Markov Stochastic Choice

Kremena Valkanova *

November 2, 2020

The most recent version can be found here.

Abstract

We propose a model of decision-making based on Markovian exploration of a choice set. It is inspired by experimental evidence that decision makers search sequentially by making stochastic pairwise comparisons. The model allows us to analyze the impact of a variety of interventions on final choices and to infer unobservable characteristics of the agent. We examine how choice can be influenced by the first fixation, comparability restrictions, time pressure, and by adding a dominated alternative. We show that the long-run choices of an agent are not susceptible to manipulation if the Markovian model is reversible and that such choices are consistent with the well-known Luce model. Further, we identify conditions on the long-run choice data that reveal the agent's consideration set and comparability restrictions. Finally, we show how one can isolate the effect of utility from salience on choice probability.

Keywords: search, bounded rationality, stochastic choice, pairwise comparison.

JEL-classification: D9, D11, D83.

---

*University of Zurich, Switzerland, e-mail: kremena.valkanova@econ.uzh.ch

I would like to thank my PhD supervisors Nick Netzer and Jakub Steiner for their guidance and constructive comments. I am grateful to 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.

1 Introduction

Classical choice theory assumes that decision makers choosing an alternative from a menu are able to perceive and evaluate all alternatives simultaneously and select the best among them according to a deterministic preference. In many real-life instances, however, this approach to decision-making is not appropriate either because the framing of the decision problem does not permit the individual to view all alternatives at once or because the size of the choice set and the complexity of the alternatives make the simultaneous comparison very difficult. In both of these situations, the agent is going to explore and evaluate the alternatives in the menu sequentially in an attempt to discover the set of available alternatives in the former case and to reduce cognitive load in the latter case.

Consider, for example, an individual shopping on an e-commerce website. Instead of displaying all alternatives at once, the platform suggests other related products for each alternative that the agent views. In this case, decision makers select the next alternative to view from the recommended list and continue to explore the choice set in this fashion. Analogous search dynamics through stochastic pairwise comparisons is documented in experiments using eye-tracking even when the whole choice set is presented simultaneously (Russo & Rosen, 1975; Reutskaja, Nagel, Camerer, & Rangel, 2011). Such evidence is in line with the idea that agents try to minimize the cost of thinking by relying on re-fixations (Shugan, 1980). The eye fixations happen very rapidly and are followed by jumps to nearby alternatives, and serve primarily for object identification (Henderson & Hollingworth, 1999; Rayner, 1998; Wedel & Pieters, 2007; Pieters, 2008).

These observations suggest that the framing of the decision problem has a potentially large influence on the final purchasing decisions. We are therefore interested in understanding and predicting the impact of different interventions on final choices, and inferring unobservable characteristics of the decision problem and the agent from choice data. Such insights are valuable for designing better-targeted interventions and increasing welfare. In order to address these issues we propose and study a model of decision-making, which incorporates the previously mentioned experimental findings[1] and at the same time offers a simple and elegant way of utilizing the much better understood models of binary choice to analyze multi-alternative choice.

We model the exploration of the choice set as a discrete-time Markov chain in which each state represents the agent's current best alternative. The process starts with the agent randomly selecting an alternative to view according to his initial beliefs and knowledge. In the next period, the incumbent alternative is compared to a competitor and the decision maker makes a transition to the competitor according to an incumbent-specific probability distribution. This distribution is independent of past incumbents, thus reducing the cognitive load on the decision maker. Each transition consists of two steps - first, the agent's attention is drawn towards a competitor and second, the agent compares the incumbent and the competitor and either accepts or rejects the competitor. We think of the first step as being influenced by the salience and of the second step as being related to the values of the incumbent and the competitor. We assume that the transition probabilities are minimally consistent across choice sets, i.e., the relative transition probabilities for each pair are independent of the other options in the menu. The decision-making process may be terminated each period with some stopping probability after which the agent chooses the current best alternative.

An important distinction of our model from other models featuring a sequential exploration of the choice set is that we allow both the search order and the transitions between alternatives to be stochastic. Different sources for this stochasticity are possible – it might be that the person's fixations on alternatives are random and do not depend on alternatives' values (as suggested in Reutskaja et al. (2011)) and/or the transition conditional on fixating on a particular competitor is noisy. The literature on stochastic choice has

[1]Other experimental findings motivating the choice model include the observations that (i) frequent eye-movements to some alternatives correlate with an increased choice frequency of these alternatives known as the gaze-cascade effect (Noguchi & Stewart, 2014; Shimojo, Simion, Shimojo, & Scheier, 2003); (ii) fixations are more frequent and longer on the chosen alternative with the last fixation being on the chosen alternative (Armel, Beaumel, & Rangel, 2008).

suggested multiple explanations for the latter including random shocks to the preferences, evidence accumulation and bounded rationality, and random mistakes among others (see for example Agranov and Ortoleva (2017)). Stochastic transitions between alternatives are also beneficial for the broader exploration of the choice set as they prevent the decision maker from getting trapped at alternatives which have the highest utility locally but not globally.

The aspect of time pressure is reflected in the model through the stopping probability parameter. Since less time pressure leads to more unbiased choices (Svenson & Maule, 1993), we devote special attention to the case in which the stopping probability parameter goes to zero. We show that the stochastic choice function of a decision maker, who is not under time pressure, converges towards the limiting distribution of the Markov chain. We refer to this procedure as the limiting Markov stochastic choice (MSC) model and use it to infer unobservable parameters of the model and the framing of the decision problem. As customary in the stochastic choice literature, we assume that choices from all subsets of the set of alternatives are observable, although for the majority of the results only data on binary sets and one larger choice set is sufficient.

Our first set of results concerns the ability to influence choice behavior through (i) adding a dominated alternative, (ii) increasing the probability of the initial fixation on a target alternative, (iii) restricting the comparability of certain pairs of alternatives, (iv) adding time pressure, and the interactions between these types of interventions. We study how choice generated by the limiting MSC model responds to the addition of an alternative to a menu which is asymmetrically dominated, meaning that it never wins a comparison against some target alternative. There is considerable experimental evidence that in such cases the choice probability of the target increases, which is referred to as the attraction effect (Huber, Payne, & Puto, 1982). The attraction effect is considered a behavioral irregularity as it cannot be explained by conventional random utility models. We identify conditions on the transition probabilities under which the limiting MSC model generates the attraction effect. Our theory predicts that restricting the direct comparison of the target and the dominated alternative eliminates the attraction effect.

There is consistent experimental evidence that drawing attention to a certain alternative increases its choice probability.[2] This can be achieved through advertising, packaging, or arranging alternatives on a screen or shelf to attract more attention to a target alternative. Our model predicts that time pressure is a crucial factor for this type of intervention to be successful. Specifically, if the decision maker is under time pressure, an increase in the initial fixation probability of the target alternative always leads to a higher choice probability of the target. This result holds irrespectively of the transition probabilities and the exact stopping probability. However, the effect of the initial fixation diminishes in the absence of time pressure when the decision maker can explore the whole choice set.

Apart from influencing the initial fixation, the positioning of the alternatives in the choice set might influence the decision maker through a different channel. Since agents tend to compare only those pairs of alternatives which are placed closer together (Chandon, Hutchinson, Bradlow, & Young, 2009), the arrangement of the choice set affects the pairs of alternatives which the decision maker cannot compare. We focus on identifying limiting MSC models for which the generated stochastic choice function is not susceptible to any comparability restriction, and thus it is not affected by the arrangement of the choice set and the cognitive limitations of the decision maker. We find that this feature characterizes limiting MSC models for which the underlying Markov chain is reversible, meaning that on average the number transitions between each pair of alternatives is equal in both directions.

The effect of time pressure on the choice probabilities in larger choice sets can be non-monotone and difficult to predict in general. However, time pressure can be used to influence choices in a predictable way through comparability restrictions. In particular, we consider the effect of restricting comparability between a pair of alternatives in a triple. We find that the difference between the choice ratios and the transition ratios of a pair determines the effect of restricting the pair comparability on choices. This result implies that data from pairs and triples can be used by policy makers to arrange alternatives on

[2] See Orquin and Mueller Loose (2013) for a review.

a list in order to increase the choice probability of a desirable alternative.

Since we have previously established the importance of the class of reversible limiting MSC models, we are interested in identifying such models from choice data. We find that in the case when all transitions happen with positive probability, the model is characterized by the well-known positivity and Independence of irrelevant alternatives (IIA) axioms. These properties are known to characterize the Luce model (Luce, 1959), which is one of the classical models of stochastic choice. We provide an extension of the characterization result and identify a family of stochastic choice functions which can only be rationalized with a reversible limiting MSC model. This result shows that observable choices can inform policy makers whether the decision maker is affected by comparability restrictions or not in the absence of time pressure.

In addition, the agent's choices are not influenced by different arrangements of the choice set if he is not subject to cognitive limitations that restrict comparability between any pair of alternatives. We provide a characterization of such limiting MSC models with only positive transitions. Based on our results an analyst can conclude whether or not the decision maker faces comparability restrictions and whether these have an influence on the choice function. Furthermore, we characterize limiting MSC models which allow the decision maker to explore the whole choice set even if some comparisons are restricted. A violation of the condition on the data suggests the existence of consideration sets. What is more, we show that choice data reveals the exact pairs of alternatives which the decision maker never compares directly. This allows an analyst to infer which alternatives belong to the same consideration set.

Our last set of results concerns isolating the effect of an alternative's value from its salience on its choice probability. For this reason, we are interested in inferring the acceptance probabilities from choice data. We show that this can be done for limiting MSC models for which the relative salience between all pairs of alternatives is symmetric, for example, if the relative salience is determined by the distance between each pair. Moreover, we can determine for every two competitors the one which is more likely to attract attention away from a given incumbent. This insight is useful for sellers to identify their closest competitors from consumer perspective.

The remainder of the paper is structured as follows: In the next section we define the modeling framework with and without time pressure. Section 3 is devoted to analyzing the impact that different interventions have on the final stochastic choices. In Section 4 we characterize different variations of the model without time pressure. Section 5 focuses on inferring information about transition probabilities from choice data. Related literature is discussed in Section 6. Section 7 concludes.

2 Model

Let N be a finite set of all possible alternatives. A menu M is a non-empty subset of N and the set of all menus is denoted by $\mathcal{M}$. The probability of choosing an alternative i from menu M is given by the stochastic choice function $p : N \times \mathcal{M} \to [0,1]$. It has the following properties: (1) if $p(i,M) > 0$, then $i \in M$ and (2) $\sum_{i \in M} p(i,M) = 1$. We denote a row vector of choice probabilities from a menu M by $\boldsymbol{p}(M)$. The stochastic choice function for all menus $M \in \mathcal{M}$ is denoted by $\boldsymbol{p}$. Since $\boldsymbol{p}$ is comprised of the actual choices of the decision maker, it is observable by an analyst.

2.1 Baseline Markov stochastic choice model

A decision maker is assumed to follow the sequential procedure in discrete time described below when choosing an alternative from a menu. The agent is endowed with a probability distribution over the alternatives in the menu such that the decision-making process starts at alternative i with probability $\pi(i,M)$. At time period $\tau = 0$ the agent randomly draws an alternative according to the distribution $\boldsymbol{\pi}(M)$, say alternative $i \in M$.

In the next period, the incumbent alternative i is compared against another available alternative. The decision maker focuses on a competitor j with probability $s_{ij}(M)$. This parameter captures the ability of j to attract attention away from i and can be interpreted as i's relative salience against j. Then, the two alternatives are compared independently from the other alternatives in the menu. The competitor wins the comparison with probability t_{ij} and loses with probability $1 - t_{ij}$, which we refer to as the acceptance or rejection probabilities, respectively. Therefore, the total probability to switch incumbents from i to j is given by:

$$q_{ij}(M) = s_{ij}(M)t_{ij} \geqslant 0.$$

The probability that i remains the incumbent[3] is given by $q_{ii}(M) = 1 - \sum_{j \neq i} q_{ij}(M) > 0$. We denote the transition probability matrix with $Q(M)$. Note that transition probabilities between a pair of states may be menu-dependent, hence it may be that $q_{ij}(M) \neq q_{ij}(M')$. Furthermore, we assume that if the choice set is binary, the decision maker is able to make transitions between the two alternatives, that is $q_{ij}(\{i,j\}) = 0$ implies that $q_{ji}(\{i,j\}) > 0$ for all $i, j \in N$.

We assume a minimal form of consistency in the transitions between each pair of alternatives across choice sets, namely that the ratio of their relative salience is not affected when the menu is enlarged. We call this condition *transition ratio independence of irrelevant alternatives (TR-IIA)* and define it formally in terms of the transition probabilities $q_{ij}(M)$ below.[4]

Transition ratio IIA. *For all $M \in \mathcal{M}$ the transition probabilities satisfy*

$$q_{ij}(\{i,j\})q_{ji}(M) = q_{ji}(\{i,j\})q_{ij}(M) \quad \forall i, j \in M. \tag{2.1}$$

Since the acceptance probabilities t_{ij} are menu-independent, the above property refers to the way relative salience is affected by enlarging the choice set. TR-IIA is in particular satisfied when the relative salience parameters between all pairs of alternatives are symmetric, i.e., $s_{ij}(M) = s_{ji}(M)$ for all $i, j \in M$. This assumption arises naturally

[3]Note that the probability to remain at the same incumbent is assumed to be positive, because in some periods the incumbent might not be challenged by another alternative.

[4]We define TR-IIA as a product rather than a ratio of transition probabilities to account for the case in which transitions occur with zero probability.

when the probability to compare a pair of alternatives depends on the distance between alternatives, which can be understood as a physical distance (positioning of items on a shelf or a screen) or a subjective distance in terms of similarity or ease of comparison (Russo & Rosen, 1975; Reutskaja et al., 2011; Cerreia-Vioglio, Maccheroni, Marinacci, & Rustichini, 2018). We assume that all MSC models satisfy TR-IIA for the remainder of the paper.

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

The described choice procedure is equivalent to a finite-state discrete-time Markov chain with states represented by the alternatives in each menu, initial distribution $\boldsymbol{\pi}(M)$, transition probability matrix $Q(M)$, and stopping probability α.

The generated stochastic choice function for a menu M is computed as shown below:

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

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 an incumbent at that time period. The following lemma presents a simplification of the above expression.

Lemma 1. *Stochastic choice functions generated by a baseline MSC model satisfy*

$$\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha \boldsymbol{\pi}(M)(I - (1-\alpha)Q(M))^{-1}. \tag{2.2}$$

Proof. Let $\tilde{Q}(M) = (1-\alpha)Q(M)$. Then, $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha \boldsymbol{\pi}(M) \sum_{\tau=1}^{\infty} \tilde{Q}(M)^{\tau-1}$. 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 $\sum_{\tau=1}^{\infty} \tilde{Q}(M)^{\tau-1} = (I - \tilde{Q}(M))^{-1}$ (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, $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha \boldsymbol{\pi}(M)(I - \tilde{Q}(M))^{-1}$. □

An alternative way of defining the baseline MSC model is as an absorbing Markov chain in which for each alternative in the menu there is one transient and one absorbing state.[5] This equivalence implies that Lemma 1 is a corollary of existing results from the literature on absorbing Markov chains. Nevertheless, the proof of the lemma is provided here for convenience.

2.2 Limiting Markov stochastic choice model

Consider the case in which the probability that the decision process is terminated approaches zero, hence the pressure to make a decision is very limited. If the Markov chain underlying a baseline MSC model is irreducible and aperiodic[6], hence ergodic, a limiting distribution exists and it is the unique stationary distribution. Naturally, if we let the decision-making process run for a long time by letting the stopping probability approach zero, the generated stochastic choice function converges to the stationary distribution of the Markov chain and does not depend on the initial distribution as shown below.

Proposition 1. *A stochastic choice function $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M))$ generated by an ergodic baseline MSC model converges to the stationary distribution of a Markov chain with transition probability matrix $Q(M)$ as $\alpha \to 0$.*

Proof. See Appendix A. □

A baseline MSC model for which $\alpha \to 0$ we call a limiting MSC model and denote its generated stochastic choice function by $\boldsymbol{\rho}(Q(M))$ for all $M \in \mathcal{M}$. Proposition 1 implies that the choice function satisfies the equality $\boldsymbol{\rho}(Q(M))(I - Q(M)) = \mathbf{0}$. Below we provide

[5]One can easily verify that the two models are equivalent by letting the transition probability matrix between transient states be $Q^*(M) = (1 - \alpha)Q(M)$ and the absorption probability matrix be $A(M) = \alpha I$, and use the fact that the stationary distribution is given by $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}^*(M), Q^*(M)) = \boldsymbol{\pi}^*(M)(I - Q^*(M))^{-1}A(M)$ (Grinstead & Snell, 1997).

[6]Note that the assumption $q_{ii}(M) > 0$ for all $i \in M$ ensures that the Markov chain is aperiodic. If the model is irreducible, the process can reach any state irrespectively of the initial state, hence there is only one communicating class. Since the Markov chain is finite, irreducibility implies positive recurrence.

the definition of a rationalizable stochastic choice function with both the baseline and limiting MSC models.

Definition 1. *A stochastic choice function $\boldsymbol{p}$ is rationalizable with*

- *a baseline MSC model if there exists a tuple $\langle Q(M), \boldsymbol{\pi}(M), \alpha\rangle$ such that $\boldsymbol{p}(M) = \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M))$ for all $M \in \mathcal{M}$ and*
- *a limiting MSC model if there exists a matrix $Q(M)$ such that $\boldsymbol{p}(M) = \boldsymbol{\rho}(Q(M))$ for all $M \in \mathcal{M}$.*

The following example shows the different rationalizability conditions for the baseline and limiting MSC model on binary menus.

Example 1. A stochastic choice function rationalizable with a baseline MSC model satisfies

$$p(i, \{i, j\}) = \frac{\alpha\pi(i, \{i, j\}) + (1 - \alpha)q_{ji}(\{i, j\})}{\alpha + (1 - \alpha)(q_{ij}(\{i, j\}) + q_{ji}(\{i, j\}))}. \tag{2.3}$$

On the other hand, if the stochastic choice function is rationalizable with a limiting MSC model then

$$p(i, \{i, j\}) = \frac{q_{ji}(\{i, j\})}{q_{ij}(\{i, j\}) + q_{ji}(\{i, j\})}, \tag{2.4}$$

and hence the ratio of transition probabilities between i and j is observable in the limiting model as it is equal to the ratio of the respective choice probabilities.

3 Choice architecture

The MSC model is useful to study and predict the effects that the framing of the decision problem has on stochastic choices. We focus on four main interventions – adding an asymmetrically dominated alternative to the choice set, manipulating the probability of the initial fixation, altering the positioning of the alternatives in the choice set, and varying the time pressure.

3.1 Effect of adding a dominated alternative

In this section we analyze the effect of an asymmetrically dominated alternative in the choice set. An alternative is asymmetrically dominated if the menu contains another alternative that surpasses the dominated one for all attributes, which does not hold for all other alternatives in the menu. A large body of experimental literature shows that the choice probability of the dominant alternative is higher in the presence of the dominated alternative (see Castillo (2020) for a recent review). This phenomenon is known as the attraction or decoy effect and was first documented in Huber et al. (1982). The finding is puzzling because it shows a violation of the regularity axiom, according to which the choice probability of an alternative cannot increase when the choice set is enlarged, and thus cannot be explained with any conventional random utility model.

Although our framework does not model attributes explicitly, we can capture the asymmetric dominance relation between the alternatives in the menu through the pairwise comparisons between them. As in most of the experimental literature, we will analyze the stochastic choices between a pair of alternatives i and j, which we will compare to the choices from a triple i, j, k, where k is dominated by i, but not by j. Because of the dominance relation between i and k, it has to hold that $t_{ki} > 0$, but $t_{ik} = 0$, therefore, $q_{ik}(\{i, k\}) = q_{ik}(\{i, j, k\}) = 0$. The trade-off between the pairs i, j and j, k implies a positive acceptance probability for each alternative. Therefore, we assume that $t_{ij} \in (0, 1)$ and $t_{jk} \in (0, 1)$. As we show below, the MSC model is flexible enough to generate the attraction effect even in the absence of time pressure.

Proposition 2. *Let $\boldsymbol{\rho}$ be a stochastic choice function over $N = \{i, j, k\}$ generated by a limiting MSC model, where $q_{ik}(\{i, k\}) = 0$. Then, $\rho(i, Q(\{i, j, k\})) > \rho(i, Q(\{i, j\}))$ iff $q_{ki}(\{i, j, k\}) > q_{ji}(\{i, j, k\})$.*

Proof. See Appendix B.1. □

Therefore, the target alternative is chosen more frequently from the triple whenever the probability to transition from decoy k to target i is greater than transitioning from j

to i. This is the case independently of the transition probabilities between j and k. Furthermore, in the special case when the relative salience of all alternatives is equal, and in particular, $s_{ki}(\{i,j,k\}) = s_{ji}(\{i,j,k\})$, the inequality $q_{ki}(\{i,j,k\}) > q_{ji}(\{i,j,k\})$ is always fulfilled, and hence the choice probability of the target is always higher in the larger set.

Proposition 2 shows that the attraction effect arises only when a specific condition on the transition probabilities is satisfied. This prediction is consistent with the heterogeneity of choice behaviors in the presence of dominated alternatives, causing some researchers to question the practical relevance of the attraction effect (Frederick, Lee, & Baskin, 2014; Yang & Lynn, 2014). Following the prediction of our model, the attraction effect can be reduced by making the comparison between the dominated and target alternatives less likely, thus reducing $s_{ki}(\{i,j,k\})$. One way to achieve that is by displaying the choice set as a list with the first and last alternative being the target and the dominated alternative, respectively. This presentation decreases the likelihood to compare the two alternatives directly and is thus expected to reduce the attraction affect.

3.2 Effect of the initial distribution

We now examine the effect of the initial fixation on the stochastic choice function. In particular, consider the situation in which a seller can influence the initial distribution such that it is more likely to start exploring the choice set at a particular target alternative and less likely to start at any other alternative. Such manipulations can be easily implemented in practice by heavily advertising the target, or by placing it in the center of a shelf, which is known to attract consumer attention (Chandon et al., 2009; Orquin & Mueller Loose, 2013).

As shown in Proposition 1, in the absence of time pressure, the generated stochastic choice function does not depend on the initial fixation when the agent is able to explore the whole choice set. However, with time pressure, the choice probability of a target alternative increases in response to an increase in the likelihood of the initial fixation irrespectively of the transition probability matrix $Q(M)$.

Proposition 3. *Fix $M \in \mathcal{M}$ and consider two baseline MSC models $\langle Q(M), \boldsymbol{\pi}(M), \alpha\rangle$ and $\langle Q(M), \boldsymbol{\pi}'(M), \alpha\rangle$. Let $\Delta\boldsymbol{\pi}(M) = \boldsymbol{\pi}'(M) - \boldsymbol{\pi}(M)$. It holds that*

$$\begin{aligned} &\Delta\pi(i, M) > 0 \\ &\Delta\pi(j, M) \leqslant 0, \quad \forall j \neq i \end{aligned} \implies \rho(i, \alpha, \boldsymbol{\pi}'(M), Q(M)) > \rho(i, \alpha, \boldsymbol{\pi}(M), Q(M)).$$

Proof. See Appendix B.2. □

This result matches to a great extent the experimental findings of Reutskaja et al. (2011) and Armel et al. (2008). The authors observe that the initial fixation does not depend on the value of the alternatives in the choice set and could be manipulated by the seller, for example by changing the location on the display or the packaging to attract more attention. They show that the first fixation effect does not wear off and has a substantial influence on the final choices.

3.3 Comparability restrictions

As previously mentioned, experimental studies on eye fixations during decision-making show that the jumps between fixations occur mostly to nearby alternatives (Chandon et al., 2009). This finding suggests that if alternatives are placed far apart the decision maker cannot compare certain pairs of alternatives directly. Hence, different arrangements of the alternatives might affect the ability of the agent to make specific pairwise comparisons.

Figure 1: Limited ability to compare directly alternatives placed on a grid.

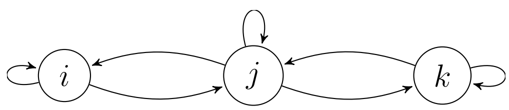
For example, consider a decision problem in which the alternatives are displayed on a grid as illustrated in Figure 1. Assume that the decision maker cannot compare directly alternatives which are placed diagonally to each other. The arrows indicate the comparisons that the agent is able to make directly. The two figures differ in the positioning of alternatives k and l, which leads to differences in the possible direct transitions. For example, we see that in the figure on the left-hand side, i and l can be compared directly, but in the figure on the right, this is not possible. Such scenarios can be captured in our model since they imply that if two alternatives i and j cannot be compared directly, the transition probability between them is equal to zero in a particular menu M, hence $q_{ij}(M) = q_{ji}(M) = 0$.

Alternatively, if we consider the situation of an online shopper, whose decision-making procedure is guided by the recommendation algorithm of an e-commerce platform, the restrictions of the pairwise comparisons are directly determined by the set of related products displayed. The designer's choice on how to present the decision problem may affect, intentionally or not, the final choice probabilities. Therefore, we are interested in characterizing the decision process of agents whose choices are not influenced by the arrangement of the alternatives in the choice set.

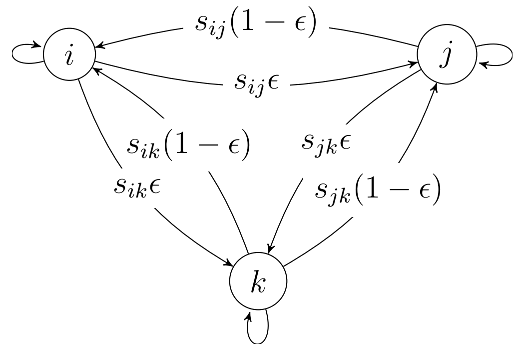
Definition 2. *A "manipulation by comparability restrictions" of a limiting MSC model $Q(M)$ makes the direct comparison for at least one pair of alternatives $i, j \in M$ impossible, such that $q_{ij}(M) \neq q'_{ij}(M) = q'_{ji}(M) = 0$, while preserving the set of alternatives that can be reached from each starting point. It satisfies the following consistency constraints:*

- *all non-restricted pairs $k, l \in M$ satisfy $q_{kl}(M) > 0 \implies q'_{kl}(M) > 0$ and $q_{kl}(M) = 0 \implies q'_{kl}(M) = 0$ and*
- *$\boldsymbol{\pi}(M)$ is independent of comparability restrictions.*

In fact, TR-IIA implies that $\frac{q'_{kl}(M)}{q_{kl}(M)} = \frac{q'_{lk}(M)}{q_{lk}(M)}$ for all distinct $k, l \in M$ for which comparability is not restricted. Note that whenever the whole choice set cannot be explored (i.e., the Markov chain is not irreducible), the stochastic choice function depends on the initial distribution. For this reason we assume that the initial distribution $\boldsymbol{\pi}(M)$ is independent



(a) Illustration of Example 2.



(b) Illustration of Example 3.

Figure 2: Reversible and non-reversible Markov chains.

of comparability restrictions.

We identify a class of limiting MSC models which are not susceptible to such manipulations in the absence of time pressure. The property that the limiting MSC model needs to satisfy is called *reversibility*, which defines one of the most extensively studied classes of Markov chains (Levin, Peres, & Wilmer, 2017). A Markov chain is reversible if it satisfies the following detailed balance conditions

$$q_{ji}(M)\rho(j, Q(M)) = \rho(i, Q(M))q_{ij}(M), \forall i, j \in M. \tag{3.1}$$

The detailed balance equations postulate that the flow of probability mass is balanced for each pair of states. A necessary and sufficient condition for reversibility is Kolmogorov's criterion. It states that a Markov chain is reversible if and only if

$$q_{i_1 i_2}(M)q_{i_2 i_3}(M)\dots q_{i_I i_1}(M) = q_{i_2 i_1}(M)q_{i_3 i_2}(M)\dots q_{i_1 i_I}(M) \tag{3.2}$$

for any finite sequence of states $i_1, i_2, \dots, i_I \in M$. We call a limiting MSC model reversible on M if the corresponding Markov chain with states M satisfies reversibility. We will say that a limiting MSC model is reversible if it is reversible on all $M \in \mathcal{M}$. Below we give examples for limiting MSC models that are reversible and non-reversible on a triple $\{i, j, k\}$.

Example 2. Imagine that the choice set is displayed as a list, which restricts the decision-maker's ability to compare all pairs of alternatives. In particular, each alternative can be directly compared only to its neighbors in the linear order as illustrated in Figure 2a for a menu of 3 alternatives. Such Markov chains are called birth-death processes (Levin et al., 2017). In this case, Kolmogorov's criterion is satisfied trivially and therefore, the limiting MSC model is reversible, independent of the other parameters.

Example 3. Consider a decision maker for whom the relative salience parameters are symmetric, that is $s_{ij}(M) = s_{ji}(M)$. Moreover, the alternatives can be arranged in a linear order such that the conditional acceptance probability $t_{ij} = \epsilon$ whenever i precedes j in the linear order, as if the agent has a deterministic preference over the alternatives, but makes mistakes with some small probability ϵ. The corresponding Markov chain is visualized for a menu of 3 alternatives in Figure 2b. Such limiting MSC models are always going to be non-reversible on M as they always violate Kolmogorov's criterion.

We state the result on comparability restrictions below.

Proposition 4. *If a limiting MSC model $Q(M)$ is reversible on M, the generated stochastic choice function $\boldsymbol{\rho}(Q(M))$ is not susceptible to manipulation by comparability restrictions.*

Proof. See Appendix B.3. □

The above result implies that any potential limitation in the ability of the agent to compare certain pairs of alternatives does not affect the stochastic choices if the decision-making process evolves as a reversible limiting MSC model. Moreover, if the decision maker would follow a reversible limiting MSC model in the hypothetical case when all pairwise comparisons happen with positive probability, the different arrangements of the alternatives in the choice set as illustrated in Figure 1 do not influence actual choices, even if the decision maker can compare only nearby alternatives. Similarly, for such decision makers the number of displayed related products on an e-commerce website does not matter for final choices as long as the agent can explore the whole choice set.

Inferring the reversibility property of a rationalizing limiting MSC model is, in general, not possible unless we add a stronger assumption on the consistency of transition probabilities across manipulations of the decision problem of a given menu and require that data on all manipulations is observable.

Definition 3. *A "strict manipulation by comparability restrictions" is a manipulation satisfying* $q'_{kl}(M) = cq_{kl}(M)$*, where* $c > 0$ *for all* $k, l \in M, k \neq l$ *for which comparability is not restricted.*

This consistency assumption concerning the pairs which are not subject to comparability restrictions is notably stronger than the one assumed previously, namely that $\frac{q'_{kl}(M)}{q_{kl}(M)} = \frac{q'_{lk}(M)}{q_{lk}(M)}$. Nevertheless, it is an intuitive assumption if one considers the acceptance probabilities t_{kl} as independent from comparability restrictions and the relative salience parameters $s_{kl}(M)$ as being determined by the distance between pairs. Then, the transition probability for the non-restricted pairs is not affected by the manipulation.

We can now state the characterization of limiting MSC models which are reversible on a particular menu M.

Proposition 5. *A stochastic choice function* $\boldsymbol{p}(M)$ *is not susceptible to any strict manipulation by comparability restrictions iff it is generated by a limiting MSC model which is reversible on* M*.*

Proof. See Appendix B.4. □

Note that to conclude that a stochastic choice function is generated by a reversible limiting MSC model (i.e., it is reversible on all $M \in \mathcal{M}$), an analyst needs to verify that the stochastic choice function does not change in response to *any* of the possible comparability restrictions when the menu equals the complete set of alternatives $M = N$. This is the case, because if stochastic choice data reveals that it is rationalizable by a limiting MSC model, which is reversible on N, then TR-IIA implies that the limiting MSC model should be reversible on all subsets. Proposition 4 would then imply that choice on all subsets of N is also not susceptible to manipulations by comparability restrictions.

3.4 Effect of time pressure

In the baseline MSC model the effect of time pressure is captured by the stopping probability parameter α. We have seen in Example 1 that the stochastic choice function is always monotone in α irrespective of the other parameters whenever the choice set is binary. This is not always the case when the choice set is larger. Then, the effect of manipulating the time pressure is, in general, difficult to predict.

In fact, comparability restrictions can be used to predict the effect of time pressure as we show below. To ease the notation we let $\rho(i, \alpha, \boldsymbol{\pi}(M), Q(M)) = \rho_{\alpha,\boldsymbol{\pi}}(i, Q(M))$.

Proposition 6. *Let $M = \{i, j, k\}$ and consider the baseline MSC models $\langle Q(M), \boldsymbol{\pi}(M), \alpha\rangle$ and $\langle Q^{'}(M), \boldsymbol{\pi}(M), \alpha\rangle$ such that $q_{ij}(M) = q'_{ij}(M) > 0$ for all $i, j \in M$ except $q_{ik}(M) \neq q'_{ik}(M) = q'_{ki}(M) = 0$. It holds that*

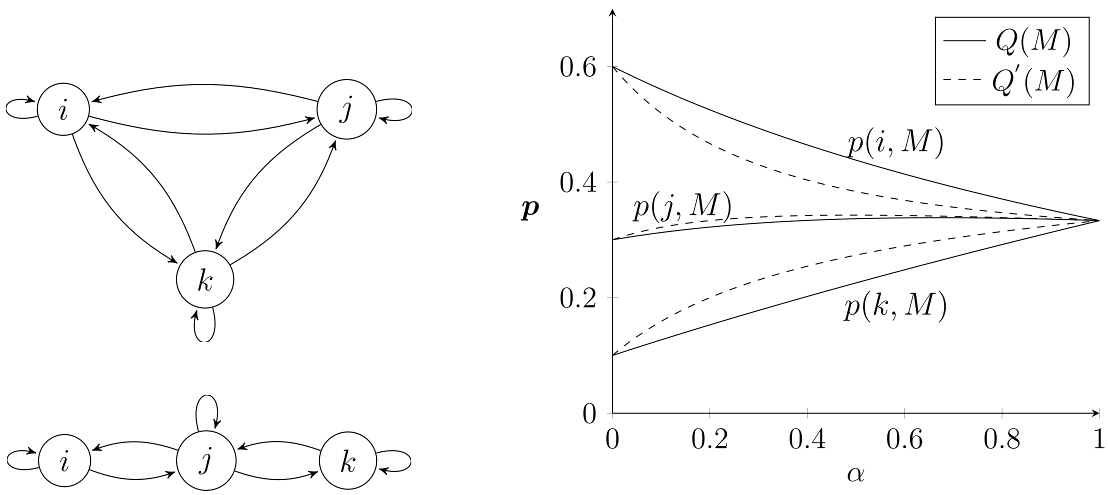
$$\frac{\rho_{\alpha,\boldsymbol{\pi}}(i, Q^{(')}(M))}{\rho_{\alpha,\boldsymbol{\pi}}(k, Q^{(')}(M))} < \frac{q_{ki}(M)}{q_{ik}(M)} \iff \begin{array}{l} \rho_{\alpha,\boldsymbol{\pi}}(i, Q(M)) - \rho_{\alpha,\boldsymbol{\pi}}(i, Q^{'}(M)) > 0, \\ \rho_{\alpha,\boldsymbol{\pi}}(k, Q(M)) - \rho_{\alpha,\boldsymbol{\pi}}(k, Q^{'}(M)) < 0 \end{array}$$

for any stopping probability $\alpha \in (0, 1)$ and initial distribution $\boldsymbol{\pi}(M)$.

Proof. See Appendix B.5. □

Proposition 6 has two important implications for nudging and predicting choice behavior. Imagine that an analyst can observe choices on binary sets when time pressure is sufficiently small or absent. As we have seen in Example 1, the ratio of binary choice probabilities approaches the ratio of the transition probabilities between any pair of alternatives as α goes to zero. Therefore, the ratio $\frac{q_{ki}(\{i,k\})}{q_{ik}(\{i,k\})}$ in the above formula can be approximated from choice data, which in turn is equal to $\frac{q_{ki}(M)}{q_{ik}(M)}$, because of TR-IIA. If the analyst observes either $\boldsymbol{\rho}_{\alpha,\boldsymbol{\pi}}(Q(\{i, j, k\}))$ or $\boldsymbol{\rho}_{\alpha,\boldsymbol{\pi}}(Q^{'}(\{i, j, k\}))$, Proposition 6 can be used to predict the effect of either restricting or enabling the direct comparison between i and k.

The result can also be applied to infer information about the ratio $\frac{q_{ki}(\{i,k\})}{q_{ik}(\{i,k\})}$, when choice data for the set $\{i, k\}$ cannot be obtained. Following the above argument, an analyst



(a) Structure of the transitions of the two models $Q(M)$ (top) and $Q^{'}(M)$ (bottom).

(b) Stochastic choice over a triple as a function of the stopping probability α.

Figure 3: Illustration of Example 4.

needs to observe choices from the triple when the comparison between i and k is permitted and restricted, which would provide an upper or lower bound on the ratio $\frac{q_{ki}(M)}{q_{ik}(M)}$. This information can be used to bound the choice probability from the binary set in the absence of time pressure.

Note that we do not make any assumptions about the properties of the transition probability matrices. Therefore, even decision makers following a reversible limiting MSC model, whose stochastic choices cannot be manipulated by comparability restrictions in the absence of time pressure, can be subject to such manipulations under time pressure. We illustrate this case with the following example.

Example 4. Consider two baseline MSC models $\langle Q(M), \boldsymbol{\pi}(M), \alpha\rangle$ and $\langle Q^{'}(M), \boldsymbol{\pi}(M), \alpha\rangle$ with $N = \{i, j, k\}$ with uniform $\boldsymbol{\pi}(M)$. Assume that there is a utility function over the alternatives in the menu such that $u(i) = 6$, $u(j) = 3$, and $u(k) = 1$. We let $q_{ji}(M) = \frac{u(i)}{|M-1|(u(i)+u(j))}$ for all $i, j \in M$ and $q'_{ji}(M) = q_{ji}(M)$ for all $i, j \in M$ except $q'_{ki}(M) = q'_{ik}(M) = 0$. The two baseline MSC models are therefore reversible. The structure of the transitions is illustrated in Figure 3a, where an arrow between a pair of alternatives means that a transition occurs with positive probability in the specified direction. If $\alpha = 1$, the choice probabilities coincide with the initial distribution and are equal to 1/3 for all alternatives. As $\alpha \to 0$, the stochastic choice function approaches the limiting distribution. Since both models are reversible and ergodic, $\lim_{\alpha\to 0} \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}, Q(M)) = \lim_{\alpha\to 0} \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}, Q^{'}(M)) = \{0.6, 0.3, 0.1\}$. In this example, the choice probability of j is higher in the birth-death process for all $\alpha > 0$. It can be clearly seen in Figure 3b that the effect of the initial distribution in the choice probabilities of the corner alternatives i and k is more significant in the case when the two cannot be compared directly.

4 Characterizations of the limiting MSC model

The results in the previous section demonstrate the influence that restrictions on pairwise comparisons have on final stochastic choices and the importance of the class of reversible limiting MSC models as being resistant to restrictions of comparability. Since such restrictions depend largely on the cognitive capacity of the decision maker, they might not always be controllable by an analyst. Therefore, we are interested in finding necessary and sufficient conditions on the stochastic choice function that ensure that the rationalizing limiting MSC model is reversible. We also characterize stochastic choice functions which are rationalizable by limiting MSC models in which all pairs of alternatives can be compared directly and in which the decision maker is able to explore the whole choice set.

4.1 Characterization of reversible limiting MSC models

As we have shown in Proposition 4, a stochastic choice function generated by a reversible limiting MSC model cannot be influenced by comparability restrictions in the absence of time pressure. We study in more detail in which cases these models emerge and whether they can be characterized in terms of testable properties of the choice function.

For now we focus only on reversible limiting MSC models for which all transition probabilities are positive. As we show below such limiting MSC models are characterized by two properties, which are well-known in the stochastic choice literature – positivity and independence of irrelevant alternatives (IIA). Positivity means that each alternative is chosen with positive probability from each menu, that is $p(i, M) > 0$ for all $i \in M$ and all $M \in \mathcal{M}$. The IIA assumption is satisfied whenever

$$p(i, \{i, j\})p(j, M) = p(j, \{i, j\})p(i, M), \quad \forall i, j \in M \text{ and } \forall M \in \mathcal{M},$$

meaning that enlarging the choice set does not have an effect on relative choice probabilities. We state the characterization result below.

Proposition 7. *A stochastic choice function* $\boldsymbol{p}$ *is rationalizable by a reversible limiting MSC model with* $q_{ij}(M) > 0$ *for all* $i, j \in M$ *and* $M \in \mathcal{M}$ *iff it satisfies positivity and IIA.*

Proof. See Appendix D.1. □

The above result together with Proposition 4 implies that if stochastic choice data satisfies the two properties – IIA and positivity – restricting additional pairwise comparisons does not change final choices provided the same alternatives can be reached from each starting point.

Reversible limiting MSC models with positive transition probabilities are an interesting class of models because the transitions between the pairs of alternatives happen as if they are determined only by the utility of the two alternatives, and not by their relative salience, as shown in the corollary below.

Corollary 1. *A limiting MSC model with strictly positive transitions is reversible iff there exists a function* $u : N \to \mathbb{R}_{++}$ *such that*

$$\frac{q_{ij}(M)}{q_{ji}(M)} = \frac{u(j)}{u(i)}. \tag{4.1}$$

Proof. See Appendix D.2. □

An example of such model is one in which $s_{ij}(M) = s_{ji}(M)$ and $t_{ij} = \frac{u(j)}{u(i)+u(j)}$ for all $i, j \in M$ and all $M \in \mathcal{M}$, where $u(.)$ can be interpreted as a utility function. Since relative transitions between all pairs of alternatives are guided only by tastes and not attention, the above result shows that we could think of decision makers following a reversible limiting MSC model as being rational, although their choices are stochastic. Moreover, a violation of the IIA and positivity axioms signifies that pairwise comparisons are affected by alternatives' salience.

Proposition 7 is interesting because it shows that the well-known Luce model, also referred to as multinomial logit, is nested in the limiting MSC model. A stochastic choice function $\boldsymbol{p}$ is a Luce rule if there exists a function $u : N \to \mathbb{R}_{++}$ such that for all $M \in \mathcal{M}$ 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, Proposition 7 can be seen as a procedural justification for the Luce model. Basically, stochastic choice consistent with the Luce model arises whenever the search dynamics evolves as a Markov process in which on average the transitions from i to j are equal to the transitions from j to i for all pairs of alternatives.

Proposition 7 is restricted to reversible limiting MSC models with positive transition probabilities. As we have seen in Proposition 4 the domain of limiting MSC models which are not susceptible to comparability restrictions is larger, and we will define in the following necessary and sufficient conditions for stochastic choice functions that are rationalizable only with reversible limiting MSC models.

First, we define a binary relation capturing violations of IIA on the set of alternatives for each menu.

Definition 4. *Let $\rhd_M$ be a binary relation over a menu $M \in \mathcal{M}$ such that for $i, j \in M$ and $i \neq j$*

$$i \rhd_M j \iff p(i, \{i, j\})p(j, M) > p(j, \{i, j\})p(i, M).$$

This means that $i \rhd_M j$ when enlarging the choice set to M decreases the relative choice probability of i against j. Therefore, the binary relation is antisymmetric, i.e. if $(i, j) \in$

$\rhd_M$, then $(j,i) \notin \rhd_M$. Note that $\rhd_M$ may not be complete, namely when the relative choice probabilities for some pairs are not affected by enlarging the choice set. A stochastic choice function satisfies IIA if and only if $\rhd_M = \varnothing$ for all $M \in \mathcal{M}$.

Furthermore, the binary relation $\rhd_M$ might not be transitive. We define a cycle of $\rhd_M$ on $M' \subseteq M$ as a set of ordered pairs $\mathcal{C}_M(M') = \{(i_1, i_2), (i_2, i_3), \ldots, (i_H, i_1)\}$ such that $\forall (j,k) \in \mathcal{C}_M(M')$ it holds that $j, k \in M'$ and $j \rhd_M k$, where some alternatives might be revisited such that $|M'| \leqslant H$. If there is no such cycle on any $M' \subseteq M$, we call $\rhd_M$ acyclical.

As shown in the theorem below we can determine based on the cyclicity property of $\rhd_M$ whether the stochastic choice function $\boldsymbol{p}(M)$ is rationalizable with a limiting MSC model which is reversible or non-reversible on M.

Theorem 1. *A stochastic choice function $\boldsymbol{p}(M)$ is rationalizable only by limiting MSC models which are reversible on M iff $\rhd_M$ is acyclical. A stochastic choice function $\boldsymbol{p}(M)$ is rationalizable by a limiting MSC model which is non-reversible on M iff $\rhd_M$ is cyclical.*

Proof. See Appendix D.3. □

According to the result above the reversibility of a limiting MSC model on a particular menu M can be determined solely by using choice data on the particular menu M and choice data on all binary choice sets between the alternatives in M. It also implies that a stochastic choice function $\boldsymbol{p}$ is rationalizable only by reversible limiting MSC models iff $\rhd_M$ is acyclical *on all* $M \in \mathcal{M}$.

Theorem 1 offers an alternative characterization of reversible limiting MSC models to the one obtained in Proposition 5. The result is particularly useful, because it does not rely on the manipulability of the decision problem by comparability restrictions, which is associated with unobservable cognitive limitations of the decision-maker. Theorem 1 and Propositions 4 and 5 can even serve as complements to make predictions and to uncover the properties of the true generating model. The above theorem together with Proposition 4 implies that if a decision maker is not under time pressure and the observed stochastic choice function satisfies acyclicity on a particular menu, then manipulating the comparability restrictions for that menu does not affect the stochastic choice. Moreover, in the case when $\triangleright_M$ is cyclical on a particular menu, Theorem 1 implies that the stochastic choice can be rationalized with a limiting MSC model which is non-reversible on M, but we cannot be certain that the true generating model is indeed non-reversible on M. However, if it is feasible to manipulate the decision problem by comparability restrictions, we can apply Proposition 5 to verify that. An analyst would need to elicit the choice function for the manipulated choice problem and observe whether this alters choices. If there exists at least one such manipulation the true generating model is indeed non-reversible.

4.2 Characterization of ergodic limiting MSC models

In this section we drop the IIA assumption of Proposition 7 and characterize limiting MSC models in which (i) the agent can directly compare all pairs of alternatives and (ii) the agent is able to explore the whole choice set.

We are in particular interested in the case when stochastic choice is rationalizable by a limiting MSC model with positive transition probabilities because it means that the agent's ability to make pairwise comparisons is not restricted by cognitive limitations similar to the ones we discussed in Section 3.3.

First, we need to define a property of the stochastic choice for each pair of alternatives in a menu.

Definition 5. *The choice over a pair $i, j \in M$ is bounded in a cycle according to $\boldsymbol{p}(M)$ if whenever $(i, j) \notin \triangleright_M$ and $(j, i) \notin \triangleright_M$, there exists a $\mathcal{C}_M(M')$ with $(i, j) \in \mathcal{C}_M(M')$ for some $M' \subseteq M$.*

The property implies that in larger choice sets, the relative choice probabilities are bounded from above and below by the relative binary choice probabilities. In particular, it follows from the definition of a cycle that it holds for all pairs $(i_h, i_{h+1}) \in \mathcal{C}_M(M')$

that

$$\frac{p(i_h,\{i_h,i_{h+1}\})}{p(i_{h+1},\{i_h,i_{h+1}\})} > \frac{p(i_h,M)}{p(i_{h+1},M)},$$

$$\frac{p(i_h,M)}{p(i_{h+1},M)} > \frac{p(i_h,\{i_{h-1},i_h\})}{p(i_{h-1},\{i_{h-1},i_h\})}\cdots\frac{p(i_2,\{i_1,i_2\})}{p(i_1,\{i_1,i_2\})}\frac{p(i_1,\{i_1,i_H\})}{p(i_H,\{i_1,i_H\})}\cdots\frac{p(i_{h+1},\{i_h,i_{h+1}\})}{p(i_h,\{i_h,i_{h+1}\})},$$

where $M' = \{i_1,\ldots,i_H\} \subseteq M$.

By requiring the boundedness condition to hold on all pairs of alternatives in the menu, we can show that the choice function is consistent with that of an agent who can transition directly between all pairs of alternatives in a menu.

Theorem 2. *A stochastic choice function* $\boldsymbol{p}$ *is rationalizable by a limiting MSC model such that*

(i) $\nexists M \in \mathcal{M}$ *and* $i,j \in M$ *for which* $q_{ij}(M) = q_{ji}(M) = 0$ *iff choice over all pairs and all menus is bounded in a cycle;*

(ii) $q_{ij}(M) > 0$ *for all* $i,j \in M$ *and* $M \in \mathcal{M}$ *iff choice over all pairs and all menus is bounded in a cycle and* $p(i,\{i,j\}) \in (0,1)$ *for all* $i,j \in N$.

Proof. See Appendix E.1. □

Theorem 2 implies that whenever a choice function satisfies the bounded in a cycle condition for pairs of alternatives and all menus, it can always be rationalized by a limiting MSC model for which $s_{ij}(M) > 0$ for all $i,j \in M$. On the other hand, choices that violate the condition indicate that the decision-maker's ability to compare certain pairs of alternatives directly has been restricted, due to the presentation of the decision problem or the cognitive limitations.

Note that the classes of models characterized in Theorems 1 and 2 do not nest each other, because a limiting MSC model might be reversible, but some transition probabilities might be equal to zero. In fact, the intersection of the two classes of limiting MSC models is the one we analyzed in Section 4.1. Therefore, if a stochastic choice function is positive and acyclical on all M, but violates IIA, then it is only rationalizable by reversible limiting MSC models, for which some transitions happen with zero probability.

Lastly, we show that choice data alone is informative about an agent's ability to explore the whole menu during the decision-making process. This is the case when the rationalizing limiting MSC model is ergodic, hence it is irreducible and aperiodic on all menus. Since this is a more general result than Theorem 2, only a subset of pairs of alternatives need to be bounded in a cycle.

Theorem 3. *A stochastic choice function* $\boldsymbol{p}$ *is rationalizable by an ergodic limiting MSC model iff* $\forall M \in \mathcal{M}$ *the alternatives in the menu can be arranged in a sequence such that choice over all consecutive pairs* $i, j \in M$ *is bounded in a cycle and* $p(i, \{i, j\}) \in (0, 1)$.

Proof. See Appendix E.2. □

Note that the ergodicity of the underlying model implies that the observed choice behavior does not depend on the initial distribution. On the other hand, if the alternatives in the menu cannot be ordered such that all consecutive pairs satisfy the boundedness condition, Theorem 3 implies that the initial fixation determines the set of alternatives that the agent can explore, which is only a subset of the whole menu. Therefore, changes in the initial distribution would affect the purchasing decisions of the agent.

The limited exploration of the choice set is related to the existence of consideration sets. One such case is when the agent approaches the choice problem by first selecting a subcategory of alternatives in the menu to view and then choosing among them. For example, if the decision maker wishes to buy a snack, he might first decide whether he wishes to buy a sweet or salty snack and explore only the items in the respective category. Such behavior is consistent with a large body of experimental literature in marketing that identifies different heuristics decision makers use to construct consideration sets, which often involves considering a subset of items that possess a certain attribute (see Hauser (2014) and references therein).

In the context of the shopper from an e-commerce website with a recommendation system, violations of the condition on the stochastic choice function indicate that the recommendation algorithm has identified distinct subgroups among the available alternatives and would suggest different sets of related alternatives depending on the first alternative that the agent views.

5 Revealed attention and acceptance probability in the limiting MSC model

In this section we analyze whether the decision-maker's choices reveal information about the unobservable parameters of the limiting MSC model. Inferring the exact transition probabilities from choice data is in general not possible, since there are multiple models that generate the same choice behavior. However, as we show in the following proposition, choice data alone allows us to determine the pairs of alternatives which the decision maker is not able to compare directly.

Proposition 8. *Let $\boldsymbol{p}$ be a positive stochastic choice function rationalizable by a limiting MSC model. If a pair $i, j \in M$ is not bounded in a cycle then $q_{ij}(M) = q_{ji}(M) = 0$.*

Proof. See Appendix F.1. □

Proposition 8 shows that choice data reveals the precise pairs of alternatives which the decision maker is not able to compare directly. This information can be used by a planner to make better targeted interventions using the instruments discussed in Section 3. Moreover, if choice data reveals the existence of consideration sets when applying Theorem 3, sellers can infer which products fall into the same consideration set as their own product, thus showing their direct competitors as seen from consumers' perspective.[7]

In order to infer the positive transition probabilities of the underlying limiting MSC model, an analyst needs to have access to additional data featuring the sequence of eye

[7] Note also that the choice behavior of the agent reveals when the only rationalizing model is one where no transitions between alternatives are allowed, namely $q_{ij}(M) = 0$ for all $i, j \in M$. According to such a model the agent chooses the alternative that he first views. This is the only rationalizing limiting MSC model if $\rhd_M$ is complete and acyclical.

fixations. However, using choice data only, we can still learn valuable information about the acceptance probabilities and relative salience.

We know from Example 1 that the ratio of the transition probabilities is revealed from choice data on binary menus. Imposing more structure on the model by strengthening the TR-IIA condition allows us to reveal additional information about the underlying parameters. Therefore, we focus now on limiting MSC models for which the relative salience parameters are symmetric, therefore $s_{ij}(M) = s_{ji}(M)$ for all $i, j \in M$ and all $M \in N$. This property of a limiting MSC model is intuitive if we consider the relative salience to be determined by the distance between the respective pairs. As shown in the corollary of Theorem 1 below, among the rationalizing limiting MSC models of a stochastic choice function, there always exists one with symmetric relative salience.

Corollary 2. *All stochastic choice functions are rationalizable by limiting MSC models with $s_{ij}(M) = s_{ji}(M)$ for all $i, j \in M$ and all $M \in N$.*

Proof. See Appendix D.4. □

This class of limiting MSC models is particularly interesting, because it allows us to infer the exact conditional acceptance probabilities t_{ij} from choice data on binary sets.

Proposition 9. *Let $\boldsymbol{p}$ be a stochastic choice function rationalizable by a limiting MSC model $Q(M)$ with $s_{ij}(M) = s_{ji}(M)$ for all $i, j \in M$. All acceptance probabilities t_{ji} can be uniquely identified from data on binary choice sets, namely $t_{ji} = p(i, \{i, j\})$ for all $i, j \in N$.*

Proof. Fix $M = \{i, j\}$. We substitute $q_{ij}(M) = s_{ij}(M)t_{ij}$ in the expression for the binary choice probabilities from Example 1:

$$p(i, \{i, j\}) = \frac{q_{ji}(\{i, j\})}{q_{ij}(\{i, j\}) + q_{ji}(\{i, j\})} = \frac{s_{ji}(\{i, j\})t_{ji}}{s_{ij}(\{i, j\})t_{ij} + s_{ji}(\{i, j\})t_{ji}} = \frac{t_{ji}}{t_{ij} + t_{ji}} = t_{ji},$$

which follows from the symmetric salience assumption and that $t_{ij} + t_{ji} = 1$. □

Proposition 9 is important to understand the driving force behind a high probability to fixate on and subsequently choose a particular alternative. The result can be used to isolate the effect of the value of an alternative from its relative salience, captured in the acceptance probabilities, on its choice probability.

As we noted previously, Proposition 8 can be used to identify whether certain competitors belong to the same consideration set. Using choice data on pairs and triples we can infer information about the probability to attract attention away from an incumbent relative to another competitor in the consideration set.

Proposition 10. *Let $M = \{i, j, k\}$ and $\boldsymbol{p}$ be a stochastic choice function rationalizable by a limiting MSC model $Q(M)$ with $s_{ij}(M) = s_{ji}(M)$ for all $i, j \in M$. Alternative j attracts more attention than k from incumbent i when*

$$-\frac{\frac{p(k,\{i,k\})}{p(i,\{i,k\})} - \frac{p(k,M)}{p(i,M)}}{\frac{p(j,\{i,j\})}{p(i,\{i,j\})} - \frac{p(j,M)}{p(i,M)}} > \frac{p(i,\{i,j\})}{p(i,\{i,k\})}.$$

Proof. See Appendix F.2. □

This condition means roughly that j attracts more attention than k if adding j to the menu $\{i, k\}$ causes a greater change in the relative choice probabilities than adding k to the pair $\{i, j\}$. Such information is valuable to firms wishing to assess the effectiveness of a product's marketing or presentation against its competitors.

6 Related literature

Recent contributions from the neuroeconomics literature by Cerreia-Vioglio et al. (2018) and Baldassi, Cerreia-Vioglio, Maccheroni, Marinacci, and Pirazzini (2020) are most related to this paper. Their proposed model is called the Metropolis DDM model which can be considered a special case of the MSC model. Notable similarities between the Metropolis DDM model and the MSC model are Markovian exploration of the choice set through stochastic pairwise comparisons and the way consideration sets are incorporated. However, in Cerreia-Vioglio et al. (2018) the exploration mechanism and the transition probabilities are defined more specifically. The decision-making procedure is assumed to follow the classical Metropolis algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953), which in our setting means that the relative salience parameters are symmetric, and that the transition probabilities are determined by the drift-diffusion model (DDM) (Ratcliff, 1978). Another difference between the two models is the way time pressure is modeled – in Cerreia-Vioglio et al. (2018) and Baldassi et al. (2020) the process stops after a given number of transitions, whereas in our setting the stopping time is probabilistic.

In recent years there has been considerable interest in the economics literature in incorporating different behavioral limitations into models of decision-making. Our paper contributes mainly to three different branches - search, reference dependence, and limited consideration. This literature is vast, so we will discuss only the most relevant papers below.

There are a number of related publications assuming a dynamic exploration of the choice set. In many of these works the search order itself is part of the decision problem (Horan, 2010; Papi, 2012; Zhang, 2016), which is not the case in the MSC model. Caplin and Dean (2011) consider a model of stochastic exploration of the choice set in which utility maximization is performed on the explored choice set. The authors use non-standard data comprised of agent's provisional choices that evolve with contemplation time to characterize the model, whereas our analysis is based only on choice data.

There are a number of papers featuring unobservable search order and deterministic preferences. Apesteguia and Ballester (2013) consider a sequential procedure guided by routes according to which agents make deterministic pairwise comparisons. As in our model, the winner of a binary comparison is compared pairwise with another alternative in the menu. Since the procedure continues until the menu is exhausted, the choice function is deterministic as opposed to the stochastic choice that we study.

Two related papers study decision-making in the context of unobservable networks of products and the impact of comparability restrictions. Masatlioglu and Nakajima (2013) and Masatlioglu and Suleymanov (2019) assume that a decision-maker is possibly unaware of the available products and searches the choice set using the recommendation system of an online platform, which also applies to the MSC model. The decision maker has a deterministic preference and chooses the best alternative from the recommended ones unless all suggested alternatives are worse than the current one, making the termination of the process endogenous.

All of the mentioned papers on search so far feature deterministic choice. A recent paper by Dutta (2020) models choices as stochastic and shares some similarities with the MSC model. Notably, both models feature pairwise comparisons and random stopping probability. However, in Dutta (2020) the comparisons and the search order are deterministic.

The MSC model can be interpreted as featuring a dynamic reference point, because the probability to transition from an incumbent to a competitor is incumbent-specific. Traditionally, reference points have been assumed to be static and exogenous (Tversky & Kahneman, 1991; Masatlioglu & Ok, 2005; Salant & Rubinstein, 2008). Ok, Ortoleva, and Riella (2015) and Tserenjigmid (2019) model reference points as endogenous depending on the choice set, but not as dynamic.

Notably, Ravid and Steverson (2019) propose the "focus, then compare" decision procedure, which features a dynamic reference dependence and stochastic binary comparisons similarly to the MSC model. They assume that the decision maker selects a focal alternative at random and compares it stochastically to all other available alternatives. If the focal option wins all pairwise comparisons, then the process stops, and if not, the procedure is repeated by selecting another focal alternative. An important feature that the MSC model shares with the model is that the probabilities to transition from an incumbent to a competitor are stable during the process and independent from past focal options. However, the stopping mechanism in Ravid and Steverson (2019) is endogenous, whereas it is exogenous in our model. Another difference is that the "focus, then compare" procedure implies that the agent is aware of all alternatives in the choice set and is able to make all pairwise comparisons.

Our paper adds to the literature on limited consideration since it is possible that the decision maker cannot explore the whole choice set. Many of the papers discussed so far also accommodate the existence of consideration sets. Two other influential papers on stochastic choice featuring consideration sets are Manzini and Mariotti (2014) and Brady and Rehbeck (2016). In these models stochastic choice is driven by the random composition of the consideration sets, unlike in the MSC model where if consideration sets exist, they are fixed for each choice set.

Finally, our paper is also related to the literature on manipulating attention and its consequences on choice behavior. Notable recent theoretical works addressing this issue are Gossner, Steiner, and Steward (2020), which features a learning mechanism and endogenous stopping, and Kovach and Tserenjigmid (2019), which enriches the Luce (1959) model by separating the alternatives into two categories depending on their ability to attract attention.

7 Conclusion

In this paper we propose a model of decision-making, which is inspired by the observation that people explore choice sets sequentially through stochastic pairwise comparisons. The model is based on the assumption that the decision-making process evolves as a Markov chain, which is terminated randomly. We use the framework to understand how and why certain interventions work and to study the properties of the decision process that can be inferred from stochastic choices, which in turn are informative for better-targeted interventions. In particular, we find that the choice probability of an alternative increases in the presence of an alternative which it dominates and when there is a higher chance that the first fixation is on that alternative.

In addition, we analyze the effect of the arrangement of alternatives and the induced comparability restrictions on choices with and without time pressure. Thereby, we identify that the class of reversible limiting MSC models is resistant to such interventions. We obtain two characterizations of this domain for different types of choice data and show that it nests the well-known Luce model.

Further, we characterize two interesting classes of limiting MSC models – one which ensures that the agent is able to compare all alternatives directly and another one according to which the agent is able to explore the complete choice set irrespective of the starting position. Finally, we show that choice data can be used to infer the exact pairs of alternatives that the agent cannot compare, thus revealing the unobservable consideration sets. Moreover, choice data reveals information on the transition probabilities, which can be used to isolate the effect of value from saliency leading to the choice of a particular alternative.

The proposed model offers different directions for future research, two of which we outline here. In the current paper, 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 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)).

References

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). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, *3*(5), 396–403.

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.

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.

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. (2018). *Multinomial logit processes and preference discovery: inside and outside the black box.*

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.

Dutta, R. (2020). *Gradual pairwise comparison and stochastic choice.*

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. *Journal of Marketing Research*, *51*(4), 487–507.

Gossner, O., Steiner, J., & Steward, C. (2020). *Attention please!*

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.*

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.

Kovach, M., & Tserenjigmid, G. (2019). *The focal Luce model.*

Kress, R. (2014). *Linear integral equations* (3rd ed., Vol. 74) (No. 470). New York: Springer-Verlag.

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., & Ok, E. A. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, *121*(1), 1–29.

Masatlioglu, Y., & Suleymanov, E. (2019). Decision making within a product network. *Economic Theory*.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, *21*, 1087–1092.

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.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Ravid, D., & Steverson, K. (2019). *Focus, then compare.*

Raymond, C. (2011). *Revealed search theory.*

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol Bull.*, *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. (2015). *Advanced linear algebra* (2nd ed.; S. Axler, F. W. Gehring, & K. A. Ribet, Eds.). New York: Springer.

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.

Shugan, S. M. (1980). The cost of thinking. *Journal of Consumer Research*, *7*(2), 99–111.

Svenson, O., & Maule, J. A. (1993). *Time pressure and stress in human judgment and decision making.* New York: Plenum Press.

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.*

Yang, S., & Lynn, M. (2014). More evidence challenging the robustness and usefulness of the attraction effect. *Journal of Marketing Research*, *51*(4), 508–513.

Zhang, J. (2016). *Stochastic choice with categorization.*

A Appendix: Proof of Proposition 1

Fix a menu $M \in \mathcal{M}$. We rearrange equation (2.2) in the following way:

$$\begin{aligned}\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) &= \frac{\alpha}{1-\alpha}\boldsymbol{\pi}(M)\left(\frac{1}{1-\alpha}I - Q(M)\right)^{-1} \\ &= \frac{\alpha}{1-\alpha}\boldsymbol{\pi}(M)\frac{1}{\det\left(\frac{1}{1-\alpha}I - Q(M)\right)}\operatorname{adj}\left(\frac{1}{1-\alpha}I - Q(M)\right),\end{aligned}$$

where adj(.) 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 as shown below:

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

where λ_l denote 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

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

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

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

Let us multiply both sides of the equation with $(I - Q(M))$ from the right:

$$\begin{aligned}\boldsymbol{\rho}(Q(M))(I - Q(M)) &= \boldsymbol{\pi}(M)\frac{1}{\prod_{l=2}^{L}(1-\lambda_l)}\operatorname{adj}(I - Q(M))(I - Q(M)) \\ &= \boldsymbol{\pi}(M)\frac{1}{\prod_{l=2}^{L}(1-\lambda_l)}\det(I - Q(M))\end{aligned}$$

Since the matrix $(I - Q(M))$ is singular, $\det(I - Q(M)) = 0$. Therefore,

$$\boldsymbol{\rho}(Q(M))(I - Q(M)) = \mathbf{0}$$

Hence, $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M))$ converges to the unique stationary distribution of the Markov chain with transition probability matrix $Q(M)$ as $\alpha \to 0$.

B Appendix: Proofs of the results on choice architecture

B.1 Proof of Proposition 2

The choice probability of i in the menus $\{i, j\}$ and $\{i, j, k\}$ equals:

$$\rho(i, Q(\{i,j\})) = \frac{q_{ji}(\{i,j\})}{q_{ji}(\{i,j\}) + q_{ij}(\{i,j\})},$$
$$\rho(i, Q(M)) = \frac{q_{jk}(M)q_{ki}(M) + q_{ji}(M)(q_{ki}(M) + q_{kj}(M))}{q_{jk}(M)q_{ki}(M) + q_{ji}(M)(q_{ki}(M) + q_{kj}(M)) + q_{ij}(M)(q_{jk}(M) + q_{ki}(M) + q_{kj}(M))},$$

where $M = \{i, j, k\}$. We plug the expressions in the inequality $\rho(i, Q(\{i, j, k\})) > \rho(i, Q(\{i, j\}))$ and rearrange it in the following way

$$q_{ij}(\{i,j\})(q_{jk}(M)q_{ki}(M)+q_{ji}(M)(q_{ki}(M)+q_{kj}(M))) > q_{ji}(\{i,j\})q_{ij}(M)(q_{jk}(M)+q_{ki}(M)+q_{kj}(M)).$$

We can simplify the inequality using that $q_{ji}(\{i, j\})q_{ij}(M) = q_{ij}(\{i, j\})q_{ji}(M) > 0$, which follows from TR-IIA, and obtain

$$q_{ij}(\{i,j\})q_{jk}(M)q_{ki}(M) > q_{ji}(\{i,j\})q_{ij}(M)q_{jk}(M).$$

We simplify the above equation, divide by $q_{ij}(\{i, j\})$ and use that $q_{ji}(M) = \frac{q_{ji}(\{i,j\})q_{ij}(M)}{q_{ij}(\{i,j\})}$ to get $q_{ki}(M) > q_{ji}(M)$. The reverse statement can be obtained in an analogous way by assuming a limiting MSC model with $q_{ki}(M) > q_{ji}(M)$ and showing that the choice probability of i is larger when the dominated alternative k is in the choice set.

B.2 Proof of Proposition 3

Fix a menu $M \in \mathcal{M}$ and let $\boldsymbol{\rho}(\alpha, \delta_i, Q(M))$ denote the stochastic choice function generated by the baseline MSC model with initial distribution equal to the degenerate distribution with point mass on alternative i. We will first show that a stochastic choice function is a linear combination of the stochastic choice functions from degenerate initial conditions. Let $V = (I - (1 - \alpha)Q(M))$ and $W = V^{-1}$, which exists for $\alpha \neq 0$. Recall from equation (2.2) that

$$\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha\boldsymbol{\pi}(M)(I - (1-\alpha)Q(M))^{-1} = \alpha\boldsymbol{\pi}(M)V^{-1} = \alpha\boldsymbol{\pi}(M)W.$$

Since we have fixed α and $Q(M)$ we will use the following simplified notation to denote the probability to choose alternative i from the menu $\rho(i, \alpha, \boldsymbol{\pi}(M), Q(M)) = \rho_{\alpha,Q}(i, \boldsymbol{\pi})$.

Note that $\rho_{\alpha,Q}(i, \delta_j) = \alpha w_{ji}$ for any $i, j \in M$. If $\boldsymbol{\pi}(M)$ is not degenerate, we have that

$$\rho_{\alpha,Q}(i, \boldsymbol{\pi}) = \sum_{j \in M} \pi(j, M)\alpha w_{ji} = \sum_{j \in M} \pi(j, M)\rho_{\alpha,Q}(i, \delta_j) \quad \forall i \in M. \tag{B.1}$$

Let i be the target alternative. We will now show that $\rho_{\alpha,Q}(i, \delta_i) > \rho_{\alpha,Q}(i, \delta_j)$ for all $j \neq i$. Denote $\max_{j \neq i}\{\rho_{\alpha,Q}(i, \delta_j)\} = \rho_{\alpha,Q}(i, \delta_m)$. Since $V \cdot W = I$ we obtain the following equation when we multiply V's m^{th} row vector with W's i^{th} column vector:

$$(1 - (1-\alpha)q_{mm}(M))w_{mi} + \sum_{j \neq m}(-1)(1-\alpha)q_{mj}(M)w_{ji} = 0$$

$$(1 - (1-\alpha)(1 - \sum_{j \neq m} q_{mj}(M)))w_{mi} - (1-\alpha)\sum_{j \neq m} q_{mj}(M)w_{ji} = 0$$

$$(\alpha + (1-\alpha)\sum_{j \neq m} q_{mj}(M))w_{mi} - (1-\alpha)\sum_{j \neq m} q_{mj}(M)w_{ji} = 0$$

$$(1-\alpha)(\sum_{j \neq m} q_{mj}(M)(w_{ji} - w_{mi})) = \alpha w_{mi}$$

Finally, we plug in $\rho_{\alpha,Q}(i, \delta_j) = \alpha w_{ji}$ and obtain

$$\sum_{j \neq m} q_{mj}(M)(\rho_{\alpha,Q}(i, \delta_j) - \rho_{\alpha,Q}(i, \delta_m)) = \frac{\alpha}{1-\alpha}\rho_{\alpha,Q}(i, \delta_m).$$

We assumed w.l.o.g. that $\rho_{\alpha,Q}(i, \delta_m) \geqslant \rho_{\alpha,Q}(i, \delta_j)$ for all $j \neq i$. If $\rho_{\alpha,Q}(i, \delta_m) \geqslant \rho_{\alpha,Q}(i, \delta_i)$, then the left-hand side of the above equation is weakly negative, which is a contradiction. Therefore, $\rho_{\alpha,Q}(i, \delta_i) > \rho_{\alpha,Q}(i, \delta_j)$ for all $j \neq i$. Consider now the difference in the choice probability of i using equation (B.1) and that $\sum_j \Delta\pi(j, M) = 0$:

$$\begin{aligned}
\rho_{\alpha,Q}(i, \boldsymbol{\pi}') - \rho_{\alpha,Q}(i, \boldsymbol{\pi}) &= \Delta\pi(i, M)\rho_{\alpha,Q}(i, \delta_i) + \sum_{j \neq i} \Delta\pi(j, M)\rho_{\alpha,Q}(i, \delta_j) \\
&= -\sum_{j \neq i} \Delta\pi(j, M)\rho_{\alpha,Q}(i, \delta_i) + \sum_{j \neq i} \Delta\pi(j, M)\rho_{\alpha,Q}(i, \delta_j) \\
&= \sum_{j \neq i} \Delta\pi(j, M)(\rho_{\alpha,Q}(i, \delta_j) - \rho_{\alpha,Q}(i, \delta_i)) > 0,
\end{aligned}$$

where the last inequality follows from the assumption that $\Delta\pi(j, M) \leqslant 0$ for all $j \neq i$.

B.3 Proof of Proposition 4

Fix a menu $M \in \mathcal{M}$ and an ergodic limiting MSC model $Q(M)$, which is reversible on M. The limiting MSC model $Q^{'}(M)$ is manipulated by comparability restrictions such that for one pair $i, j \in M$ holds $q_{ij}(M) \neq q'_{ij}(M) = q'_{ji}(M) = 0$. We denote their respective stochastic choice functions by $\boldsymbol{\rho}(Q(M))$ and $\boldsymbol{\rho}(Q^{'}(M))$.

We first show that the manipulated limiting MSC model is reversible on M and ergodic. Since the non-manipulated model $Q(M)$ is reversible, it satisfies Kolmogorov's criterion on all sequences of states, i.e., for all $i_1, i_2, \dots, i_I \in M$,

$$q_{i_1 i_2}(M) q_{i_2 i_3}(M) \dots q_{i_{I-1} i_I}(M) q_{i_I i_1}(M) = q_{i_2 i_1}(M) q_{i_3 i_2}(M) \dots q_{i_I i_{I-1}}(M) q_{i_1 i_I}(M).$$

Since a valid manipulation requires that the set of reachable alternatives is preserved, $Q^{'}(M)$ has to be ergodic.

Consider all sequences of alternatives in M that do not involve i and j. TR-IIA and the consistency requirement on non-manipulated pairs ensures that the model $Q^{'}(M)$ satisfies Kolmogorov's criterion on all such sequences of alternatives. Since $q'_{ij}(M) = q'_{ji}(M) = 0$ all remaining sequences satisfy Kolmogorov's criterion trivially and the manipulated limiting MSC model is also reversible on M. Since both models satisfy the detailed balance condition we have that

$$\frac{\rho(l, Q(M))}{\rho(k, Q(M))} = \frac{q_{kl}(M)}{q_{lk}(M)} = \frac{q'_{kl}(M)}{q'_{lk}(M)} = \frac{\rho(l, Q^{'}(M))}{\rho(k, Q^{'}(M))} \qquad \forall k, l \in M \backslash \{i, j\} \text{ and } q_{lk}(M) > 0.$$

Next, we use the fact that rationalizable stochastic choice functions satisfy by definition the equality $\boldsymbol{\rho}(Q(M))Q(M) = \boldsymbol{\rho}(Q(M))$. The system can also be written as [8]

$$\sum_{k \neq l} \rho(k, Q(M)) q_{kl}(M) = \rho(l, Q(M)) \sum_{k \neq l} q_{lk}(M),$$
$$\sum_{k \neq l} \rho(k, Q^{'}(M)) q'_{kl}(M) = \rho(l, Q^{'}(M)) \sum_{k \neq l} q'_{lk}(M),$$

for all $l \in M$. Our goal is to show $\sum_{k \neq l} \rho(k, Q(M)) q'_{kl}(M) = \rho(l, Q(M)) \sum_{k \neq l} q'_{lk}(M)$ for all $l \in M$, which ensures that $\boldsymbol{\rho}(Q(M))$ is a stationary distribution of $Q^{'}(M)$. Set $l = i$ and

[8] More details can be found in the proof of Lemma 3 in Appendix C.

consider

$$\sum_{k\neq i}\rho(k,Q(M))q'_{ki}(M)-\rho(i,Q(M))\sum_{k\neq i}q'_{ik}(M)=$$
$$=\sum_{\substack{k\neq i\\ q'_{ki}\neq 0\\ q'_{ik}\neq 0}}\left(\rho(k,Q(M))q'_{ki}(M)-\rho(i,Q(M))q'_{ik}(M)\right)+\sum_{\substack{k\neq i\\ q'_{ki}\neq 0\\ q'_{ik}=0}}\rho(k,Q(M))q'_{ki}(M)$$
$$-\rho(i,Q(M))\sum_{\substack{k\neq i\\ q'_{ki}=0\\ q'_{ik}\neq 0}}q'_{ik}(M)+\sum_{\substack{k\neq i\\ q'_{ki}=0\\ q'_{ik}=0}}\left(\rho(k,Q(M))q'_{ki}(M)-\rho(i,Q(M))q'_{ik}(M)\right).$$

Consider the case in which $q'_{ki}(M)\neq 0$ and $q'_{ik}(M)=0$. The consistency assumption about the non-manipulated pairs implies that $q_{ki}(M)\neq 0$ and $q_{ik}(M)=0$. Detailed balance ensures that $\rho(k,Q(M))=0$. This means that the second and third terms in the above expression are equal to zero. The fourth term is equal to zero trivially. We also use the assumption $\frac{q'_{ki}(M)}{q'_{ik}(M)}=\frac{q_{ki}(M)}{q_{ik}(M)}$ to modify the first term. Therefore, the resulting equation after the simplification is the following:

$$\sum_{k\neq i}\rho(k,Q(M))q'_{ki}(M)-\rho(i,Q(M))\sum_{k\neq i}q'_{ik}(M)=$$
$$=\sum_{\substack{k\neq i\\ q_{ki}\neq 0\\ q_{ik}\neq 0}}\frac{q'_{ki}(M)}{q_{ki}(M)}\left(\rho(k,Q(M))q_{ki}(M)-\rho(i,Q(M))q_{ik}(M)\right).$$

Since the limiting MSC model $Q(M)$ is reversible on M, detailed balance needs to holds on all pairs. Therefore, $\rho(k,Q(M))q_{ki}(M)-\rho(i,Q(M))q_{ik}(M)=0$ for all $k\neq i$ and we have shown that the whole expression is equal to zero.

We can repeat the same argument for all other $l\in M$. Thus, $\boldsymbol{\rho}(Q(M))$ is also the stationary distribution of $Q^{'}(M)$. Since the manipulated limiting MSC model is irreducible, the stationary distribution of the Markov chain $Q^{'}(M)$ is unique and therefore $\boldsymbol{\rho}(Q(M))=\boldsymbol{\rho}(Q^{'}(M))$.

We have now shown that restricting the comparability between a pair of alternatives does not affect the stationary distribution of the new Markov chain. If the manipulation affects more than one pair of alternatives, the argument of this proof can be applied again by taking $Q^{'}(M)$ as the non-manipulated matrix and constructing a new manipulated matrix $Q^{''}(M)$ which does not permit the transition between a different pair of alternatives. This can be done until the actual manipulated limiting MSC model is obtained.

Finally, the same argument extends trivially to the case in which the initial Markov chain is not irreducible. Since the initial distribution of the manipulated and non-manipulated limiting MSC model is the same, and also because after the manipulation the decision maker is able to reach the same alternatives from each starting point as before the manipulation, the generated choice distributions of the two models are identical.

B.4 Proof of Proposition 5

The necessity part of the argument follows directly from Proposition 4. In order to show the sufficiency part, fix a menu M and let $\boldsymbol{p}(M)$ be a stochastic choice function rationalizable by a limiting MSC model $Q(M)$ and all its manipulations by comparability restrictions $Q^{'}(M)$. We consider in particular the manipulations in which the comparison between a single pair is prohibited, let this be $i, j \in M$. It follows by the rationalizability of $\boldsymbol{p}$ that[9]

$$\begin{aligned}\sum_{k\neq l} p(k,M)q_{kl}(M) &= p(l,M)\sum_{k\neq l} q_{lk}(M),\\ \sum_{k\neq l} p(k,M)q'_{kl}(M) &= p(l,M)\sum_{k\neq l} q'_{lk}(M),\end{aligned} \tag{B.2}$$

for all $l \in M$. Setting $l = i$ and using that $\frac{q'_{ki}(M)}{q'_{ik}(M)} = \frac{q_{ki}(M)}{q_{ik}(M)}$ and $q'_{ki}(M) > 0 \implies q_{ki}(M) > 0$ for $k \neq i, j$, we can express the second inequality as follows:

$$\begin{aligned}&\sum_{k\neq i} p(k,M)q'_{ki}(M) - p(i,M)\sum_{k\neq i} q'_{ik}(M) =\\ &= \sum_{\substack{k\neq i\\ q'_{ki}\neq 0}} \left(p(k,M)q'_{ki}(M) - p(i,M)q'_{ik}(M)\right) - p(i,M) \sum_{\substack{k\neq i\\ q'_{ki}=0\\ q'_{ik}\neq 0}} q'_{ik}(M) = 0.\end{aligned}$$

[9] More details can be found in the proof of Lemma 3 in Appendix C.

We use the assumption that $q'_{ki}(M) = cq_{ki}(M)$ for all $k, i \in M$ for which comparability is not restricted and obtain

$$\sum_{\substack{k \neq i \\ q'_{ki} \neq 0}} (p(k,M)q_{ki}(M) - p(i,M)q_{ik}(M)) - p(i,M) \sum_{\substack{k \neq i \\ q'_{ki}=0 \\ q'_{ik} \neq 0}} q_{ik}(M) = 0.$$

Since only the transition between i and j is restricted, the above equation becomes

$$\sum_{\substack{k \neq i,j \\ q_{ki} \neq 0}} (p(k,M)q_{ki}(M) - p(i,M)q_{ik}(M)) - p(i,M) \sum_{\substack{k \neq i,j \\ q_{ki}=0 \\ q_{ik} \neq 0}} q_{ik}(M) = 0. \tag{B.3}$$

Consider now the first equation in (B.2) for $l = i$:

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

Together with (B.3) this implies that detailed balance needs to be satisfied for the manipulated pair i, j, hence

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

Therefore, we see that if the stochastic choice function is not susceptible to the comparability restriction of a particular pair, detailed balance is satisfied for the restricted pair. Since the stochastic choice function $\boldsymbol{p}(M)$ is not susceptible to any manipulation, this is only possible if it holds for the rationalizing limiting MSC model that detailed balance holds on all pairs, and hence that it is reversible on M.

B.5 Proof of Proposition 6

Let $M = \{i, j, k\}$ and consider the following baseline MSC models $\langle Q(M), \boldsymbol{\pi}(M), \alpha \rangle$ and $\langle Q'(M), \boldsymbol{\pi}(M), \alpha \rangle$ such that $q_{ij}(M) = q'_{ij}(M) > 0$ for all $i, j \in M$ except $q'_{ik}(M) =$

$q'_{ki}(M) = 0$. Recall that all stochastic choice functions generated by a baseline MSC model satisfy equation (2.2)

$$\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \alpha\boldsymbol{\pi}(M)(I - (1-\alpha)Q(M))^{-1}.$$

We multiply both sides of the equation with the matrix $(I - (1-\alpha)Q(M))$ and obtain the following:

$$\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M))(I-(1-\alpha)Q(M)) = \alpha\boldsymbol{\pi}(M) = \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q^{'}(M))(I-(1-\alpha)Q^{'}(M)),$$

since both models have the same stopping probability and initial distribution. We simplify the above equations in the following way

$$\begin{aligned} &\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) - \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q^{'}(M)) = \\ &= (1-\alpha)(\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M))Q(M) - \boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q^{'}(M))Q^{'}(M)). \end{aligned} \tag{B.4}$$

We solve the above system for $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M))$ and let $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q(M)) = \boldsymbol{p}$ and $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}(M), Q^{'}(M)) = \boldsymbol{p'}$ to simplify the notation. We obtain the following

$$\begin{aligned} p(i, M) - p'(i, M) &= -\frac{f_i(\alpha, Q)(q_{ik}(M)p'(i, M) - q_{ki}(M)p'(k, M))}{g(\alpha, Q)} \text{ and} \\ p(k, M) - p'(k, M) &= \frac{f_k(\alpha, Q)(q_{ik}(M)p'(i, M) - q_{ki}(M)p'(k, M))}{g(\alpha, Q)}, \end{aligned}$$

where

$$\begin{aligned} f_i(\alpha, Q) &= (1-\alpha)((1-\alpha)(q_{ji}(M) + q_{jk}(M) + q_{kj}(M)) + \alpha) > 0, \\ f_k(\alpha, Q) &= (1-\alpha)((1-\alpha)(q_{ji}(M) + q_{ij}(M) + q_{jk}(M)) + \alpha) > 0, \\ g(\alpha, Q) &= ((1-\alpha)(q_{ik}(M) + q_{ki}(M)) + \alpha)((1-\alpha)(q_{ji}(M) + q_{jk}(M)) + \alpha) + \\ &+ (1-\alpha)q_{kj}(M)((1-\alpha)(q_{ji}(M) + q_{ik}(M)) + \alpha) + \\ &+ (1-\alpha)q_{ij}(M)((1-\alpha)(q_{ki}(M) + q_{jk}(M) + q_{kj}(M)) + \alpha) > 0. \end{aligned}$$

Therefore, if $q_{ik}(M)p'(i, M) - q_{ki}(M)p'(k, M) < 0$ it holds that $p(i, M) - p'(i, M) > 0$ and $p(k, M) - p'(k, M) < 0$.

If we solve (B.4) for $\boldsymbol{\rho}(\alpha, \boldsymbol{\pi}, Q'(M))$ instead we obtain the following result:

$$p(i, M) - p'(i, M) = -\frac{f_i(\alpha, Q)(q_{ik}(M)p(i, M) - q_{ki}(M)p(k, M))}{g'(\alpha, Q)} \text{ and}$$
$$p(k, M) - p'(k, M) = \frac{f_k(\alpha, Q)(q_{ik}(M)p(i, M) - q_{ki}(M)p(k, M))}{g'(\alpha, Q)},$$

where

$$g'(\alpha, Q) = \alpha^2 + (1-\alpha)(\alpha(q_{jk}(M) + q_{kj}(M)) + q_{ji}(M)(\alpha + (1-\alpha)q_{kj}(M)) + \\ + q_{ij}(M)(\alpha + (1-\alpha)(q_{jk}(M) + q_{kj}(M)))) > 0.$$

We make a similar conclusion that if $q_{ik}(M)p(i, M) - q_{ki}(M)p(k, M) < 0$ it holds that $p(i, M) - p'(i, M) > 0$ and $p(k, M) - p'(k, M) < 0$.

C Appendix: General results on rationalizability

Let $d_{ji}(M) = p(j, M) - \frac{p(j,\{i,j\})}{p(i,\{i,j\})} p(i, M)$ for $p(i, \{i, j\}) \neq 0$. Therefore, we can rewrite Definition 4 as $i \rhd_M j \iff d_{ji}(M) > 0, \forall i, j \in M$. Here we show some auxiliary results which will be useful in the proofs of Theorems 1-3.

Lemma 2. *Let $\boldsymbol{p}$ be a stochastic choice function rationalizable by a limiting MSC model. For all $i, j \in M$ for which $p(i, \{i, j\}) \in (0, 1)$ it holds that $d_{ij}(M)q_{ij}(M) = -d_{ji}(M)q_{ji}(M)$.*

Proof. Since $p(i, \{i, j\}) \in (0, 1)$, $q_{ij}(\{i, j\}) > 0$ and $q_{ji}(\{i, j\}) > 0$. The statement is trivially satisfied for $q_{ij}(M) = q_{ji}(M) = 0$. For $q_{ij}(M) > 0$ and $q_{ji}(M) > 0$, TR-IIA implies that $q_{ij}(M) = \frac{p(j,\{i,j\})}{p(i,\{i,j\})} q_{ji}(M)$. Therefore,

$$\begin{aligned} d_{ij}(M)q_{ij}(M) &= \left(p(i, M) - \frac{p(i, \{i, j\})}{p(j, \{i, j\})} p(j, M)\right) \frac{p(j, \{i, j\})}{p(i, \{i, j\})} q_{ji}(M) \\ &= \left(\frac{p(j, \{i, j\})}{p(i, \{i, j\})} p(i, M) - p(j, M)\right) q_{ji}(M) \\ &= - d_{ji}(M)q_{ji}(M). \end{aligned}$$

□

The following lemma shows a necessary and sufficient condition for rationalizability.

Lemma 3. *Let $Q(M)$ be a limiting MSC model satisfying TR-IIA. A stochastic choice function $\boldsymbol{p}(M)$ is rationalizable iff for all $M \in \mathcal{M}$ it holds that*

$$\sum_{\substack{i \neq j \\ q_{ji} \neq 0}} d_{ji}(M) q_{ji}(M) - \sum_{\substack{i \neq j \\ q_{ji} = 0}} d_{ij}(M) q_{ij}(M) = 0 \quad \forall j \in M. \tag{C.1}$$

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

$$\sum_{\substack{i \neq j \\ q_{ij} \neq 0}} p(i, M) q_{ij}(M) + p(j, M) \left(1 - \sum_{\substack{i \neq j \\ q_{ji} \neq 0}} q_{ji}(M) \right) = p(j, M)$$

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

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

We apply the TR-IIA property and obtain the following equation:

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

We know from Example 1 that the ratio of transition probabilities for binary sets is equal to the choice probability ratio. We can rewrite the above equality using function $d()$ as shown below and obtain the final result:

$$\sum_{\substack{i \neq j \\ q_{ji} \neq 0 \\ q_{ij} \neq 0}} d_{ji}(M) q_{ji}(M) - \sum_{\substack{i \neq j \\ q_{ji} = 0 \\ q_{ij} \neq 0}} d_{ij}(M) q_{ij}(M) + \sum_{\substack{i \neq j \\ q_{ji} \neq 0 \\ q_{ij} = 0}} d_{ji}(M) q_{ji}(M) = 0$$

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

□

D Appendix: Proofs of the characterization results of reversible limiting MSC models

D.1 Proof of Proposition 7

Necessity: Let $Q(M)$ be a reversible limiting MSC model with strictly positive transition probabilities between all pairs of alternatives. Fix a menu $M \in \mathcal{M}$ and a pair $i, j \in M$. We will show that positivity and IIA hold for this pair and hence, for all other pairs and menus.

Since the transition probabilities between all pairs are strictly positive, the stationary distributions are such that $\rho(i, Q(M)) > 0$ for all $i \in M$ and all $M \in \mathcal{M}$. Hence, the induced stochastic choice function is positive. Reversibility implies that

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

Because of TR-IIA all fractions in the above equations are equal and hence IIA is satisfied.

Sufficiency: Let $\boldsymbol{p}$ be a positive stochastic choice function that satisfies IIA, that is

$$p(j, \{i,j\})p(i, M) = p(i, \{i,j\})p(j, M), \quad \forall i, j \in M, \forall M \in \mathcal{M}.$$

Note that $d_{ji}(M) = p(j, M) - \frac{p(j,\{i,j\})}{p(i,\{i,j\})} p(i, M) = 0$ for all $i, j \in M$. Lemma 3 in Appendix C implies that the stochastic choice function is rationalizable by a limiting MSC model with strictly positive transition probabilities. We will show that all such rationalizing models are reversible, hence

$$q_{ij}(M)p(i, M) = q_{ji}(M)p(j, M), \quad \forall i, j \in M, \forall M \in \mathcal{M}.$$

We apply IIA and Example 1 and 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, TR-IIA and the above equation imply that

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

Therefore, detailed balance holds on all pairs and menus and the rationalizing model is reversible for all menus.

D.2 Proof of Corollary 1

Necessity: If there exists a utility function such that the ratios of transition probabilities satisfy equation (4.1), reversibility of the limiting MSC model is trivial as it follows directly from Kolmogorov's criterion.

Sufficiency: Let $Q(M)$ be a reversible limiting MSC model with strictly positive transitions. In Proposition 7 we show that the model is characterized by IIA and positivity. Therefore, a stochastic choice function generated by the limiting MSC model $Q(M)$ is a Luce rule. Hence, there exists an increasing function $u : N \to \mathbb{R}_{++}$ such that for all $M \in \mathcal{M}$ and $i \in M$ such that

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

Since the model is reversible, it satisfies detailed balance on each pair of alternatives. Hence, it holds for all $i, j \in M$ and all $M \in \mathcal{M}$ that

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

and the result follows.

D.3 Proof of Theorem 1

Necessity: Fix a menu $M \in \mathcal{M}$. We consider a limiting MSC model $Q(M)$ with generated stochastic choice function $\boldsymbol{\rho}(Q(M)) = \boldsymbol{p}(M)$. We will show that if a limiting MSC model violates reversibility on M, the generated stochastic choice function $\boldsymbol{p}(M)$ is cyclical.

The proof is structured as follows. First, we show that a limiting MSC model, which is not reversible on M, generates stochastic choice functions which satisfy $\rhd_M \neq \varnothing$. Then, we represent the equations in (C.1) in matrix notation for each menu. If the Markov chain is not reversible on M, the system has a positive solution. 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 cyclicity property of $\rhd_M$.

If a limiting MSC model is not reversible on M, then detailed balance is violated for at least one pair $i, j \in M$, hence:

$$q_{ij}(M)p(i,M) \neq q_{ji}(M)p(j,M). \tag{D.1}$$

Assume by contradiction that $\rhd_M = \varnothing$. This means that it holds for i, j that

$$p(j,\{i,j\})p(i,M) = p(i,\{i,j\})p(j,M). \tag{D.2}$$

If $p(i,M) = p(j,M) = 0$, then inequality (D.1) is violated. If $p(i,M) > 0, p(j,M) = 0$, equation (D.2) implies that $p(i,\{i,j\}) = 0$ and hence $q_{ji}(\{i,j\}) = 0$. It follows from TR-IIA that $q_{ji}(M) = 0$, which would then violate (D.1). Finally, if $p(i,M) > 0, p(j,M) > 0$ implies together with (D.2) that $p(j,\{i,j\}) > 0$ and $p(i,\{i,j\}) > 0$. This in turn means that $q_{ji}(\{i,j\}) = 0$ and $q_{ij}(\{i,j\}) = 0$. TR-IIA is satisfied when either $q_{ij}(M) = q_{ji}(M) = 0$ or when both are strictly positive and

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

Both of these cases violate (D.1). Therefore, if a limiting MSC model is not reversible on M then $\rhd_M \neq \varnothing$.

Recall that $d_{ji}(M) = p(j,M) - \frac{p(j,\{i,j\})}{p(i,\{i,j\})}p(i,M)$ for $p(i,\{i,j\}) \neq 0$ as defined in Appendix C. Let $f_M : M \to [1,\ldots,|M|]$ be a bijective mapping assigning an index to each alternative in the menu M. Further, we denote by G_M a set of ordered pairs in the menu such that for all $i, j \in M$, $(i,j) \in G_M$ unless $\rho(j, Q(\{i,j\})) = 0$ or $(j,i) \in G_M$ or $d_{ij}(M) = 0$. Note that since there is some $i, j \in M$ for which $d_{ij}(M) \neq 0$, $G_M \neq \varnothing$. We denote with $g_M : G_M \to [1,\ldots,|G_M|]$ a bijective mapping on the set of ordered pairs. Further, we define a vector of transition probabilities $\boldsymbol{\gamma}(M)$ between the pairs in G_M, that is

$\gamma_{g_M(i,j)}(M) = q_{ij}(M)$. Finally, we denote by $\mathcal{D}(M)$ a matrix with dimensions $|M| \times |G_M|$ and elements

$$\mathscr{d}_{f_M(k),g_M(i,j)}(M) = \begin{cases} d_{ij}(M), & \text{if } k = i, \\ -d_{ij}(M), & \text{if } k = j, \\ 0 & \text{if } k \neq i, k \neq j, \end{cases}$$

where $k \in M$ and $(i,j) \in G_M$. Note that all elements of the matrix given by the functions $d_{ij}(M)$ are defined since if $\rho(j, Q(\{i,j\})) = 0$, $(i,j) \notin G_M$. Further, none of the columns of the matrix are equal to $\mathbf{0}$ because if $d_{ij}(M) = 0$, $(i,j) \notin G_M$. We claim that the system

$$\mathcal{D}(M)\boldsymbol{\gamma}(M) = \mathbf{0} \tag{D.3}$$

is equivalent to the one in (C.1) and is thus necessary and sufficient for rationalizability. In order to verify this claim, we consider an equation from the system in (D.3) corresponding to an arbitrary matrix row $f_M(j)$:

$$\sum_{\substack{i \in M \\ (j,i) \in G_M}} \mathscr{d}_{f_M(j),g_M(j,i)}(M)\gamma_{g_M(j,i)}(M) + \sum_{\substack{i \in M \\ (i,j) \in G_M}} \mathscr{d}_{f_M(j),g_M(i,j)}(M)\gamma_{g_M(i,j)}(M) = 0$$

$$\sum_{\substack{i \in M \\ (j,i) \in G_M}} d_{ji}(M)q_{ji}(M) - \sum_{\substack{i \in M \\ (i,j) \in G_M}} d_{ij}(M)q_{ij}(M) = 0$$

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

where we obtain the last equation by applying Lemma 2 and letting $d_{ij}(M)q_{ij}(M) = -d_{ji}(M)q_{ji}(M)$ for all $(i,j) \in G_M$ for which $q_{ji}(M) \neq 0$. For all such pairs for which $q_{ji}(M) > 0$, TR-IIA implies that $q_{ji}(\{i,j\}) > 0$ and hence, $p(i, \{i,j\}) > 0$, which ensures that $d_{ji}(M)$ is defined.

If the limiting MSC model is non-reversible on M, there must be at least one pair i, j for which $d_{ij}(M) \neq 0$ and $q_{ij}(M) > 0$ and/or $q_{ji}(M) > 0$, which means that $\boldsymbol{\gamma}(M) \geqslant \mathbf{0}$. We apply Gordan's theorem[10] which states that there exists a strictly positive solution

[10] See for example Theorem 15.1(2) in Roman (2015).

$\boldsymbol{\gamma}(M) \geqslant \mathbf{0}$ to the linear system in (D.3) if and only if there does not exist a vector $\boldsymbol{v} \in \mathbb{R}^{|M|}$ for which

$$\boldsymbol{v}\mathcal{D}(M) \gg \mathbf{0}. \tag{D.4}$$

Let us denote the elements of the vector $\boldsymbol{v} = (v_i, v_j, \dots)$ such that $\boldsymbol{v}_{f_M(i)} = v_i$. The linear system in (D.4) is equivalent to

$$v_i \boldsymbol{d}_{f(i),g_M(i,j)}(M) + v_j \boldsymbol{d}_{f(j),g_M(i,j)}(M) > 0, \quad \forall (i,j) \in G_M,$$

which is in turn equal to

$$(v_i - v_j) d_{ij}(M) > 0, \quad \forall (i,j) \in G_M. \tag{D.5}$$

Hence, each inequality in (D.5) specifies whether $v_i > v_j$ or $v_j > v_i$ depending on the sign of $d_{ij}(M)$ for all $(i,j) \in G_M$ (recall that for all $(i,j) \in G_M$, $d_{ij}(M) \neq 0$). As we argued above, if the Markov chain is non-reversible on M, the system in (D.3) has a strictly positive solution, and hence there does not exist a vector $\boldsymbol{v}$ satisfying (D.5). This is the case when the elements of $M' \subseteq M$ can be ordered in a sequence $(i_1, i_2, \dots, i_I)$ such that $d_{i_1 i_2}(M) > 0, d_{i_2 i_3}(M) > 0, \dots, d_{i_{I-1} i_I}(M) > 0, d_{i_I i_1}(M) > 0$. Combining this insight with the definition of the binary relation $\rhd_M$ implies that $\rhd_M$ is cyclical.

Sufficiency: The sufficiency part of the proof has a similar structure to the necessity proof. Fix a menu $M \in \mathcal{M}$. We first consider the case in which $\rhd_M = \varnothing$, therefore

$$p(j, \{i,j\}) p(i, M) = p(i, \{i,j\}) p(j, M), \quad \forall i, j \in M. \tag{D.6}$$

This means that $d_{ij}(M) = 0$ for all $i, j \in M$ for which it is defined. Lemma 3 implies that there exists a limiting MSC model that rationalizes it. We now show that it is reversible on M. Consider an arbitrary pair $i, j \in M$. If $p(i, \{i,j\}) p(j, M) > 0$ then necessarily $p(j, \{i,j\}) > 0$ and $p(i, M) > 0$. This in turn implies that $q_{ij}(\{i,j\}) > 0$ and $q_{ji}(\{i,j\}) > 0$. If $q_{ij}(M) = q_{ji}(M) = 0$, detailed balance is satisfied trivially. It is not possible that $q_{ij}(M) > 0$ and $q_{ji}(M) = 0$ or vice versa, because it would violate TR-IIA. If both $q_{ij}(M) > 0$ and $q_{ji}(M) > 0$ it holds that

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

and detailed balance is satisfied on the pair. If $p(i,\{i,j\})p(j,M)=0$ and $p(i,M)=0$ there are two possible subcases: $p(i,\{i,j\})=0$ or/and $p(j,M)=0$. Detailed balance is trivially satisfied in the latter. In the former case, $p(i,\{i,j\})=0$ implies that $q_{ji}(\{i,j\})=0$. Using TR-IIA we get that $q_{ji}(M)=0$ and detailed balance is again satisfied. Analogously, if $p(j,M)=0$ and $p(j,\{i,j\})=0$, detailed balance is satisfied. We can conclude that in the case when $\rhd_M=\varnothing$ the rationalizing model is reversible on M. We will assume for the remainder of the proof that $\rhd_M\neq\varnothing$.

In the second part of the proof, we start by creating the matrix $\mathcal{D}(M)$ and use Gordan's theorem, similarly to the necessity proof. If $\rhd_M$ is cyclical, there is a strictly positive solution to the system in (D.3) and if it is acyclical, the only non-negative solution is the zero vector. Using the solution to the homogeneous system, we can create transition probability matrices such that TR-IIA is satisfied and $\boldsymbol{p}$ is a left eigenvector.

Let $\boldsymbol{p}$ be a stochastic choice function such that $\rhd_M\neq\varnothing$. Fix a menu M and define the set G_M and the functions $f_M(.)$ and $g_M(.)$ as shown in the necessity part of the proof. We also define the matrix $\mathcal{D}(M)$ analogously. If $\rhd_M$ is acyclical there is a vector $\boldsymbol{v}\in\mathbb{R}^{|M|}$ for which

$$\boldsymbol{v}\mathcal{D}(M)\gg\mathbf{0}. \tag{D.7}$$

On the contrary, if $\rhd_M$ is cyclical, there is no such vector. Applying Gordan's theorem we know that the system

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

has a strictly positive solution if $\rhd_M$ is cyclical and no strictly positive solution if $\rhd_M$ is acyclical.

We can now use the solution to the above system to construct a transition probability matrix $Q(M)$. In particular, we define

$$q_{ij}(M)=\begin{cases}\gamma_{g_M(i,j)} & \text{if } (i,j)\in G_M,\\ 0 & \text{if } p(j,\{i,j\})=0,\\ (0,1) & \text{if } d_{ij}(M)=0.\end{cases} \tag{D.9}$$

We then let

$$q_{ji}(M) = \begin{cases} q_{ij}(M)\frac{p(i,\{i,j\})}{p(j,\{i,j\})} & \text{if } (i,j) \in G_M \text{ or } d_{ij}(M) = 0, \\ 0 & \text{if } p(i,\{i,j\}) = 0, \\ (0,1) & \text{if } d_{ji}(M) = 0 \text{ and } p(j,\{i,j\}) = 0. \end{cases} \tag{D.10}$$

We ensure that $\sum_{j \neq i} q_{ij}(M) < 1$ by scaling down all non-diagonal entries in the matrix by the same factor. Finally, we let $q_{ii}(M) = 1 - \sum_{j \neq i} q_{ij}(M)$.

Note that the constructed matrix satisfies trivially TR-IIA for all pairs i, j for which $p(i, \{i, j\}) \neq 0$. Also if $p(i, \{i, j\}) = 0$ this implies that $q_{ji}(\{i, j\}) = 0$. Since then $q_{ji}(M) = 0$, TR-IIA is satisfied as well.

Since the systems (D.8) and (C.1) are equivalent, Lemma 3 implies that the stochastic choice function is rationalizable with the constructed limiting MSC model on M. The procedure can be repeated analogously on all $M \in \mathcal{M}$.

D.4 Proof of Corollary 2

We show that among the rationalizing models there is always going to be a model with symmetric relative salience parameters by letting $t_{ij} = p(j, \{i, j\})$ for all $i, j \in N$. We then generate the rationalizing limiting MSC model from choice data analogously as shown in the proof of Theorem 1. If $q_{ij}(M) = q_{ji}(M) = 0$, then $s_{ij}(M) = s_{ji}(M) = 0$. Having $q_{ij}(M) > 0$ and $q_{ji}(M) = 0$ implies $q_{ji}(\{i, j\}) = 0$ and $t_{ji} = 0$. We can then let $s_{ji}(M) = s_{ij}(M)$. Finally, if $q_{ij}(M) > 0$ and $q_{ji}(M) > 0$, TR-IIA implies that $s_{ji}(M) = s_{ij}(M)\frac{t_{ij}}{t_{ji}}\frac{p(i,\{i,j\})}{p(j,\{i,j\})} = s_{ij}(M)$. Therefore, the constructed model rationalizes the choice data and has symmetric relative salience parameters.

E Appendix: Proofs of Theorems 2 and 3

E.1 Proof of Theorem 2

Necessity: Let $Q(M)$ be a limiting MSC model with no $i, j \in M$ for which $q_{ij}(M) = q_{ji}(M) = 0$ for all $M \in \mathcal{M}$. Assume by contradiction that there is a choice set $M \in \mathcal{M}$ and a pair $k, l \in M$ which violates the boundedness condition. This is the case when $k \rhd_M l$ and there is no cycle such that $(k, l) \in \mathcal{C}_M(M')$. This assumption implies that $\rhd_M \neq \varnothing$. We construct the set G_M, the matrix $\mathcal{D}(M)$ and the vector $\boldsymbol{\gamma}(M)$ as shown in the proof of Theorem 1. Since there is no $i, j \in M$ for which $q_{ij}(M) = q_{ji}(M) = 0$, $\boldsymbol{\gamma}(M) \gg \mathbf{0}$. We use Stiemke's lemma, which states that there exists a strongly positive solution $\boldsymbol{\gamma}(M)$ to the linear system in (D.3) if and only if there does not exist a vector $\boldsymbol{v} \in \mathbb{R}^{|M|}$ for which

$$\boldsymbol{v}\mathcal{D}(M) \geqslant \mathbf{0}. \tag{E.1}$$

Analogously to (D.5), this system can be written as

$$(v_i - v_j)d_{ij}(M) \geqslant 0, \forall (i, j) \in G_M \text{ and } (v_i - v_j)d_{ij}(M) > 0 \text{ for at least one } (i, j) \in G_M.$$

Since $\boldsymbol{\gamma}(M) \gg \mathbf{0}$, such vector $\boldsymbol{v}$ does not exist. However, since $(k, l) \notin \mathcal{C}_M(M')$, one can set $v_l > v_k$ and therefore $(v_k - v_l)d_{kl}(M) > 0$, which is a contradiction. Therefore, only when it holds for all pairs of alternatives $(i, j) \in G_M$ that there is some cycle $\mathcal{C}_M(M')$ such that $(i, j) \in \mathcal{C}_M(M')$, we can ensure that $v_i = v_j$ for all $(i, j) \in G_M$.

If the limiting MSC model is such that $q_{ij}(M) > 0$ for all $M \in \mathcal{M}$ and all $i, j \in M$, this means that $q_{ij}(\{i, j\}) > 0$ for all $i, j \in N$. Therefore, it holds that $\rho(i, Q(\{i, j\})) > 0$ for all $i, j \in N$.

Sufficiency: We now show that if choice over all pairs of alternatives in all menus is bounded in a cycle, there always exists a rationalizing limiting MSC model such that there is no $i, j \in M$ for which $q_{ij}(M) = q_{ji}(M) = 0$ for all $i, j \in M$. We construct the set G_M, the matrix $\mathcal{D}(M)$ and the vector $\boldsymbol{\gamma}(M)$ as shown in the proof of Theorem 1. We use

Stiemke's lemma, which states that there exists a strongly positive solution to the linear system in (D.8) if and only if there does not exist a vector $\boldsymbol{v} \in \mathbb{R}^{|M|}$ for which

$$\boldsymbol{v}\mathcal{D}(M) \geqslant \mathbf{0}. \tag{E.2}$$

We can alternatively write this linear system as

$$(v_i - v_j)d_{ij}(M) \geqslant 0, \forall (i,j) \in G_M \text{ and } (v_i - v_j)d_{ij}(M) > 0 \text{ for at least one } (i,j) \in G_M.$$

We will show that the assumption that all pairs of alternatives are bounded in a cycle implies that there is no pair $(i,j) \in G_M$ for which $(v_i - v_j)d_{ij}(M) > 0$. Since all pairs of alternatives are bounded in a cycle, it holds for all $(i,j) \in G_M$, that there exist a cycle $\mathcal{C}_M(M')$ such that $(i,j) \in \mathcal{C}_M(M')$, where $M' \subseteq M$ might be different for some pairs in the set. The inequality (E.1) implies that for all alternatives $\{i'_1, i'_2, \ldots, i'_I\}$ belonging to a cycle on M' it holds that $v_{i'_1} = v_{i'_2} = v_{i'_3} = \cdots = v_{i'_I}$. It holds, therefore, that $(v_i - v_j)d_{ij}(M) = 0$ for all $(i,j) \in \mathcal{C}_M(M')$. The only possible way of having a pair $(i,j) \in G_M$ for which $(v_i - v_j)d_{ij}(M) > 0$ is that i and j do not belong to the same cycle. However, if that is true, the boundedness condition would imply that $(i,j) \notin \rhd_M$ and $(j,i) \notin \rhd_M$, which in turn implies that $(i,j) \notin G_M$. This means, however, that $\boldsymbol{v}\mathcal{D}(M) = \mathbf{0}$. Therefore, the boundedness condition implies that the solution of the system (D.8) is strongly positive. We construct a transition probability matrix $Q(M)$ as shown in the proof of Theorem 1. Note in particular that it holds for all pairs $(i,j) \notin G_M$ that $q_{ij}(M) > 0$ unless $p(j, \{i,j\}) = 0$ following the proposed construction of the stochastic choice function. This is the case since the boundedness condition implies positivity of the stochastic choice function of M. The constructed Markov chain rationalizes $\boldsymbol{p}$ and satisfies TR-IIA.

If the stochastic choice function is positive for all binary menus, we can immediately see from the construction of the transition probability matrix that all entries will be strictly positive.

E.2 Proof of Theorem 3

Necessity: Let $Q(M)$ be an ergodic limiting MSC model. Hence, for each menu M there is a sequence of alternatives $I(M)$ such that it holds for all consecutive elements (i, j) that $q_{ij}(M) > 0$. We want to show that the generated stochastic choice function is positive on binary sets and that for all subsequent pairs of alternatives in $I(M)$ it holds that the induced choice over the pair is bounded in a cycle.

The positivity of the stochastic choice function on binary sets is an immediate consequence of the irreducibility of the Markov chain. Note that irreducibility on binary menus requires that $q_{ij}(\{i, j\}) \in (0, 1)$ for all $i, j \in N$.

We now need to show that it holds for all consecutive pairs in $I(M)$ that the generated choice is bounded in a cycle. Assume by contradiction that this is not the case for some pair i, j. Then, $i \triangleright_M j$, but there is no cycle $\mathcal{C}_M(M')$ for which $(i, j) \in \mathcal{C}_M(M')$. Proposition 8 then implies that all rationalizable models of $\boldsymbol{\rho}(Q(M))$ are such that $q_{ij}(M) = q_{ji}(M) = 0$, which is a contradiction.

Sufficiency: Let $\boldsymbol{p}$ be a stochastic choice function which is positive on binary choice sets and for which there is a sequence of the alternatives in the menu $I(M) = (i, j, \dots)$ (possibly with repetitions) such that all consecutive pairs (i, j) are bounded in a cycle. We will first show that the rationalizing model is such that for all $M \in \mathcal{M}$ and $i, j \in M$ it holds that either $q_{ij}(M) > 0, q_{ji}(M) > 0$ or $q_{ij}(M) = q_{ji}(M) = 0$. Then, we will show that there exists an irreducible limiting MSC model, in particular that for all consecutive elements in the above sequence the transition probability is strictly positive. Since limiting MSC models are aperiodic by definition, proving irreducibly implies that the limiting MSC model is ergodic.

Suppose that there is a rationalizing limiting MSC model with $q_{ij}(M) > 0$ and $q_{ji}(M) = 0$. Since $\boldsymbol{p}$ is positive on binary choice sets, we must have $q_{ij}(\{i, j\}) > 0$ and $q_{ji}(\{i, j\}) > 0$. Under these assumptions, TR-IIA is violated and we have found a contradiction.

We now show the existence of a limiting MSC model with strictly positive transitions between every two consecutive alternatives in $I(M)$. We start by constructing a stochastic choice function $\boldsymbol{p}'$ from $\boldsymbol{p}$, where we adjust $p(k, \{k, l\})$ for all $k \triangleright_M l$ that are not bounded in a cycle such that $d_{kl}(M)' = 0$. Note that the constructed stochastic choice function $\boldsymbol{p}'$ has to satisfy the boundedness condition on all pairs $i, j \in M$. It also satisfies positivity on binary sets because we assume it for $\boldsymbol{p}$.

Theorem 2 implies that $\boldsymbol{p}'$ is rationalizable with a limiting MSC model $Q^{'}(M)$ with only strictly positive transitions. Now let $Q(M)$ be such that $q_{ij}(M) = q_{ij}(M)'$ for all pairs i, j that are bounded in a cycle and $q_{ij}(M) = q_{ji}(M) = 0$ for all i, j that are not bounded in a cycle. The constructed matrix $Q(M)$ satisfies equation (C.1) and hence rationalizes the stochastic choice function $\boldsymbol{p}$. Since we only adjust the transition probability between pairs of alternatives that do not belong to the sequence $I(M)$, in the constructed model the transition probabilities between the subsequent elements of $I(M)$ are strictly positive. Since the sequence spans over the whole menu, the constructed Markov chain is ergodic.

F Proofs of the identification results

F.1 Proof of Proposition 8

Let $\boldsymbol{p}$ be a stochastic choice function rationalizable by a limiting MSC model $Q(M)$ which is positive on binary choice sets and it holds for at least one pair $i, j \in M$ that the choice over the pair is not bounded in a cycle. This implies that $i \triangleright_M j$ w.l.o.g. and there is no cycle $\mathcal{C}_M(M')$ such that $(i, j) \in \mathcal{C}_M(M')$. Similarly to the proof of Theorem 1, we will use the function $d(M)$ defined in Appendix C, which is always defined because $\boldsymbol{p}$ is positive on binary sets. Note that $i \triangleright_M j$ implies $d_{ji}(M) > 0$. We assume by contradiction that there is a rationalizing model $Q(M)$ with $q_{ij}(M) > 0$. Because of TR-IIA it holds also that $q_{ji}(M) > 0$. Since $\boldsymbol{p}$ violates the boundedness condition, it follows from Theorem 2 that there has to be at least one pair $k, l \in M$ for which $d_{kl}(M) \neq 0$ and $q_{kl}(M) = 0$.

Since $\triangleright_M \neq \varnothing$, we define the set G_M and the functions $f_M(.)$, $g_M(.)$, and a matrix $\mathcal{D}(M)$

in the same way as we do in the sufficiency proof of Theorem 1 (see Appendix D.3). Let w.l.o.g. that $(i, j) \in G_M$ and $\mathcal{P}(G_M)$ denote the power set of $G_M \backslash (i, j)$ with characteristic element $\mathcal{G}$.

We show the contradiction by constructing stochastic choice functions $\boldsymbol{p}'$ from $\boldsymbol{p}$ for each set in $\mathcal{P}(G_M)$ by adjusting $p(k, \{k, l\})$ such that $d_{kl}(M)' = 0$ for all $(k, l) \in \mathcal{G}$, for all $\mathcal{G} \in \mathcal{P}(G_M)$. What this means is that $G'_M \subset G_M$ and that $\boldsymbol{\gamma}(M)'$ has less dimensions than $\boldsymbol{\gamma}(M)$. Among those stochastic choice functions there needs to be one, say $\boldsymbol{p}^*$, in which exactly those pairs of alternatives are excluded from G^*_M for which there is zero transition probability in $Q(M)$, that is, for all $k, l \in M$ for which $q_{kl}(M) = 0$ and $(k, l) \in G_M$ holds that $(k, l) \notin G^*_M$ and $d_{kl}(M)^* = 0$, and for all $k, l \in M$ for which $q_{kl}(M) > 0$ and $(k, l) \in G_M$ holds that $(k, l) \in G^*_M$.

Let $Q(M)^*$ be such that $q_{mn}(M)^* = q_{mn}(M)$ for all $m, n \in M$ for which $q_{mn}(M) > 0$ and $q_{mn}(M)^* \in (0, 1)$ and $q_{nm}(M)^* = q_{mn}(M)^* \frac{p(m,\{m,n\})^*}{p(n,\{m,n\})^*}$ for all $m, n \in M$ for which $q_{mn}(M) = 0$. In particular, since $q_{ij}(M) > 0$, $q_{ij}(M)^* = q_{ij}(M)$. Due to the way $\boldsymbol{p}^*$ is constructed from $\boldsymbol{p}$, the choice over the pair i, j according to $\boldsymbol{p}^*$ is not bounded in a cycle.

The constructed limiting MSC model rationalizes $\boldsymbol{p}^*$ and has strictly positive transition probabilities. Theorem 2 implies that all pairs of alternatives are bounded in a cycle. However, the pair i, j violates the condition, which is a contradiction. Therefore, we must have $q_{ij}(M) = 0$.

F.2 Proof of Proposition 10

We start by recalling a result in Lemma 3 in Appendix C, which follows directly from the definition of rationalizable stochastic choice functions with the limiting MSC model

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

We simplify the above equation, using that $M = \{i, j, k\}$ and that $q_{ij}(M) \neq 0$ for all $i, j \in M$ and obtain:

$$(p(i, M)q_{ij}(M) - p(j, M)q_{ji}(M)) + (p(i, M)q_{ik}(M) - p(k, M)q_{ki}(M)) = 0.$$

We use Proposition 9 and substitute $q_{ij}(M) = s_{ij}(M)p(j, \{i, j\}) = s_{ji}(M)p(j, \{i, j\})$ for all $i, j \in M$ and obtain

$$\frac{s_{ij}(M)}{s_{ik}(M)} = \frac{p(k, M)p(i, \{i, k\}) - p(i, M)p(k, \{i, k\})}{p(i, M)p(j, \{i, j\}) - p(j, M)p(i, \{i, j\})}.$$

We have that $s_{ij}(M) > s_{ik}(M)$ if

$$\frac{p(k, M)p(i, \{i, k\}) - p(i, M)p(k, \{i, k\})}{p(i, M)p(j, \{i, j\}) - p(j, M)p(i, \{i, j\})} > 1.$$

We can now factor out $p(i, M)p(i, \{i, k\})$ from the nominator and $p(i, M)p(i, \{i, j\})$ from the denominator to obtain the following inequality:

$$\frac{p(i, M)p(i, \{i, k\}) \left(\frac{p(k,M)}{p(i,M)} - \frac{p(k,\{i,k\})}{p(i,\{i,k\})} \right)}{p(i, M)p(i, \{i, j\}) \left(\frac{p(j,\{i,j\})}{p(i,\{i,j\})} - \frac{p(j,M)}{p(i,M)} \right)} > 1.$$

We can cancel out the term $p(i, M)$ and rearrange to obtain the final result.