0$. Since $\mathbf{p}(M)$ violates Theorem 2(ii), there has to be at least one pair $k, l \in M$ for which $q_{kl}(M) = 0$. Let us denote the set of such pairs with $S(Q(M))$.

We construct a stochastic choice function $\mathbf{p}(\mathcal{N})^*$ such that $\mathbf{p}(M)^* = \mathbf{p}(M)$ and for all $k, l \in S(Q(M))$ we adjust the choice probability from the binary sets so that $\delta_{kl}(\mathbf{p}(M)^*) = 0$. Note that $\mathbf{p}(M)^* \neq \mathbf{p}(M)$ since following the proof of Theorem 2, there is at least one pair with $\delta_{kl}(\mathbf{p}(M)) \neq 0$ and $q_{kl}(M) = 0$. Note further that $\mathbf{p}(M)^*$ is not bounded in a cycle over the pair i, j , since according to the described procedure, $\mathcal{C}_{+/-}^S(\mathbf{p}(M)^*) \subseteq \mathcal{C}_{+/-}^S(\mathbf{p}(M))$ and $\delta_{ij}(\mathbf{p}(M)^*) \neq 0$.

We now construct an MSC model with $Q(M)^*$ such that $q_{kl}(M)^* = q_{kl}(M)$ for all $k, l \notin S(Q(M))$ and let $q_{kl}(M)^* \in (0, 1)$ and $q_{lk}(M)^* = q_{kl}(M)^* \frac{p(k, \{k, l\})^*}{p(l, \{k, l\})^*}$ for all $k, l \in S(Q(M))$. Observe that this model is pairwise comparable.

Applying Lemma 4 shows that the constructed MSC model $Q(M)^*$ rationalizes $\mathbf{p}(M)^*$. This is the case because $\delta_{kl}(\mathbf{p}(M)^*) q_{kl}(M)^* = \delta_{kl}(\mathbf{p}(M)) q_{kl}(M)$ for all $k, l \in M$. Theorem 2 implies that $\mathbf{p}(M)^*$ should be bounded in a cycle over all pairs of alternatives, which is a contradiction. Therefore, all rationalizing models of $\mathbf{p}(\mathcal{N})$ should be such that $q_{ij}(M) = 0$.

A.7 Proof of Theorem 3

Consider a stochastic choice function $\mathbf{p}(\mathcal{N})$ and a rationalizing MSC model on $M \in \mathcal{N}$, i.e. $\rho(\pi(M), Q(M)) = \mathbf{p}(M)$.

A.7.1 Equivalence between (i) and (ii)

(i) \implies (ii): Let $Q(M)$ be a reversible and fully comparable MSC model. We fix an arbitrary pair $i, j \in M$, for which we will show that the stochastic choice function satisfies positivity and IIA.

Since the model is fully comparable, the stationary distributions are such that $\rho(i, Q(M)) > 0$ for all $i \in M$ and all $M \in \mathcal{M}$, as there are no transient states. Hence, the generated stochastic choice function is positive. Reversibility means that the detailed balance condition is satisfied for all pairs and menus, hence

$$\frac{\rho(i, Q(M))}{\rho(j, Q(M))} = \frac{q_{ji}(M)}{q_{ij}(M)} \text{ and } \frac{\rho(i, Q(\{i, j\}))}{\rho(j, Q(\{i, j\}))} = \frac{q_{ji}(\{i, j\})}{q_{ij}(\{i, j\})}.$$

It follows from Assumption 3 that all fractions in the above equations are equal and hence IIA is satisfied on the pair i, j and thus, on all pairs and menus.

(ii) \implies (i): Let $\mathbf{p}(\mathcal{N})$ be a positive stochastic choice function that satisfies IIA, thus $\delta_{kl}(\mathbf{p}(M)) = 0$ for all $k, l \in M$. Lemma 3 in Appendix A.2 implies that there exists a fully comparable rationalizing MSC model. We will show that the detailed balance condition given in Equation (2) holds on an arbitrary pair $i, j \in M$.

Using the positivity and IIA properties of the choice function, we obtain

$$\frac{p(j, M)}{p(i, M)} = \frac{p(j, \{i, j\})}{p(i, \{i, j\})} = \frac{q_{ji}(\{i, j\})}{q_{ij}(\{i, j\})}.$$

Finally, Assumption 3 and the above equation imply detailed balance. Thus, the condition is satisfied on all pairs and menus and the rationalizing model is reversible.

A.7.2 Equivalence between (i) and (iii)

(iii) \implies (i): If there exists a utility function $u : N \rightarrow \mathbb{R}_{++}$ such that the ratios of transition probabilities satisfy $\frac{q_{ij}(M)}{q_{ji}(M)} = \frac{u(j)}{u(i)}$ for all $i, j \in M$, reversibility of the MSC model is trivial as it follows directly from Kolmogorov's criterion given in Equation (3).

(i) \implies (iii): Let $Q(M)$ be a reversible and fully comparable MSC model. Hence, the detailed balance condition (2) is satisfied for all pairs of alternatives. As we showed previously, the generated stochastic choice function is positive and IIA. It follows from Luce (1959) that there exists an increasing function $u : N \rightarrow \mathbb{R}_{++}$ such that for all $M \in \mathcal{M}$ and $i \in M$ such that

$$(18) \quad p(i, M) = \frac{u(i)}{\sum_{j \in M} u(j)}$$

Plugging in Equation (18) into the detailed balance condition, we obtain that for all $i, j \in M$ and all $M \in \mathcal{M}$

$$q_{ji}(M) \frac{u(j)}{\sum_{k \in M} u(k)} = \frac{u(i)}{\sum_{k \in M} u(k)} q_{ij}(M)$$

and the result follows.

A.8 Proof of Theorem 4

The MSC model is finite and aperiodic by definition. If it is irreducible in addition, the Markov chain is ergodic. Thus, the equivalence of (i) and (ii) follows trivially from Proposition 1, since the stationary distribution of an ergodic Markov process does not depend on the initial distribution.

A.8.1 Proof of equivalence between statements (i) and (iii)

Consider a stochastic choice function $\mathbf{p}(\mathcal{N})$ and a rationalizing MSC model $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ on $M \in \mathcal{N}$. Note that a rationalizing model always exists (see Appendix A.3.2).

(i) \implies (iii): Let $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ be an irreducible model, i.e. every state can be reached from every other state. In other words, for a given pair $(i, j) \in M$, there exists some $C(M')$ for which $q_{kl}(M) > 0$ for all $(k, l) \in C(M') \setminus (j, i)$. Note that if we take the union of all such cycles corresponding to each of the pairs, we obtain a cycle $C(M)$ for which $q_{kl}(M) > 0$ for all $(k, l) \in C(M)$. We need to show that the generated $\mathbf{p}(M)$ is bounded in a cycle over all $(k, l) \in C(M)$, i.e. if $\delta_{kl}(\mathbf{p}(M)) \neq 0$, then $\exists C(M'')$ with $(k, l) \in C(M'')$ and $C(M'') \in \mathcal{C}^S(\mathbf{p}(M))$ for some $M'' \in \mathcal{M}$.

Suppose by contradiction that there is a pair $(k, l) \in C(M)$, such that $\delta_{kl}(\mathbf{p}(M)) \neq 0$, but there is no sign-consistent cycle $C(M'')$ with $(k, l) \in C(M'')$ and $C(M'') \in \mathcal{C}^S(\mathbf{p}(M))$.

Proposition 3 implies that all rationalizable models of $\mathbf{p}(M)$ are such that $q_{ij}(M) = q_{ji}(M) = 0$, which is a contradiction. Thus, $\mathbf{p}(M)$ needs to be bounded in a cycle over the pairs of the cycle $C(M)$.

(iii) \implies (i): Let $\mathbf{p}(M)$ be bounded in a cycle over all pairs of alternative of some cycle $C(M)$. We will show that there is a model $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ such that $q_{kl}(M) > 0$ for all $(k, l) \in C(M)$, as it implies that every state can be reached from every other state and the Markov chain is irreducible.

We construct a stochastic choice function $\mathbf{p}(\mathcal{N})^*$ such that $\mathbf{p}(\mathcal{N})^* = \mathbf{p}(\mathcal{N})$ except for the pairs $k, l \in M$ for which $\mathbf{p}(M)$ is not bounded in a cycle, we adjust the choice probability from the binary sets so that $\delta_{kl}(\mathbf{p}(M)^*) = 0$. Thus, $\mathbf{p}(M)^*$ is bounded in a cycle over all pairs $k, l \in M$. Following Theorem 2, $\mathbf{p}(M)^*$ is rationalizable with a pairwise comparable MSC model $\langle Q^*(M), \boldsymbol{\pi}(M) \rangle$.

We let $Q(M)$ be such that $q_{ij}(M) = q_{ij}(M)^*$ for all pairs $i, j \in M$ over which $\mathbf{p}(M)$ is bounded in a cycle and $q_{ij}(M) = q_{ji}(M) = 0$ for the remaining pairs in M . It follows from the rationalizability of $\mathbf{p}(\mathcal{N})^*$ and Lemma 4 in Appendix A.2, that $\mathbf{p}(M)$ is rationalizable by $\langle Q(M), \boldsymbol{\pi}(M) \rangle$ since $\delta_{kl}(\mathbf{p}(M)^*)q_{kl}(M)^* = \delta_{kl}(\mathbf{p}(M))q_{kl}(M)$ for all $k, l \in M$. By definition, $q_{ij}(M) = q_{ij}(M)^* > 0$ for all pairs $i, j \in C(M)$ since $\mathbf{p}(M)$ is bounded in a cycle over those pairs and the model is irreducible.

B Examples

B.1 Details on Example 1

We first show that $\frac{\rho(i, Q(\{i, j, k\}))}{\rho(j, Q(\{i, j, k\}))} > \frac{\rho(i, Q(\{i, j\}))}{\rho(j, Q(\{i, j\}))}$ whenever $q_{ik}(\{i, k\}) = q_{ik}(\{i, j, k\}) = 0$. The rationalizability of $\mathbf{p}(\mathcal{N})$ implies $\mathbf{p}(M)(I - Q(M)) = 0$ for all $M \in \mathcal{N}$. Therefore, it holds for all $j \in M$ and $M \in \mathcal{N}$ that

$$\sum_{i \neq j} \rho(j, Q(M))q_{ji}(M) - \sum_{i \neq j} \rho(i, Q(M))q_{ij}(M) = 0.$$

In particular, it holds $\rho(j, Q(N))q_{ji}(N) + \rho(k, Q(N))q_{ki}(N) = \rho(i, Q(N))q_{ij}(N)$. After applying Assumption 3 and rearranging, we obtain

$$(19) \quad \frac{\rho(j, Q(N))}{\rho(j, \{i, j\})} + \rho(k, Q(N))\frac{q_{ki}(N)}{q_{ji}(N)} = \frac{\rho(i, Q(N))}{\rho(i, \{i, j\})}.$$

Since the model is irreducible, the induced choice function is positive (Theorem 4). Thus, all terms in the equation are positive and the first result follows.

We now prove the second part of the statement that $q_{ki}(\{i, j, k\}) > q_{ji}(\{i, j, k\})$ if and only if $\rho(i, Q(\{i, j, k\})) > \rho(i, Q(\{i, j\}))$. Starting from Equation (19), we use that all choice probabilities from the set N should sum up to 1:

$$\frac{\rho(j, Q(N))}{\rho(j, \{i, j\})} + (1 - \rho(j, Q(N)))\frac{q_{ki}(N)}{q_{ji}(N)} = \frac{\rho(i, Q(N))}{\rho(i, \{i, j\})} + \frac{q_{ki}(N)}{q_{ji}(N)}\rho(i, Q(N)).$$

Rearranging the equality and using that $\rho(i, Q(N)) > \rho(i, Q(\{i, j\}))$, we obtain

$$\frac{\rho(j, Q(N))}{\rho(j, \{i, j\})}q_{ji}(N) + (1 - \rho(j, Q(N)))q_{ki}(N) > q_{ji}(N) + q_{ki}(N)\rho(i, \{i, j\}),$$

$$(q_{ji}(N) - q_{ki}(N))(\rho(j, Q(N)) - \rho(j, \{i, j\})) > 0.$$

Note that $\rho(i, Q(N)) > \rho(i, Q(\{i, j\}))$ together with the fact that all choice probabilities from a given menu sum up to 1 implies that $\rho(j, \{i, j\}) > \rho(j, Q(N))$ and the result follows.

B.2 Details on Example 2

Note that Assumption 3 implies $\frac{q_{ji}(M)}{q_{ij}(M)} = \frac{q_{ji}(\{i, j\})}{q_{ij}(\{i, j\})} = \frac{p(i, \{i, j\})}{p(j, \{i, j\})}$. Since the stochastic choice function on binary sets is the same, all rationalizing MSC models in Examples 2-4 are such that:

$$(20) \quad I - Q(M) = \begin{pmatrix} q_{ij} + q_{ik} + q_{il} & -q_{ij} & -q_{ik} & -q_{il} \\ -q_{ij} & q_{ij} + q_{jk} + q_{jl} & -q_{jk} & -q_{jl} \\ -q_{ik} & -\frac{3}{2}q_{jk} & q_{ik} + \frac{3}{2}q_{jk} + q_{kl} & -q_{kl} \\ -q_{il} & -q_{jl} & -\frac{2}{3}q_{kl} & q_{il} + q_{jl} + \frac{2}{3}q_{kl} \end{pmatrix}.$$

We show that no rationalizing MSC model will have positive transition probabilities to and from state k . Using the definition of rationalizability, it has to hold that

$$\begin{aligned} p(i, M)q_{ik} + p(j, M)q_{jk} - p(k, M)(q_{ik} + \frac{3}{2}q_{jk} + q_{kl}) + p(l, M)\frac{2}{3}q_{kl} &= 0, \\ q_{ik} + q_{jk} - 2(q_{ik} + \frac{3}{2}q_{jk} + q_{kl}) + \frac{2}{3}q_{kl} &= 0, \end{aligned}$$

where the second equation results from plugging in the stochastic choice function and simplifying. Since transition probabilities cannot be negative, the only solution is $q_{ik}(M) = q_{jk}(M) = q_{kl}(M) = 0$.

One can easily verify that $\mathbf{p}(M) = (0.2, 0.2, 0.4, 0.2)$ is the stationary distribution of an MSC model with the following transition probability matrix and $\pi(k, M) = 0.4$

$$Q_2(M) = \begin{pmatrix} 0.8 & 0.1 & 0 & 0.1 \\ 0.1 & 0.8 & 0 & 0.1 \\ 0 & 0 & 1 & 0 \\ 0.1 & 0.1 & 0 & 0.8 \end{pmatrix}.$$

Note that restricting the comparability between each one of the pairs results in another rationalizing MSC model when the initial distribution is the same. For example, if we let $R(M)$ be such that $r_{il}(M) = 0$ and all other off-diagonal elements equal to 1, $R(M) \odot Q_2(M)$ rationalizes $\mathbf{p}(M)$. This holds even if the comparability restriction is weak, for example $r_{il}(M) = 0$, $r_{ij}(M) = 2$ and all other off-diagonal elements equal to 1.

B.3 Details on Example 3

As pointed out in Appendix B.2, the transition probability matrix should satisfy Equation (20) and $\rho(\pi(M), Q(M))(I - Q(M)) = \mathbf{0}$. One can verify that the below transition probability matrix satisfies these requirements for $\mathbf{p}(M) = (0.25, 0.28, 0.2, 0.27)$:

$$Q_3(M) = \begin{pmatrix} 0.7 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.72 & 0.16 & 0.02 \\ 0.1 & 0.24 & 0.57 & 0.09 \\ 0.1 & 0.02 & 0.06 & 0.82 \end{pmatrix}.$$

B.4 Details on Example 4

First, we show that there is no rationalizing model which assigns a positive transition probability between the pairs (j, l) and (k, l) in either direction. One of the equations that a

rationalizing MSC model needs to satisfy is

$$p(i, M)q_{il} + p(j, M)q_{jl} + p(k, M)q_{kl} - p(l, M)(q_{il} + q_{jl} + \frac{2}{3}q_{kl}) = 0,$$

$$0.07q_{jl} + 0.06q_{kl} = 0,$$

where the second equation results from plugging in the stochastic choice function and simplifying. Hence, the only solution to the above equation is when $q_{jl}(M) = q_{kl}(M) = 0$.

As in the previous examples, the transition probability matrix should satisfy Equation (20) and $\rho(\pi(M), Q(M))(I - Q(M)) = \mathbf{0}$. The following matrix satisfies these requirements for $\mathbf{p}(M) = (0.24, 0.3, 0.22, 0.24)$:

$$Q_4(M) = \begin{pmatrix} 0.4 & 0.1 & 0.3 & 0.2 \\ 0.1 & 0.7 & 0.2 & 0 \\ 0.3 & 0.3 & 0.4 & 0 \\ 0.2 & 0 & 0 & 0.8 \end{pmatrix}.$$

References

- Adler, S. J., Schöniger, M. K., Lichters, M., & Sarstedt, M. (2024). Forty years of context effect research in marketing: a bibliometric analysis. *Journal of Business Economics*, 94(3), 437–466.
- Agranov, M., & Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1), 40–68.
- Apesteguia, J., & Ballester, M. A. (2013). Choice by sequential procedures. *Games and Economic Behavior*, 77(1), 90–99.
- Armel, C. K., Beaumel, A., & Rangel, A. (2008). Biassing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, 3(5), 396–403.
- Atalay, S., Bodur, H., & Rasolofoarison, D. (2012, 12). Shining in the center: Central gaze cascade effect on product choice. *Journal of Consumer Research*, 39.
- Baldassi, C., Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Pirazzini, M. (2020). A behavioral characterization of the Drift Diffusion Model and its multialternative extension for choice under time pressure. *Management Science*, 1–19.

- Blanchet, J., Gallego, G., & Goyal, V. (2016). A Markov chain approximation to choice modeling. *Operations Research*, 64(4), 886-905.
- Brady, R., & Rehbeck, J. (2016). Menu-dependent stochastic feasibility. *Econometrica*, 84(3), 1203–1223.
- Caplin, A., & Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, 6(1), 19–48.
- Caplin, A., Dean, M., & Leahy, J. (2018). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies*, 86(3), 1061-1094.
- Castillo, G. (2020). The attraction effect and its explanations. *Games and Economic Behavior*, 119, 123–147.
- Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Rustichini, A. (2022). Multinomial logit processes and preference discovery: Inside and outside the black box. *The Review of Economic Studies*, 90(3), 1155-1194.
- Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, 73, 1–17.
- Chen, Y.-L., Chen, J.-M., & Tung, C.-W. (2006). A data mining approach for retail knowledge discovery with consideration of the effect of shelf-space adjacency on sales. *Decision Support Systems*, 42(3), 1503-1520.
- Demirkan, Y. (2024). *Visual Choice*.
- Demuynck, T., & Seel, C. (2018). Revealed preference with limited consideration. *American Economic Journal: Microeconomics*, 10(1), 102–131.
- Drèze, X., Hoch, S. J., & Purk, M. E. (1994). Shelf management and space elasticity. *Journal of Retailing*, 70(4), 301-326.
- Dutta, R. (2020). Gradual pairwise comparison and stochastic choice. *Theoretical Economics*, 15(4), 1335–1364.
- Gossner, O., Steiner, J., & Stewart, C. (2021). Attention please! *Econometrica*, 89(4), 1717–1751.
- Grinstead, C. M., & Snell, J. L. (1997). Markov Chains. In *Introduction to probability* (chap. 11). American Mathematical Society.

- Hauser, J. R. (2014). Consideration-set heuristics. *Journal of Business Research*, 67(8), 1688–1699.
- Henderson, J., & Hollingworth, A. (1999). High-level scene perception. *Annual review of psychology*, 50, 243–271.
- Horan, S. (2010). *Sequential search and choice from lists*.
- Huang, B., Juaneda, C., Sénécal, S., & Léger, P.-M. (2021). “Now you see me”: The attention-grabbing effect of product similarity and proximity in online shopping. *Journal of Interactive Marketing*, 54, 1-10.
- Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1), 90–98.
- Keel, A. L., & Padgett, D. (2015). The effects of adjacent competitors and promotion on brand sales. *Journal of Consumer Marketing*, 32(1), 43–50.
- Kibris, Ö., Masatlioglu, Y., & Suleymanov, E. (2024). A random reference model. *American Economic Journal: Microeconomics*, 16(1), 155–209.
- Kovach, M., & Tserenjigmid, G. (2022). The focal Luce model. *American Economic Journal: Microeconomics*, 14(3), 378–413.
- Kress, R. (2014). Linear integral equations. In *Applied Mathematical Sciences* (3rd ed., Vol. 82). New York: Springer.
- Kumar, R., Tomkins, A., Vassilvitskii, S., & Vee, E. (2015). Inverting a steady-state. In *Proceedings of the eighth ACM international conference on web search and data mining* (p. 359–368). New York, NY, USA: Association for Computing Machinery.
- Levin, D. A., Peres, Y., & Wilmer, E. L. (2017). *Markov chains and mixing times* (2nd ed.). Providence, Rhode Island: American Mathematical Society.
- Luce, D. (1959). *Individual choice behavior: A theoretical analysis*. New York: John Wiley & Sons.
- Manzini, P., & Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82(3), 1153–1176.
- Masatlioglu, Y., & Nakajima, D. (2013). Choice by iterative search. *Theoretical Economics*, 8, 701–728.

- Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5), 2183–2205.
- Masatlioglu, Y., & Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, 121(1), 1–29.
- Masatlioglu, Y., & Suleymanov, E. (2021). Decision making within a product network. *Economic Theory*, 185–209.
- Mattis, N., Groot Kormelink, T., Masur, P., Moeller, J., & van Atteveldt, W. (2024). Nudging news readers: A mixed-methods approach to understanding when and how interface nudges affect news selection. *Digital Journalism*, 1–21.
- Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, 22(1), 67-74.
- Morimura, T., Osogami, T., & Idé, T. (2013). Solving inverse problem of markov chain with partial observations. *Advances in neural information processing systems*, 26.
- Noguchi, T., & Stewart, N. (2014). In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition*, 132(1), 44–56.
- Ok, E. A., Ortoleva, P., & Riella, G. (2015). Revealed (p)reference theory. *American Economic Review*, 105(1), 299–321.
- Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144(1), 190–206.
- Papi, M. (2012). Satisficing choice procedures. *Journal of Economic Behavior & Organization*, 84(1), 451–462.
- Pieters, R. (2008). A review of eye-tracking research in marketing. *Review of Marketing Research*, 4, 123–147.
- Ragain, S., & Ugander, J. (2016). Pairwise choice Markov chains. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 29). Curran Associates, Inc.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108.
- Ravid, D., & Stevenson, K. (2019). *Focus, then compare*.

- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101(2), 900–926.
- Roman, S. (2008). Advanced linear algebra. In S. Axler & K. A. Ribet (Eds.), *Graduate texts in mathematics* (3rd ed.). New York: Springer.
- Rubinstein, A., & Salant, Y. (2006). A model of choice from lists. *Theoretical Economics*, 1(1), 3–17.
- Russo, J. E., & Rosen, L. D. (1975). An eye-fixation analysis of multialternative choice. *Memory & Cognition*, 3, 267–276.
- Salant, Y., & Rubinstein, A. (2008). (A, f): Choice with frames. *The Review of Economic Studies*, 75, 1287–1296.
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6(12), 1317–1322.
- Tserenjigmid, G. (2019). Choosing with the worst in mind: A reference-dependent model. *Journal of Economic Behavior and Organization*, 157, 631–652.
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039–1061.
- Wedel, M., & Pieters, R. (2007). Informativeness of eye movements for visual marketing: Six cornerstones. *Visual marketing: From attention to action*.
- Zhang, J. (2017). *Sequential Luce with Categorization*.