Visual Choice

Yusufcan Demirkan *

October 19, 2023

Abstract

I propose a procedural model of random choice called visual choice based on the eye movements of individuals. The decision maker fixates randomly on alternatives, and performs saccades to other alternatives based on the fixation point. The choice probabilities are determined using a random tie-breaking between these saccaded alternatives and weighting according to the fixation probabilities, both formulated using the well-known Luce rule. I discuss the cases of fixation-independent and fixation-dependent saccades. The former connects to the recent literature extending the Luce rule giving a procedural interpretation for these choice rules. For the latter, I discuss a special case called satisficing visual choice which determines the choice probability of an alternative x by calculating the weighted average of the contribution coming from the fixated alternatives strictly dominated by x. The main results of the model are related to the rationality properties of this procedure, and a behavioral characterization result based on conditions imposed on the choice data.

1 Introduction

Imagine that you want to cook a dinner for your family, and you decided to prepare something involving fish. Being indecisive about which fish to cook, you go to the market and look at the fish section. At a moment, you find yourself looking at the nicely arranged salmon options. However, this does not stop you to check other options in a quick fashion even without rotating your head. Of course, this does not imply that you are going to buy anything you see, but if you can find something better than salmon during these quick eye movements, you might end up buying it. At the end of the day, you cook turbot at home, and decide to watch a movie together after the dinner. The streaming channel you own presents multiple suggestions, some of which attract your attention immediately. Finding yourself looking at the second suggested option, you realize that there is a movie of your favourite actor just below it. Because the topic of that movie seems boring, you decide to watch the movie suggested by the platform.

What is common to these stories? A decision-maker (henceforth DM) has a certain goal (like cooking a dinner or watching a movie) and there is a set of associated choice options constituting the choice problem (the alternatives present in the fish section or the movies on the online streaming platform). The DM sees these options as they are presented to her, but whether she will be aware of these or not will depend on her vision. The salient options might attract her attention towards these, resulting in a relatively longer period of looking at these alternatives.

* *New York University, Department of Economics, 19 West 4th Street, 6th Floor, New York, NY 10013. yd1131@nyu.edu*

In the literature of vision science, one says that the DM *fixates* on these options, and it is further known that an individual can focus on only a small region of her visual field at a single point of time. I fix this to be a single alternative available in the menu (such as salmon or the suggested movie). Moreover, it is known that the human eye performs *saccade* movements after fixating at a certain point. If the DM further likes an option she saccaded to (turbot or the movie of the favourite actor), then she might end up choosing this instead of what she fixated. I present a model with the goal of capturing these type of situations, which can be seen as a part of the fundamental process underlying how people decide due to its connection with how they see their environment.

The first goal of this paper is to provide a tractable framework that reflects how people make decisions based on what they see. Consider the following fairly general choice rule:

$$\rho_x(M) := \sum_{y \in M} q(M,y) \cdot \mathbf{p}_{(M,y)}(x)$$

The mapping q determines the *fixation distribution* of the DM in *menu* (or *choice problem*) M. The DM looks at y with probability $q(M,y)$ in M. This pair (M,y) then determines the conditional choice probabilities of all alternatives available in M, denoted by $\mathbf{p}_{(\mathbf{M},\mathbf{y})}$.[1] However, this rule is empirically uninteresting due to its generality, but provides the foundation for the choice rule that will be called *visual choice*. Assume that the DM is endowed two (strictly) positive value functions u and v, which are interpreted respectively as the salience value that determines the fixation probabilities and saccade probabilities.[2] Furthermore, assume that the DM is also endowed with a *saccade correspondence* s that determines the alternatives DM performs saccadic eye movements to conditional on a fixation point. The choice rule is then defined using the well-known *Luce rule*:

$$\rho_x(M) := \sum_{y \in M} \frac{u(y)}{u(M)} \cdot \begin{cases} \frac{v(x)}{v(s(y,M))} & \textit{if} \quad s(y,M) \neq \emptyset \ \& \ x \in s(y,M) \\ 1 & \textit{if} \quad s(y,M) = \emptyset \ \& \ x = y \\ 0 & \textit{o.w.} \end{cases}$$

This choice rule reflects the process of visual decision-making described through examples previously. The DM looks at an alternative y in M with a probability equal to its relative salience value in M.[3] After fixating at this point, she performs saccades to alternatives in $s(y,M)$, and then chooses an alternative $x \in s(y,M)$ again with its relative salience value in these saccaded alternatives. When the DM performs no saccades, the fixation point is chosen with probability 1, which can interpreted as a *status-quo effect*.[4] If furthermore saccade correspondence is *rational* in the sense that it satisfies the *weak axiom*, then this choice procedure becomes *rational*

[1]This can be interpreted also as a stochastic version of *choice problem with a status-quo/reference point* as in Masatlioglu and Ok [34] after the DM fixates at y in M.

[2]The reason for assuming two separate salience mappings is due to the dynamic nature of fixation and saccade. More precisely, if one imagines a two-period decision problem, then one can think of u as the salience mapping for time 0, while v is the salience mapping for time 1. Even without temporal interpretation, it is possible that fixation and saccade are affected by different salience functions. Alternatively, v can be interpreted as Luce value which is used to tie-break between different alternatives saccaded by the DM.

[3]See Bundesen [4] for a support of the view that salience is determined using such formula, although with some additional ingredients that are not modeled in this paper.

[4]As opposed to the usual notion of choice correspondence, I allow saccade correspondence to be empty-valued, but mostly use the induced correspondence which is everywhere nonempty due to the assumption I made about the status-quo effect. More through explanation will be provided later on.

visual choice. I discuss two cases of the visual choice that depend on the properties of the saccade correspondence s.

First, I assume that s is *fixation-independent*, i.e. $s(x,M) = s(y,M)$ for any fixation point $x,y \in M$, which connects the visual choice to the recent literature that tries to extend the Luce rule in order to deal with its implausible implications such as nonzero choice probabilities. Looking at this special case has two benefits: the choice rules in the literature which do not have a procedural interpretation gain an interpretation in this terms using visual choice, and also their characterization results provide an empirical test for the fixation-independent visual choice. In particular, if s is nonempty-valued and fixation-independent, then the visual choice becomes equivalent to the *general Luce model* of Echenique and Saito [13]. Thus, general Luce can be interpreted as a visual choice rule in which fixations do not matter in determining the saccades, and choice probabilities are only determined by the saccades made in the corresponding menu. If furthermore the visual choice is rational, meaning that the saccades performed can be rationalized, then one has the *preference-oriented Luce rule* of Dogan and Yildiz [11]. Two extreme cases of this result in the *Luce rule*: when the DM perform no saccades or she performs saccades to all alternatives in the menu.

After discussing the fixation-independent case, I relax the visual choice to be fixation-dependent. Let $\succeq$ be a reflexive, transitive and antisymmetric binary relation (a partial order) which represents the *dominance relation* of the DM: for example, if alternatives can be represented as bundles of features, then an alternative dominates another one iff it has any feature the other alternative has (the vector ordering).[5] I assume that the DM performs saccades only to alternatives that are strictly dominate the fixation point, and if there are none, then the status-quo effect is in power, meaning that the DM chooses the fixation point. I call this specific choice rule *satisficing visual choice*, where the satisficing level is determined by the fixation point.

A core goal of the satisficing visual choice (svc) is to analyze the rationality of this choice procedure once top-down processes (determined by the dominance relation) is allowed to influence the bottom-up processes (determined by visual salience functions u and v). Therefore, I start by discussing the rationality of the svc. This discussion makes use of several properties used in the literature (such as *regularity*, *monotonicity*, and *stochastic transitivity*), the recent *stochastic rationality measure* developed by Ok and Tserenjigmid [38], and finally models of *deliberate randomization*. To connect the top-down to the bottom-up, I define the notion of *compatibility* and *strong compatibility*. On top of the rationality stemming from the consistency of the saccade correspondence conditional on the fixation point, these notions further provide connection by making the visual value functions u and v compatible with preferences $\succeq$.[6] Without any condition of compatibility, satisficing visual choice allows violations of regularity, monotonicity, and even weak stochastic transitivity, and it is able to capture well-known behavioral regularities such as *attraction effect* and *compromise effect*. If furthermore u is compatible, then svc satisfies moderate stochastic transitivity, but still not necessarily strong stochastic transitivity (even under strong compatibility). This is consistent with the experimental evidence that people obey moderate stochastic transitivity, even though they frequently violate the strong version (see Tversky [48] and Rieskamp et al. [42]).

The more developed results use the stochastic rationality measure developed by Ok and Tserenjigmid. A random choice function (rcf) is called *maximally (minimally) rational* if it is more

[5]This will be the main interpretation I will use.

[6]Compatibility of u here means that if $x \succeq y$ ($x \succ y$) then $u(x) \geq u(y)$ ($u(x) > u(y)$). It is similarly defined for v. The notion of strong compatibility will be defined in the main body.

(less) rational compared to any other rcf. For the first result I fix the preferences and assume further it is complete. In addition, I assume that both of the visual value functions are compatible with the underlying preferences, that is, both u and v are compatible with $\succeq$. I show that even in this case no svc is maximally or minimally rational, although one can always compare different rcf's which are svc's by changing the parameters u and v. The second result goes one step further and characterizes when the DM's choices can be called maximally rational under the assumption of strong compatibility. This happens either when she cannot compare any alternatives (which is equivalent to the Luce rule) *or* there is a unique alternative that dominates the rest of the alternatives with the remaining alternatives being incomparable (which is equivalent to a slight variant of Luce rule I call an *almost Luce rule*). Surprisingly, both these cases involve *minimal levels of comparison* with respect to the preferences, while the choice data might reveal that the DM is maximally rational. In a sense, if the DM has minimal level of decisiveness and "leaves" herself only to the bottom-up processes (involving usually unconscious processes), then the choices of the DM are as if they are maximally rational.

After discussing the rationality features, I provide an empirical characterization result. The identification of preferences is easily handled from binary choice problems, and it is by definition unique. However, the identification for the visual value functions u and v are more subtle. Indeed, full identification is impossible, and without identifying u one cannot identify v. First, I provide when and why identification of u is possible, and in which cases it cannot be identified. Then I do a similar analysis for the saccade value v. The characterizing axioms can be separated into three groups. Those which are mainly related to rationalizability and hence the underlying preferences (*dominance transitivity*, *rationalizability*, and *independence from incomparable alternatives*), those that are related to the construction of the fixation probabilities and relate it to regularity violations (*cyclical independence*, *fixation consistency*, and *bounded fixation ratio*), and finally the condition that imposes consistency on saccade values (*path independent saccade*).[7] Rationalizability puts a minimal requirement for an alternative to be chosen with strictly positive probability in a menu: an alternative should either dominate at least one other alternative or should be incomparable to the rest. Dominance transitivity says that if x and y are chosen with probability 1 in menus $\{x,y\}$ and $\{y,z\}$, then x should be chosen with probability 1 in $\{x,z\}$. Both independence from incomparable alternatives and cyclical independence are relaxations of *Luce's IIA*, which in a sense restricts the independence menus that only differ by incomparable alternatives to the rest of the menu and paths that involve alternatives that consists of consecutively incomparable alternatives. The rest of the conditions guarantee that fixation and saccade probabilities are well-defined and consistent. Fixation-consistency posits that the probability of not looking at an alternative gives the relative fixation probability increase due to adding that alternative, while bounded fixation ratio puts an upper bound on this relative fixation ratio using the relative probability changes and hence the relative size of regularity violations induced by this addition.

This paper is mainly related to two literatures. Firstly and more closely, the procedural model proposed in this paper is a random choice model, a growing literature that is partly inspired from the cognitive mechanisms in the decision-making process. Three literatures stand out in close connection to visual choice: the models generalizing the well-known Luce rule, the models that treat attention as a random process while keeping preferences fixed, and finally the models that assume that reference-points affect the choices. There are several models in the recent literature

[7]Some of these conditions are familiar from the literature. Dominance transitivity and cyclical independence are used by Echenique and Saito [13] to characterize general Luce model, even though the underlying dominance relation in this paper and here are different.

that deal with the limitations of the Luce rule such as *general Luce model* of Echenique and Saito [13], *limited consideration Luce model* of Ahumada and Ulku [1], *nested logit model* of Kovach and Tserenjigmid [28], *focal Luce model* of Kovach and Tserenjigmid [29], and *perception-adjusted Luce model* by Echenique et al. [14]. Visual choice also makes use of Luce rule formulation, but the model uses it for the formation of salience and not try to solve the problems related to it. Another stream of these models fix the preferences of the DM by assuming it satisfies the main properties of rationality, and relax the requirement that people are perfectly attentive, treating it as a random process. Two prominent models of random attention are proposed by Manzini and Mariotti [33] and Cattaneo et al. [6]. The main connection of this model to the models of attention is that it proposes a particular process of attention formation using the eye movements of individuals. Even though visual choice is not nested by the random attention model of Cattaneo et al., imposing the rationality condition through the weak axiom on saccade correspondence makes it a model of random attention, which implies that there exists a particular type of random attention process and an associated complete preference relation that models the same random choices. A final type of models focus on the impact of the *reference points* on the final choices of the DM which is initiated by the deterministic model presented by Masatlioglu and Ok [34], recently endogenized using a random reference selection by Kibris et al. [26]. The fixation point can be interpreted as a reference point for the DM, and indeed the specific case satisficing visual choice bears close resemblances to logit random reference model, even though it is not nested by it. An advantage of svc compared to this model is that it uses a small number of primitive parameters, while in random reference model there is a complete preference relation associated with each alternative in addition to the salience function.

In addition to the random choice literature, the model presented in this paper is related to the literature on the visual system of humans. Indeed, this literature forms the basis of the procedural model presented here. The literature on vision expands into different branches of sciences; including neuroscience, medicine, psychology, physics and computer science, and it has its own branch called vision science. Although the literature is vast, the model presented here relies on the fundamental eye movements, and postulates a choice procedure based upon these. It is known that eye performs variuous types of movements such as saccadic movements and smooth pursuit movements (although these are not the exclusive list of movements). Even though both movements determine the gaze direction, the fixation point is largely determined by the slower smooth pursuit movements, while saccadic movements are mostly unaware quick movements after eyes are fixated at a certain point. The notion of attention, which is also very prominent in the economics literature, usually depends on how the visual system works and what is perceived to be *salient* in a natural scene. Because the information load in a typical scene is very high, it is suggested that the human visual perception occurs in two stages. In the first stage, the simple features in the environment are processed in a parallel fashion (*preattentive mode*), while in the second stage the focus of attention is directed at a particular location in the visual field using the data collected in the first stage (*attentive mode*) (see Treisman [46] and Treisman and Gelade [47]). It is articulated that the neural system prepares a *saliency/conspicuity map* which is responsible to filter out objects that are salient compared to the other objects in the environment, which is computed using the the features objects have. A prominent approach is the *winner-take-all mechanism* which chooses the most salient point in the environment, and passes to the next most salient location using the *inhibition-of-return* (see Itti and Koch [24] and Koch and Ullman [27]). Luce rule is a normalization to measure the relative saliency and hence the probability of being the fixation point, and it is compatible with the principles related to this approach. The model here also accounts for the *saccadic movements*, which are hypothesized to be predetermined conditional on the fixation point, and this is reflected through the saccade correspondence of the DM that depends on the fixated point. Thus, the overall procedure postulates that at a moment of fixation at x, the individual makes a probabilistic selection depending on the saccade correspondence conditional on x, and the final choice probabilities are determined by aggregating different fixation points through the probability of viewing them using their relative saliences.

The structure of the paper is as follows. In Section 2, I provide the definition of the visual choice model. The deterministic version of the model and its characterization is provided in Section 3. Then, I start analyzing the (random) visual choice by assuming fixation-independence in Section 4. I relax the assumption of fixation independence in Section 5, and focus on a specific type of saccade correspondence. First, I discuss the rationality properties of the choice procedure induced by this correspondence in Section 5.1, and then provide a characterization result in Section 5.2. In Section 5.3, I discuss the comparative statics of the model, and also its relation to the random choice literature that is closest to this work. In Section 6, I discuss further issues not discussed in the main body, and conclude in Section 7 by pointing out to limitations in visual choice. All the proofs that are not in the main body are in the Appendix, as well as some further related literature discussion.

2 Visual Choice

Let X be the grand set of alternatives. The decision-maker (henceforth DM) chooses from a *menu* M which is a nonempty finite subset of X. The set of all such menus is denoted as $\mathbb{X}$. A **random choice function (rcf)** is a probability distribution $\rho : X \times \mathbb{X} \to [0,1]$ such that $\rho(x,M) = 0$ if $x \notin M$ and $\sum_{x \in M} \rho(x,M) = 1$. I will denote $\rho(x,M)$ simply as $\rho_x(M)$. Consider an arbitrary menu M. First, I am going to provide a general model which is consistent with the visual interpretation. The DM *fixates* at alternatives in M randomly, which can be modeled by a probability distribution $q : \mathbb{X} \times X \to [0,1]$ such that $q(M,x) = 0$ for any $x \notin M$ and $\sum_{x \in M} q(M,x) = 1$ for any menu M. Once DM fixates her eyes at y in M, she chooses an alternative in the menu (including possibly y) with some probability depending on y. Formally, let $\mathbf{p}_{(M,y)}(x)$ denote the probability of choosing x conditional on y in menu M. The induced choice probability gives the following choice rule[8]:

$$\rho_x(M) := \sum_{y \in M} q(M,y) \cdot \mathbf{p}_{(M,y)}(x)$$

This choice rule is unfalsifiable. To see, consider a random choice data $\tilde{\rho}$ and assume that $\mathbf{p}_{(M,y)}(x) = 1$ iff $x = y$, letting $q(M,y) = \tilde{\rho}_y(M)$. This shows that any choice data can be accommodated by the above rule. Therefore, I am going to put additional structure that is inspired by the studies in vision science, and then define the visual choice.

The main concepts coming from the vision literature is the notion of salience mapping which determines where the DM looks at in a particular environment, and the notion of ballistic movement for the saccades, that says the saccadic movements are predetermined given a particular fixation point. Assume that the DM is endowed with a pair of functions (u,v) defined on X both taking strictly positive real values. It is possible that these two are equal to each other, which I am going to assume in some cases. Let $u(M) := \sum_{y \in M} u(y)$ and $v(M) := \sum_{y \in M} v(y)$.

[8] Towal et al. [45] formulated a very similar model with much different ingredients, which relies on a **drift-diffusion model (ddm)**.

For the saccadic movements, let $s: X \times \mathbb{X} \to \mathbb{X}$ be a *saccade correspondence* such that $s(x,M) \subseteq M$ for any $M \in \mathbb{X}$, which is allowed to be empty.[9] If the DM does not perform any saccades from x so that $s(x,M) = \emptyset$, I assume that the DM chooses x itself conditional on fixating at x. Once fixation at x occurs, I assume that the DM will perform saccades to $s(x,M)$.

Definition 1. *A rcf ρ is called **visual choice (vc)** if for each menu $M \in \mathbb{X}$ and $x \in M$ there is a pair of functions (u,v) such that $u,v : X \to \mathbb{R}_{++}$ and a saccade correspondence $s : X \times \mathbb{X} \to \mathbb{X}$ where $s(x,M) \subseteq M$ for any $M \in \mathbb{X}$ such that:*

$$\rho_x(M) := \sum_{y \in M} \frac{u(y)}{u(M)} \cdot \begin{cases} \frac{v(x)}{v(s(y,M))} & \text{if} \quad s(y,M) \neq \emptyset \ \& \ x \in s(y,M) \\ 1 & \text{if} \quad s(y,M) = \emptyset \ \& \ x = y \\ 0 & \text{o.w.} \end{cases} \tag{1}$$

The analysis of visual choice can be categorized into two subcases: *fixation-independent* and *fixation-dependent* visual choice. In fixation-dependent visual choice, the alternatives that are saccaded by the DM conditional on the fixation point does not depend on the fixation point, so $s(x,M) = s(y,M)$ for any $x,y \in M$ and $M \in \mathbb{X}$. As will be shown later, the fixation-independent case is connected to the recent literature that extends the Luce model to deal with the limitations such as the nonallowance of zero choice probabilities. While considering this case, I assume that s is nonempty. A special case is when s is *rational* in the sense that it satisfies *weak axiom of revealed preference*: $s(M') \cap M = s(M)$ for any $M \subseteq M'$ provided that $\hat{s}(M') \cap M \neq \emptyset$.

For the case of fixation-dependence, I will focus on a special case which exemplifies a type of *rational* visual choice, called *satisficing visual choice*. One can define the following *induced saccade correspondence* $\hat{s}$ if s is allowed to be fixation-dependent for ease of notation:

$$\hat{s}(x,M) := \begin{cases} s(x,M) & \text{if} \quad s(x,M) \neq \emptyset \\ \{x\} & \text{o.w.} \end{cases}$$

Using the induced correspondence, the rationality condition amounts to the following for the fixation-dependent case:

- $\hat{s}(x,M') \cap M = \hat{s}(x,M)$ for any $x \in M \subseteq M'$ provided that $\hat{s}(x,M') \cap M \neq \emptyset$.

Let $\succeq$ be a partial order defined on X, which is intended to capture the *dominance relation*. Interpreting an alternative x using the feature space approach, I will use the vector ordering induced by the feature comparisons, so $x \succeq y$ iff x is at least good as y with respect to all features. If two alternatives x and y are incomparable according to $\succeq$, I will denote this as $x \bowtie y$. Partition the alternatives in X according to their position with respect to y using $\succeq$. Finally, let $D(y) := \{x \in X : x \succ y\}$, $W(y) := \{x \in X : y \succ x\}$, and $Inc(y) := \{x \in X : x \bowtie y\}$. Denote the restriction of $D(y)$ to menu M as $D(y,M)$, and similarly for others. The satisficing visual choice is defined next.

Definition 2. *A rcf ρ is called **satisficing visual choice (svc)** if it is visual choice such that there is a partial order $\succeq$ and $s(x,M) = D(x,M)$, which implies that:*

$$\rho_x(M) := \sum_{y \in M : x \succeq y} \frac{u(y)}{u(M)} \cdot \begin{cases} \frac{v(x)}{v(\hat{s}(y,M))} & \text{if} \quad x \in \hat{s}(y,M) \\ 0 & \text{o.w.} \end{cases} \tag{2}$$

[9] This is just a different interpretation of choice correspondence with a reference point/status-quo formulated by Masatlioglu and Ok [34] and Rubinstein and Zhou [43].

To see that this exemplifies rational visual choice, note that warp conditional on the fixation point implies that for each fixation point y, there is a complete preference relation $\succsim_y$ that rationalizes saccade correspondence, so $\hat{s}(y,M) \equiv \max(M, \succsim_y)$. Thus, $\succsim_y$ will be the *satisficing relation* that is defined with respect to y. Assume that $x \sim_y z$ if one of the following holds: $x,z \in D(y)$, $x,z \in W(y) \cup Inc(y)$ or $x = y = z$. The strict part of $\succsim_y$, $\succ_y$, is defined as follows: $x \succ_y z$ if $x \in D(y)$ and $z \notin D(y)$ or $x = y$ and $z \in W(y) \cup Inc(y)$. Given fixation at y in menu M, any alternative that is satisficing with respect to y can be saccaded from y. This implies according to the definition of $\succsim_y$ that:

$$\max(M, \succsim_y) = \begin{cases} \{y\} & \textit{if} \quad D(y,M) = \emptyset \\ D(y,M) & \textit{if} \quad D(y,M) \neq \emptyset \end{cases}$$

so that $\max(M, \succsim_y) = \hat{D}(y,M)$.

The main interpretation of visual choice is based on *feature spaces*. Assume that $X \subseteq \mathbb{R}^k_+$ is the feature space for $k \geq 1$ where x_i denotes the level of feature i in alternative x. A special case is when each feature is binary and any alternative x is defined as a vector of $0-1$'s where $x_i = 1$ iff x has feature i. The salience value u and the utility value v can be computed using the underlying feature space. In particular, I am going to use an *additive* specification in this paper. Accordingly, let $u(x) = \sum_{i \in \{1,\ldots,k\}} \alpha_i x_i$ and $v(x) = \sum_{i \in \{1,\ldots,k\}} \beta_i x_i$ where $\alpha_i, \beta_i > 0$ for all i. Since u and v are assumed to be strictly positive, I will take any $x \in X$ to have at least one feature i such that $x_i > 0$. The feature space can also determine the saccade correspondence. For example, an individual might only perform saccades to alternatives that dominate the fixated alternative in all dimensions, or alternatives that are sufficiently differentiated from it. I will assume that both fixation probabilities and conditional choices due to saccades are the results of the Luce rule, where the latter is only applied on the saccade correspondence. When using this interpretation, I will focus on the $k = 2$ case unless otherwise specified.

The visual choice can be interpreted as follows. The features of each alternative y determine their relative salience in a menu M, computed using the Luce rule $\frac{u(y)}{u(M)}$, which gives the probability of fixating at y in M. Given fixation at y, the individual performs saccadic eye movements to other alternatives in the menu, determined through saccade correspondence $s(y,M)$. Each alternative in the correspondence has a probability of being chosen according to their relative salience computed using v and Luce rule, but now only using the alternatives that are in the saccade correspondence. So, if $x \in s(y,M)$, then x is chosen with (at least) $\frac{u(y)}{u(M)} \cdot \frac{v(x)}{v(\hat{s}(x,M))}$ probability from M due to being saccaded from x. The DM does not fixate on only one alternative due to the inhibition of return principle, and continue to fixate on other alternatives with a probability equal to their relative salience value computed using the same rule, performing saccadic movements conditional on these fixation points. Thus, the choice probability of an alternative x in M is determined by the weighted aggregation of all these saccades from fixation points.

It is important to note that there are two limitations of this formulation. First, the strict positivity of u implies that the DM fixates on all alternatives with positive probability, which might not be the case in reality. I do not view this as a serious limitation, since the alternatives that DM does not look at can be modeled using sufficiently small but positive u value. The more important limitation is that the salience value of an alternative determined through u and v is independent from the menu, and only depends on its underlying features. This can be also unrealistic in certain contexts. I will try to address this limitation briefly in further discussion using the salience function formulated by Bordalo et al. [3].

Before proceeding to analyze the visual choice, I will provide an illustrative example in order to illuminate the choice procedure proposed in the paper.

Example 1.

Assume that u,v values are equal for all alternatives. According to the feature space interpretation, this holds when $\alpha_i = \beta_i$ for any dimension i. I will consider two cases: in the first one, the saccades will be independent from the fixation point. A case in which this is satisfied is when $s(x,M) = s(M)$ for any $x \in M$ and $M \in \mathbb{X}$. When this is the case, the individual fixates at each alternative with equal probability of $\frac{1}{|M|}$. Since the individual is assumed to saccade to all alternatives in M (including x, which is called a *refixation* in the vision science terminology), each alternative in M gets an equal share from fixation at x, equal to $\frac{1}{|M|}$. Therefore, the choice probability of any alternative x in M is given by:

$$\rho_x(M) = \sum_{y \in M} \frac{1}{|M|} \cdot \frac{1}{|M|} = \frac{1}{|M|}$$

so that each alternative is chosen with equal probability. Thus, if each alternative has equal salience and utility value with the assumption that the DM performs saccades to all alternatives, then each alternative is chosen with equal probability according to visual choice.

For the latter example, assume that the individual is endowed with a complete and asymmetric preference relation $\succ$ on X. Let $k(M)$ denote the k^{th} worst alternative in M, and assume that the individual only saccades to alternatives that strictly dominate the fixation point.[10] This implies the following choice probabilities:

$$\rho^{k(M)}(M) = \frac{1}{|M|} \sum_{j<k} \frac{1}{|M| - j}$$

if $k(M) \leq |M| - 1$, and $\rho_{|M|(M)}(M) = \rho^{|M|-1}(M) + \frac{2}{|M|}$. So, the DM fixates on each alternative with a uniform probability, and the contribution of each is again uniformly distributed over the alternatives that dominate it.

Example 2.

Again consider the previous setting with $u = v$. The choice probabilities can be written more generally as an adjusted Luce rule in the following form for the case of svc:

$$\begin{aligned}
\rho_x(M) &= \sum_{x \succ y} \frac{u(y)}{u(M)} \cdot \frac{u(x)}{u(D(y,M))} \\
&= \frac{u(x)}{u(M)} \cdot \sum_{x \succ y} \frac{u(y)}{u(D(y,M))} \\
&\equiv p^u_x(M) \cdot \sum_{x \succ y} \tilde{u}(y,M)
\end{aligned}$$

where p^u denotes the Luce choice probability with value function u and $\tilde{u}(y,M)$ is the *rationalizability-adjusted value* of an alternative y in M. The assumption that $u = v$ makes the procedural interpretation more apparent. More precisely, the first component is the *fixation probability* of alternative x in menu M determined by salience values u, and the second component finds the total *saccade probability* from alternatives that x strictly dominates. While the first expression

[10] Alternatively, one can interpret this as if the individual saccades to all alternatives in the menu, but only the ones that are strictly better are taken seriously by the DM.

is exactly the Luce rule, note that the second expression is not exactly so because the denominator only accounts the value of alternatives that dominate y. Let $X = \{x,y,z\}$ with $x \succ y \succ z$. In any binary menu, choice becomes deterministic and the alternative which dominates the other one is chosen with probability 1. In the case of three alternatives, the induced choice probabilities are as follows:

$$\rho_x(X) = \frac{u(xy)}{u(X)} + \frac{u(z)}{u(X)} \frac{u(x)}{u(xy)}$$
$$\rho_y(X) = \frac{u(y)}{u(X)} \frac{u(z)}{u(xy)}$$
$$\rho_z(X) = 0$$

Thus, the worst alternative is chosen with zero probability, while other alternatives are chosen with strictly positive probability. In particular, the difference in probabilities of these alternatives can be written as follows:

$$\begin{aligned}\rho_x(X) - \rho_y(X) &= \frac{u(xy)}{u(X)} + \frac{u(z)}{u(X)} \cdot \frac{u(x) - u(y)}{u(x) + u(y)} \\ &\equiv p^u_{xy}(X) + p^u_z(X) \cdot \sigma^{bgs}(x,y)\end{aligned}$$

The last line says that the difference of choice probabilities are an expression in terms of choice probabilities of the Luce form and salience function (σ) defined by Bordalo et al [3] (bgs).

Next, I am going to provide a deterministic version of the visual choice in order to make the procedural aspect of the model clearer, and after that start to analyze the visual choice.

3 Deterministic Visual Choice

In this section, I will focus for simplicity on choice functions that outputs a single alternative for each menu as the choice data. Assume that the salience value u induces an asymmetric complete ordering over the alternatives, denoted as $\succ_u$, and similarly for v, denoted $\succ_v$. The DM fixates on a unique alternative, which is the most salient alternative from the menu:

$$f(M) := \max(M, \succ_u)$$

where $f(M)$ denotes the fixated alternative in M. This alternative determines which alternatives are looked upon conditional on this, $\hat{s}(f_i(M),M)$. The deterministic visual choice can be defined as:

$$c(M) := \max(\hat{s}(f_i(M),M), \succ_v)$$

There are two cases to consider given the definition of visual choice:

- s is fixation-independent.
- s is fixation-dependent.

Assume s is fixation-independent. Then, one can view s as a *consideration set*. In the literature, there are several models that characterizes choices using different assumptions on s. Two relevant cases (which will become clear in the analysis of visual choice) are when s satisfies the *attention filter property (af)* and *weak axiom of revealed preference (warp)*. The former is defined and characterized by Masatlioglu et al. [36], which assumes that $\hat{s}(M) = \hat{s}(M \setminus \{x\})$ for

$x \notin \hat{s}(M)$. The latter property corresponds to the case of rational visual choice, and this means that $\hat{s}(x,M)$ is equivalent to $\max(M, \succsim)$ for a complete preference relation.

Assume s is fixation-dependent. Applying the satisficing visual choice to the deterministic setting, let $s(x,M) = D(x,M)$ given the partial order $\succeq$, and define the choice function as:

$$c(M) = \max(\hat{D}(f(M),M), \succ_v)$$

Assume that a choice reversal occurs when one removes an alternative from M that is not chosen from it, i.e. $x \neq c(M)$ and $c(M) \neq c(M \setminus \{x\})$. This is only possible if x is itself the fixation point. To see, note that if x is not the fixation point, then it can change the choice only if it is one of the saccaded alternatives. If there is a unique saccaded alternative, then it is equal to the chosen alternative, so this cannot be the case. So, $|\hat{D}(x,M)| > 1$, and by the definition of D, $\hat{D}(x,M) \setminus \{y\} = \hat{D}(x,M \setminus \{y\})$. Hence, the chosen alternative cannot change unless y is itself the fixation point. Let $f^R(M)$ be the revealed fixation point in menu M and $\succ_u^R$ denote the revealed preferences according to u. Thus:

$$x \neq c(M) \neq c(M \setminus \{x\}) \rightarrow x = f^R(M) \;\;\&\;\; x \succ_u^R z \quad \forall z \in M \setminus \{x\}$$

By rationality of the visual choice, x will be the fixation point in any menu M' that is a subset of M and $x \in M'$. Any z except x cannot be the fixation point in M' provided that $x \in M'$. If one furthermore knows that $z \neq c(M')$, then removal of z from M' cannot change the choice of the DM.[11] This shows the necessity of the following condition.

Condition 1. *Let M, M' be two menus s.t. $\{x, c(M)\} \subseteq M$ and $x \in M'$. If $c(M) \neq c(M \setminus \{x\})$ and $x \neq c(M)$, then $c(M') = c(M' \setminus \{z\})$ for all $z \in M \setminus \{x, c(M')\}$.*

If there is no choice reversal, then Condition 1 implies that $c(M) = c(M')$ for any $M' \subseteq M$ with $c(M) \in M'$. If this is the case, let $c(M) \succ_u^R x$ for $x \neq c(M)$ and $x \in M$. Thus, $f^R(M) = c(M)$ in this case.

The previous part provided the identification strategy for the fixation point and hence the revealed u-order. For the identification of the saccade correspondence, one can rely on this and the information from binary menus. Assume that $x = f^R(M)$ for some $M \supseteq \{x,y\}$ and $c(xy) = y$. Because of rationality, x is also the fixated point in $\{x,y\}$, and therefore $c(\{x,y\}) = y$ implies that $y \succeq^R x$ where $\succeq^R$ denotes the revealed dominance relation. Using this, one can construct:

$$D^R(M) = \{x \in M \setminus \{f^R(M)\} : c(\{x, f^R(M)\}) = x\}$$

From $D^R(M)$, one can define the induced correspondence $\hat{D}^R(M)$. The following is a necessary condition which guarantees that the revealed dominance relation is a partial order.

Condition 2. *Let $x_1, \dots, x_k$ be a sequence of alternatives such that $f^R(\{x_i, x_{i+1}\}) = x_{i+1}$ for all $i \leq k-1$. If $c(\{x_i, x_{i+1}\}) = x_i$ for all $i \leq k-1$, then $c(\{x_1, x_k\}) = x_1$.*

Finally, let $\succ_v^R$ be the revealed v-order and $x \succ_v^R y$ if $x, y \in D^R(z, M)$ for some $M \supseteq \{x,y,z\}$ and $x = c(M)$. The following is implied by the acyclicity of the v-order.

Condition 3. *Let $M_1, M_2, \dots, M_k$ be a sequence of menus such that $\{x_i, x_{i+1}\} \subseteq M_i$ with $x_1, \dots, x_k$ being distinct and $x_{k+1} \equiv x_1$. If $x_{i+1} \in D^R(M_i)$ for $i \leq k-1$ and $c(M_i) = x_i$ for $i \leq k$, then $x_1 \notin D^R(M_k)$.*

[11]Note that allowing for multiple saccaded alternatives only change the statement of the condition without making a change in the identification.

Theorem 1. *A choice function c is deterministic rational visual choice if and only if Conditions 1-3 hold.*

A special case occurs if one assumes the individual saccades at most to one point, and the choice rule becomes:

$$c(M) = \max(\max(M, \succ_{f(M)}), \succ_v) = \max(M, \succ_{f(M)})$$

This special case coincides with *triggered choice* of Rubinstein and Salant [44], *choice by association* of Demirkan [10], *choice by salience* of Giarlotta et al. [18] and *conspicuity-based reference point* representation of Kibris et al. [39]. This choice rule can be characterized by using an axiom that only allows a single choice reversal in a menu, and that reversal can only occur if the fixation point (the most salient point) is dropped from the menu (except the chosen alternative). Formally, this condition is read as follows: The above stated result with single saccade follows as a corollary. When the DM performs only one saccade, then $\hat{D}^R(M) = \{c(M)\}$ for any menu M. Thus, the Condition 3 is trivially satisfied when the DM performs only one saccade, because there can be no x_{i+1} in $D^R(M_i)$ since $x_{i+1} \neq c(M)$. Similarly, Condition 2 is trivially satisfied because the fixation point should be equal to the chosen alternative. Therefore, one can conclude the following.

Corollary 1. *A choice function c is deterministic rational visual choice with single saccade if and only if Condition 1 holds.*

Finally, note that if $\succeq$ is complete, then the DM always saccades to the set of maximal alternatives in the menu with respect to $\succeq$. Because c is a choice function, $c(M) \in \max(M, \succeq)$ when $\succeq$ is complete. If $\succeq$ is not complete, then the DM does not necessarily choose an alternative in $MAX(M, \succeq)$. A condition that guarantees this is satisfied is when top-down preferences affect the bottom-up salience functions. For this, assume that the order induced by u (strictly) preserves the order of $\succeq$, that is, if $x \succeq y$, then $x \succ_u y$. This implies that even if $k = 1$, the first fixated alternative is an alternative in $MAX(M, \succeq)$. Therefore, $c(M) \in MAX(M, \succeq)$.

4 Fixation Independent Visual Choice

In this section, I will assume that the saccade correspondence is fixation-independent, and show that in this case one can connect visual choice to some models and their characterization results in the literature. Furthermore, I assume that s is nonempty. Still, note that if $s(M) = \emptyset$ for all M, then the DM does not perform any saccades, then this implies that conditional on fixation at a certain alternative x, the DM chooses this without comparing it to any other alternative. So, the choice probabilities become:

$$\rho_x(M) = \frac{u(x)}{u(M)}$$

for all $x \in M$ and any menu M. Thus, all alternatives are chosen with probability equal to their relative salience values with respect to u. The other extreme case happens when $s(x, M) = M$ for all M and $x \in M$, which is allowed here. The DM performs saccades to all alternatives available in the menu, which results in:

$$\rho_x(M) = \frac{v(x)}{v(M)}$$

for all $x \in M$ and any menu M. This means that all alternatives are chosen with probability equal to their relative salience values with respect to v, the function that determines saccade values. So, the extreme cases (the former being excluded) of this induce two different Luce rules. Luce [32] characterized this using two well-known conditions in the literature: *positivity* and *Luce's IIA*.

Definition 3. *For any rcf ρ:*

- *ρ satisfies positivity if $\rho_x(M) > 0$ for any menu M and $x \in M$.*
- *ρ satisfies Luce's IIA if $\frac{\rho_x(M)}{\rho_y(M)} = \frac{\rho_x(M')}{\rho_y(M')}$ for any M, M' such that $x, y \in M \cap M'$.*

Both of these cases satisfy warp. In general, if the saccade correspondence is *rational* in this sense, then one can find a complete and transitive binary relation $\succsim$ such that $s(M) = \max(M, \succsim)$. Under the assumption that s is nonempty-valued, this is equivalent to the *preference-oriented Luce rule (polr)* discussed by Dogan and Yildiz [11]. They characterize this rule using a simple condition called *odds modularity*, where odds is defined as $o_x(M) := \frac{1-\rho_x(M)}{\rho_x(M)}$.

Definition 4. *ρ satisfies odds modularity if $o_x(M) + o_x(M') = o_x(M \cup M')$ for any menu M, M'.*

In general, visual choice reduces to the following choice rule when s is fixation-independent:

$$\rho_x(M) = \begin{cases} \frac{v(x)}{\sum_{y \in s(M)} v(y)} & if \quad x \in s(M) \\ \frac{u(x)}{u(M)} & if \quad s(M) = \emptyset \\ 0 & o.w. \end{cases}$$

This choice rule generalizes the *general Luce model (glm)* defined by Echenique and Saito [13] and *limited consideration Luce model (lcl)* by Ahumada and Ulku [1], which are introduced in order to generalize the Luce rule to deal with zero-probability choices. Since s is nonempty, the fixation-independent visual choice coincides with these models. Echenique and Saito provides a characterization of this model using the cyclical independence condition, which is defined next.

Definition 5. *(Cylical Independence) Consider a sequence of alternatives $x_1, \ldots, x_n$ in X. If there exists a sequence of menus $M_1, \ldots, M_n$ such that $\rho_{x_i}(M_i), \rho_{x_i}(M_{i+1}) > 0$ for all $i \in \{1, \ldots, n\}$, then:*

$$\frac{\rho_{x_1}(M_n)}{\rho_{x_n}(M_n)} = \frac{\rho_{x_1}(M_1)}{\rho_{x_2}(M_1)} \frac{\rho_{x_2}(M_2)}{\rho_{x_3}(M_2)} \cdots \frac{\rho_{x_{n-1}}(M_{n-1})}{\rho_{x_n}(M_{n-1})}$$

Using the characterization theorems and these observations, one can directly state the following result for the fixation-independence case with s being nonempty.

Proposition 1. *Consider any rcf ρ.*

- *ρ satisfies cyclical independence iff ρ is a visual choice such that s is nonempty and fixation-independent.*
- *ρ satisfies odds modularity iff ρ is a visual choice such that s is nonempty, fixation-independent and satisfies warp.*

5 Fixation Dependent Visual Choice

Now, I allow the saccades to be fixation dependent. I already defined the special choice rule I am going to use for this case, called *satisficing visual choice*. I assume that the DM *only* chooses from alternatives that is better with respect to the fixation point, and potentially chooses all such alternatives with positive probability because she is not able to find a unique maximum among them due to cognitive limitations. From a neuroscience perspective, the fixation point leads to a certain value calculation in the brain, and any alternative that passes this threshold available in the menu is saccaded after fixating at some alternative. So, this special case proposes a certain type of saccade correspondence that relies on the notion of satisficing. The brain is able to compute the value of the alternatives in the scene in a very fast manner, and eliminates using the saccade correspondence any alternative that does not dominate the fixation point. Alternatively, this can be interpreted as if the DM saccades to all alternatives in the environment, but only considers that are at least good as the fixation point. So, in this sense, one can view this choice rule as a search over all of the set of alternatives because the DM fixates on all alternatives, while the final choice of the DM depends on how much alternatives are satisficing with respect to the fixation points. To sum up, the satisficing visual choice model combines different aspects of bounded rational choice behavior and cognitive limitations. The visual process implies that individuals fixate on certain alternatives that depend on their relative salience values, and they perform saccades conditional on the fixation point. One of the most well-established bounded rationality procedures is satisficing, which is incorporated here by assuming that the individual saccades to only alternatives that are at least good as the fixation point with a random tie-breaking among these alternatives. These two parts determine the final choice probabilities. The next section is devoted to understanding the rationality properties of this procedure, and then I will provide an empirical characterization.

5.1 Rationality of SVC

I start by discussing the rationality of satisficing visual choice. Although the literature on rationality of deterministic choice correspondences is quite developed, the evaluation of random choice functions from this perspective is not substantial as the former. First, one can discuss whether *regularity*, a much discussed property satisfied by the *random utility model*, is satisfied by svc. I will look at the contextual effects that can be explained by svc, which constitute examples that violate regularity, so svc allows for this type of violation. Another much discussed property of rationality for random choice is the notion of *stochastic transitivity*. I will show that svc also allows violation of even *weak stochastic transitivity*. However, under a certain condition on the relation between preferences and visual value functions, svc satisfies *moderate stochastic transitivity*[12], while it continues to violate *strong stochastic transitivity* even under stronger conditions. This is plausible given the evidence that strong stochastic transitivity is frequently violated in experiments (see Rieskamp [42]). A recent work that deals extensively with measuring the rationality of random choice functions is Ok and Tserenjigmid [38]. I am going to use this stochastic rationality measure developed by Ok and Tserenjigmid to show two results regarding the rationality of svc. Finally, I will conclude by discussing how svc fits within the deliberate randomization framework.

The results in this section rely on the relation between the preferences of the DM and the

[12]This is analyzed in detail by Natenzon and He [21] who used it to characterize *moderate expected utility* building on the earlier work.

visual value functions, that is, the relation between $\succeq$ and (u,v). For this, I need the following definition.

Definition 6.

- *A function u is said to be compatible with $\succeq$ if $x \succeq (\succ) y$ implies $u(x) \geq (>) u(y)$.*
- *(u,v) is said to be strongly compatible if both u and v are compatible and $x \bowtie y$ implies either $u(x) > u(y)$ & $v(x) < v(y)$ or $u(y) > u(x)$ & $v(y) < v(x)$.*

5.1.1 Regularity and Contextual Effects

Satisficing visual choice takes a very simple form when the DM faces binary menus. Consider an arbitrary menu $\{x,y\}$ such that $x \neq y$. Either these two alternatives are comparable or they are not. In the former, if say $x \succ y$ wlog, then DM chooses x conditional on a fixation at x making no saccades, and again chooses x by doing a saccade from the fixation at y. Thus, $\rho_x(xy) = 1$ if $x \succ y$, and $\rho_y(xy) = 0$. So, the choice becomes deterministic when one of the alternatives is strictly better than the other. On the other hand, if $x \bowtie y$, then no saccades are performed from the fixation point, implying that:

$$\rho_x(xy) = \frac{u(x)}{u(xy)} \;\;\&\;\; \rho_y(xy) = \frac{u(y)}{u(xy)}$$

Thus, in binary choice problems, decisiveness about the choice problem implies deterministic choices, while indecisiveness implies interior choice probabilities and hence random choice. Adding a third alternative to the menu enables to capture more relevant behavioral phenomena. Next, I present two examples to show that svc can capture some important behavioral phenomena in this case. Both examples are related to contextual effects resulting in the violations of regularity, which says that the choice probability of an alternative decreases as one moves to larger sets.

Example 3. (Attraction Effect)
Let $X = \{x,y,z\}$. Assume that $x \succ z$ with alternatives being incomparable otherwise. The attraction effect arises when the choice problems of the DM are $\{x,y\}$ and $\{x,y,z\}$. More precisely, attraction effect is the observation that the choice probability of x increases relative to y when a third option which is only dominated by x is added to the menu. In choice problem $\{x,y\}$, $\rho_s(xy) = \frac{u(s)}{u(xy)}$ for any $s \in \{x,y\}$. Thus, choice probabilities of the products depend on the probability that the DM fixates on the alternatives. In particular, if $u(x) = u(y)$, then both alternatives are chosen with probability $\frac{1}{2}$. The choice probabilities in $\{x,y,z\}$ are given as follows:

$$\rho_x(xyz) = \frac{u(xz)}{u(xyz)}, \quad \rho_y(xyz) = \frac{u(y)}{u(xyz)}, \quad \rho_z(xyz) = 0$$

Observe that x's choice probability increases although the menu is now larger, which shows that attraction effect can be accomodated. More generally, the attraction effect holds only when $y \succ x$ or $x \bowtie y$, because otherwise $\rho_x(xy) = 1$. In the former situation, it is straightforward because $\rho_x(xy) = 0$ and $\rho_x(xyz) > 0$. It is also evident when $x \succ z$ but $y \bowtie z$ as in the above example. On the other hand, whether attraction effect works for x when x and y both dominate z depends on a further comparison. In particular, the attraction effect works only for one of the alternatives, and it is true for x meaning that $\rho_x(xyz) > \rho_x(xy)$ iff:

$$\frac{v(x)}{v(xy)} > \frac{u(x)}{u(xy)}$$

showing that the attraction effect works for the alternative which has a higher relative v value compared to u. In the feature spaces, this is equivalent to:

$$\frac{\sum_i \alpha_i x_i}{\sum_i \beta_i x_i} < \frac{\sum_i \alpha_i (x_i + y_i)}{\sum_i \beta_i (x_i + y_i)}$$

Assume that $k = 2$, i.e. there are two features. Then, the attraction effects works for x when:

$$\rho_x(xyz) > \rho_x(xy) \quad when \quad \begin{cases} \frac{\alpha_1}{\alpha_2} < \frac{\beta_1}{\beta_2} & if \quad \frac{x_1}{x_2} > \frac{y_1}{y_2} \\ \frac{\alpha_1}{\alpha_2} > \frac{\beta_1}{\beta_2} & if \quad \frac{x_1}{x_2} < \frac{y_1}{y_2} \end{cases}$$

and both stay the same when $\frac{x_1}{x_2} = \frac{y_1}{y_2}$.

Example 4. (Compromise Effect)

Consider the same setting as in the previous example, but now assume that $x \succ y \succ z$. The compromise effect arises when the choice probability of y increases by the introduction of product z to the menu $\{x, y\}$. Observe that $\rho_x(xy) = 1$ and $\rho_y(xy) = 0$. On the other hand, in the larger choice problem $\{x, y, z\}$:

$$\rho_x(xyz) = 1 - \frac{u(z)}{u(xyz)} \cdot \frac{v(y)}{v(xy)}, \quad \rho_y(xyz) = \frac{u(z)}{u(xyz)} \cdot \frac{v(y)}{v(xy)}, \quad \rho_z(xyz) = 0$$

Observe that y violates the regularity property after z is added. The increase in the choice probability is due to the possibility that the DM can first fixate on z, and then saccade to y. Note that compromise effect holds always, but one can show that the impact of the compromise effect changes as one adds more options that are dominated by y. To see, let t be a fourth alternative that is also dominated by z, so $x \succ y \succ z \succ t$. The change in the choice probability of y in these two situations are given as follows:

$$\begin{aligned} \rho_y(xyz) - \rho_y(xy) &= \frac{u(z)}{u(xyz)} \cdot \frac{v(y)}{v(xy)} \\ \rho_y(xyzt) - \rho_y(xyz) &= \left(\frac{u(z)}{u(xyzt)} - \frac{u(z)}{u(xyz)} \right) \frac{v(y)}{v(xy)} + \frac{u(t)}{u(xyzt)} \cdot \frac{v(y)}{v(xyz)} \end{aligned}$$

First, one can show that the compromise effect in the latter case holds iff:

$$\frac{v(xy)}{v(xyz)} \geq \frac{u(z)}{u(xyz)}$$

Thus, if this condition is violated, then there is no compromise effect for y. Assume that this condition holds. For simplicity, let $u = v$. Then, the compromise effect decreases iff:

$$\frac{u(z)}{u(t)} \geq \frac{u(xy) - u(z)}{u(xyzt)}$$

I conclude the discussion of regularity by providing a simple example which shows that regularity violation is possible even in the simplest case. This further illustrates that regularity violation is characteristic of satisficing visual choice.

Example 5. (Uniform SVC)
Assume that $\succeq$ is complete and u, v values are equal for all alternatives. Recall that the choice probabilities in the uniform case with satisficing visual choice is given by:

$$\rho^{k(M)}(M) = \frac{1}{|M|}\sum_{j<k}\frac{1}{|M|-j}$$

if $k(M) \leq |M|-1$, and $\rho_{|M|(M)}(M) = \rho^{|M|-1}(M) + \frac{2}{|M|}$.

Adding an alternative that dominates x never increases the choice probability of x. Hence, consider adding an alternative that is dominated by x, say y. Assuming that $x = k(M)$ and $y = l(M)$ for some $l < k$ where $k < |M|$ (so it is not the best alternative in the menu), the change in choice probability of x is given by:

$$\begin{aligned}\rho^{k(M)+1}(M \cup \{y\}) - \rho^{k(M)}(M) &= \left[(\frac{1}{|M|+1} - \frac{1}{|M|})\sum_{j<k}\frac{1}{|M|-j}\right] + \frac{1}{|M|+1}\frac{1}{|M|} \\ &= \frac{1}{|M|+1}\frac{1}{|M|}\left(1 - \sum_{j<k}\frac{1}{|M|-j}\right)\end{aligned}$$

Thus, the violation of regularity occurs iff:

$$1 > \sum_{j<k}\frac{1}{|M|-j}$$

Note that the right-hand side of this expression increases as the alternative becomes better, that is, as k increases. This implies that violating regularity becomes harder as the alternative's ranking becomes higher. In fact, the regularity is never violated by the best alternative in M. To see, note that the change in probability is slightly different:

$$\rho^{|M|+1}(M \cup \{y\}) - \rho^{|M|}(M) = -\frac{1}{|M|+1}\frac{1}{|M|}\left(\sum_{j<k}\frac{1}{|M|-j} + 1\right)$$

which is always negative, and hence regularity cannot be violated.

5.1.2 Stochastic Transitivity

Transitivity is seen as one of the main tenets of rational choice behavior. In the case of stochastic choices, this property is translated as *stochastic transitivity*. Let $\geq_\rho$ be the *stochastic preferences* defined as $x \geq_\rho y$ if $\rho_x(xy) \geq 1/2$. Stochastic transitivity comes in different forms. The following definition provides three versions used in the literature.

Definition 7. *Assume that $x \geq_\rho y \geq_\rho z$. ρ is:*

- *weak stochastic transitive (WST) if $\rho_x(xz) \geq \frac{1}{2}$.*
- *moderate stochastic transitive (MST) if $\rho_x(xz) \geq min\{\rho_x(xy), \rho_y(yz)\}$.*
- *strong stochastic transitive (SST) if $\rho_x(xz) \geq max\{\rho_x(xy), \rho_y(yz)\}$.*

Proposition 2. *Satisficing visual choice does not necessarily satisfy weak stochastic transitivity. However, if u is compatible, then moderate stochastic transitivity satisfied, while strong stochastic transitivity can be violated even if (u,v) is strongly compatible.*

Proof. Assume that $x \geq_\rho y \geq_\rho z$. First, let me show that even wst can be violated without any compatibility assumption. To see, let $x \bowtie y \bowtie z$ but $z \succ x$. Then, $\rho_x(xz) = 0 < \frac{1}{2} \leq \min\{\rho_x(xy), \rho_y(yz)\}$ since u is strictly positive. Observe that this is ruled out when u is compatible with $\succeq$: $\rho_x(xy) \geq \frac{1}{2}$ implies that $u(x) > u(y)$, and similarly $\rho_y(yz) \geq \frac{1}{2}$ implies $u(y) > u(z)$, which results in $u(x) > u(z)$. This contradicts the compatibility since $z \succ x$.

Now assume that u is compatible. The goal is to show that $\rho_x(xz) \geq min\{\rho_x(xy), \rho_y(yz)\}$. Note that $x \geq_\rho y$ implies that either $x \succ y$ or $x \bowtie y$. Similarly, $y \geq_\rho z$ implies either $y \succ z$ or $y \bowtie z$. If $x \succ z$, then the conclusion follows directly in any case, because $\rho_x(xz) = 1$. So, I will assume that $x \bowtie z$.

First, assume that $x \succ y$, which implies by compatibility $u(x) > u(y)$. If $y \succ z$, then transitivity implies $x \succ z$, which is excluded above. So, let $y \bowtie z$. Because $y \geq_\rho z$, it follows that $u(y) \geq u(z)$, and hence $u(x) > u(z)$. Since $x \bowtie z$ and $u(x) > u(y)$:

$$\rho_x(xz) = \frac{u(x)}{u(xz)} > \frac{u(y)}{u(yz)} = \rho_y(yz)$$

This concludes the proof for $x \succ y$.

Now assume that $x \bowtie y$. If $y \succ z$, then $\rho_y(yz) = 1$ and also $u(y) > u(z)$ by compatibility. Moreover, the incomparability of x and y and the assumption that $x \geq_\rho y$ implies that $u(x) \geq u(y)$ since $\rho_x(xy) \geq \frac{1}{2}$, and hence $u(x) > u(z)$. Since $x \bowtie z$ and $u(y) > u(z)$:

$$\rho_x(xz) = \frac{u(x)}{u(xz)} > \frac{u(x)}{u(xy)} = \rho_x(xy)$$

Finally, consider the case when $y \bowtie z$. The assumption that $\rho_x(xy) \geq \frac{1}{2}$ and $\rho_y(yz) \geq \frac{1}{2}$ implies that $u(x) \geq u(y) \geq u(z)$. So, $\rho_x(xz)$ is in fact greater than both $\rho_x(xy)$ and $\rho_y(yz)$, implying the conclusion.

I conclude the proof by showing that SST is not satisfied even when (u,v) is strongly compatible. To wit, assume $x \succ y \bowtie z$ and $x \bowtie z$. Then, $\max\{\rho_x(xy), \rho_y(yz)\} = 1$, and $\rho_x(xz) = \frac{u(x)}{u(xz)} < 1$, which shows that SST is violated.

$\square$

5.1.3 Stochastic Rationality Measure

Towards developing the stochastic rationality measure developed by Ok and Tserenjigmid, I need to provide some definitions first. Consider any random choice function ρ. The λ*-choice correspondence* induced by ρ is a mapping $C_{\rho,\lambda} : \mathbb{X} \to \mathbb{X}$ such that:

$$C_{\rho,\lambda}(M) := \begin{cases} \{x \in M : \rho^*_x(M) \geq \lambda\} & \text{if} \quad \lambda \in (0,1] \\ \{x \in M : \rho_x(M) > 0\} & \text{if} \quad \lambda = 0 \end{cases}$$

where ρ^* is defined as:

$$\rho^*_x(M) := \frac{\rho_x(M)}{\max_{y \in M} \rho_y(M)}$$

The collection $\{C_{\rho,\lambda} : \lambda \in [0,1]\}$ is referred as *Fishburn family associated with* ρ. ρ is said to be λ-rational if $C_{\rho,\lambda}$ is rational. Thus, showing λ-rationality is violated is equivalent to the violation of the characterizing features for the deterministic rationality. These are stated as follows for any deterministic choice function C:

- (*Chernoff*): $\forall S,T \in \mathbb{X}$ such that $S \subseteq T$: $C(T) \cap S \subseteq C(S)$.
- (*Condorcet*): $\forall S \in \mathbb{X}$ and $x \in S$: $x \in C(xy)$ $\forall y \in S$ implies that $x \in C(S)$.
- (*No Cycle*): $\forall x,y,z \in X$: $x = C(xy)$ and $y = C(yz)$ implies $x = C(xz)$.

OT shows that a λ-rational random choice function can be characterized by the λ-stochastic analogues of the conditions which characterize a rational deterministic choice correspondence. These are stated as follows:

- (λ*-Chernoff*): $\forall S,T \in \mathbb{X}$ such that $S \subseteq T$: $\rho_x^*(T) \geq \lambda$ implies $\rho_x^*(S) \geq \lambda$.
- (λ*-Condorcet*): $\forall S \in \mathbb{X}$ and $x \in S$: $\rho_x^*(xy) \geq \lambda$ $\forall y \in S$ implies that $\rho_x^*(S) \geq \lambda$.
- (λ*-No Cycle*): $\forall x,y,z \in X$: $\lambda \rho_x(xy) > \rho_y(xy)$ and $\lambda \rho_y(yz) > \rho_z(yz)$ implies $\lambda \rho_x(xz) > \rho_z(xz)$.

Consider two arbitrary random choice functions $\rho, \tilde{\rho}$. The *comparative rationality ordering*, a partial order denoted as $\trianglerighteq_r$, is defined as follows: $\rho \trianglerighteq_r \tilde{\rho}$ iff ρ is λ-rational for all $\lambda \in [0,1]$ whenever $\tilde{\rho}$ is λ-rational. A rcf ρ is *maximally rational* if $\rho \trianglerighteq_r \tilde{\rho}$ for any rcf $\tilde{\rho}$, and similarly *minimally rational* if $\tilde{\rho} \trianglerighteq_r \rho$ for any rcf $\tilde{\rho}$.

The svc rule has the primitives $(\succeq, u, v)$. The following proposition shows that if one knows that $\succeq$ is complete, then any svc can be compared (by varying u or v) with respect to the rationality ordering defined here. Furthermore, no svc with complete $\succeq$ is maximally or minimally rational.

Theorem 2. *If $\succeq$ is complete and (u,v) is compatible, then:*

- *Any svc is $\trianglerighteq_r$-comparable.*
- *No svc is maximally or minimally rational.*

Next, I am going to characterize the cases svc is maximally rational. Say a rcf ρ is an *almost Luce rule (alr)* if there is unique alternative x^* such that:

$$\rho_y^{alr}(M) := \begin{cases} 1 & x^* \in M \ \& \ y = x^* \\ 0 & x^* \in M \ \& \ y \neq x^* \\ \frac{u(y)}{u(M)} & x^* \notin M \end{cases}$$

So, alr differs from the Luce rule only when the menu contains x, becoming deterministic when this is the case. The following result shows that svc is maximally rational iff it is either Luce or an almost Luce.

Theorem 3. *If (u,v) is strongly compatible, then svc is maximally rational iff one of the following holds:*

- *All alternatives are incomparable (Luce rule).*

- *There is unique x^* such that $x^* \succ y$ for all $y \neq x$ and otherwise all alternatives are incomparable (almost Luce rule).*

The rationality of svc is not monotonic with respect to the completeness of the preferences. Above, I showed that when $\succeq$ is complete, svc is not maximally rational, while it is maximally rational in the case of fully incomparability or when the DM has a favourite alternative which dominates the rest without any other comparable pair. These show that increasing the level of decisiveness is not necessarily better for the DM from the viewpoint of rationality.

Remark 1. *Ok and Tserenjigmid shows that if the menus are restricted to be binary, as it is usually the case in experimental applications, then moderate stochastic transitivity implies maximal rationality. This implies that svc is maximal rational because svc satisfies this by Proposition 2 when u is compatible.*

Remark 2. *If one allows $\succeq$ to include ties among distinct alternatives, then also the case of $\succeq = X \times X$ is maximally rational. When this is the case, the choice probabilities are as follows:*

$$\begin{aligned}\rho_y(M) &= \sum_{x\in M} \frac{v(y)}{u(M)} \cdot \frac{u(x)}{v(M)} \\ &= \frac{v(y)}{u(M)v(M)} \cdot \sum_{x\in M} u(x) \\ &= \frac{v(y)}{v(M)}\end{aligned}$$

If furthermore v is compatible, then $\rho_y(M) = \frac{1}{|M|}$.

In fact, if $\succeq$ consists of either incomparabilities or indifferences, then this holds true. Let $I(x,M)$ denote the set of indifferent alternatives to x in menu M. Assume that the conditional choice probability due to saccades are modified as:

$$\mathbf{p}^s_{(M,x)}(y) = \begin{cases} \frac{v(x)}{v(D(y,M))} & \text{if} \quad x \in D(y,M) \\ \frac{v(x)}{v(I(y,M))} & \text{if} \quad D(y,M) = \emptyset \;\; \& \;\; x \in I(y,M) \\ 0 & o.w. \end{cases}$$

Since preferences consists of either incomparabilities or indifferences, one can show that svc is a Luce rule with u, using the compatibility of v. The choice probability of x in M is given as follows:

$$\rho_x(M) = \begin{cases} \frac{u(I(x,M))}{u(M)} \cdot \frac{v(x)}{v(I(x,M))} & \text{if} \quad I(x,M) \neq \emptyset \\ \frac{u(x)}{u(M)} & o.w. \end{cases}$$

The compatibility of both u and v with $\succeq$ implies that $u(x) = u(y)$ and $v(x) = v(y)$ for any $y \in I(x,M)$. Thus, $\frac{v(x)}{v(I(x,M))} = \frac{1}{|I(x,M)|}$ and $u(I(x,M)) = |I(x,M)|u(x)$, which implies that the above choice probabilities reduces to $\rho_x(M) = \frac{u(x)}{u(M)}$. This shows that svc is also maximally rational in this case.

5.1.4 Deliberate Randomization

In the previous section, I discussed the rationality properties of satisficing visual choice using stochastic transitivity and the rationality index developed by Ok and Tserenjigmid. A related inquiry is to evaluate svc from the viewpoint of *deliberate randomization*. This approach views the random choices of the DM as the result of optimizing expected utility. Let $\Delta(M)$ be the set of probability distributions on M. Fudenberg et al. [17] provides a model of random choice which solves a maximization problem with perturbed expected utility. This distribution is found by solving the following maximization problem, and a rcf ρ is a member of this whenever:

$$\rho(M) \in \underset{\rho\in\Delta(M)}{\operatorname{argmax}} \sum_{x\in M} u^*(x)\rho_x(M) - k_{M,x}(\rho_x(M))$$

where u^* is the utility function of the DM, and k is a convex perturbation function. When $\rho(M)$ can be found using this maximization problem, it is said to have an *additive perturbed utility (apu)* representation. This version where the cost function depends on both the alternative x and menu M is empirically vacuous. On the other hand, the versions where k is menu-independent ($k_{A,x}(\cdot) = k_x(\cdot)$), called *menu-independent apu*, and in addition alternative independent ($k_{A,x}(\cdot) = k(\cdot)$), called *invariant apu*, satisfy regularity, and therefore does not nest svc. Therefore, svc cannot be represented using these type of cost functions. However, assume that the cost function is alternative-independent but depends on the menu, called *item-independent apu*, so $k_{A,x}(\cdot) = k_A(\cdot)$. The next result shows that svc can be represented using an item-independent apu under certain conditions.

Proposition 3. *If $\succeq$ is complete and v is compatible, then svc can be represented by an alternative-independent apu.*

Proof. The alternative-independent apu is characterized using a condition called *item-acyclicity*. Let $x \succsim_i y$ if $\rho_x(M) \geq \rho_y(M)$ for some menu $M \supseteq \{x,y\}$ and $x \sim_i y$ if both $x \succsim_i y$ and $y \succsim_i x$. Item-acyclicity says that $\succsim_i$ should be acyclic. First, I will show that the compatibility of v implies monotonicity, which implies when $\succeq$ is complete that $\rho_x(M) \geq \rho_y(M)$ iff $\rho_x(M') \geq \rho_y(M')$ for any $M \cap M' \ni x, y$. To see that monotonicity is satisfied, consider $x, y \in X$ that are comparable, and assume wlog that $x \succ y$. For any $z \in M$ such that $y \succ z$, it is the case that $x \succ z$ by transitivity. Furthermore, compatibility assumption implies that $v(x) > v(y)$. This shows that the contribution to x from the items it dominates is at least the contribution y gets similarly, which concludes the proof.

$\square$

Another paper that lays out a model of deliberate randomization is= Cerreia-Vioglio et al. [8]. Assume that $X = [0,1]$ and Δ is the set of lotteries on X.[13] Define $\rho^*(M) := \sum_{x\in M} \rho_x(M) \cdot x$, and let $\gg$ denote the *first-order stochastic dominance relation*. Next, I provide the definition for deliberate random choice:

Definition 8. *A rcf ρ is called a deliberate random choice (drc) if there is a complete preorder $\succsim$ on Δ such that:*

- *For all $M \in \Delta$: $\rho^*(M) \succsim p$ for any $p \in co(M)$.*
- *If $p \gg q$, then $p \succ q$.*

[13] The only change is that instead of an arbitrary interval in $\mathbb{R}$ as in their model, I assume X is the unit interval from 0 to 1. Also, they assume the topology of weak convergence on Δ.

Deliberate random choice does not necessarily satisfy regularity. Still, one can show that svc is not the result of deliberate randomization according to this definition. It is shown that drc is characterized by the following condition called *rational hedging*:

Definition 9. *Consider a collection of menus $M_1,\dots,M_k$ such that $k \geq 2$. If $\rho^*(M_k) \in co(M_{k-1})$ for any k, then $q \in co(M_k)$ implies that $\neg(q \gg \rho^*(M_1))$.*

To see that svc is not drc, consider the set of degenerate lotteries which gives a prize $x \in [0,1]$ with certainty. In particular, let $\{x,z\} = M_2 \subset M_1 = \{x,y,z\}$ with $x < y < z$, and $\succeq = \geq$. Note that $co(M_1) = co(M_2) = [x,z]$ and svc procedure implies that $\rho_z(M_2) = 1$ while $\rho_y(M_1), \rho_z(M_1) > 0$. Thus, $\rho^*(M_2) = z$ and $\rho^*(M_1) < z$. This contradicts the rational hedging condition since $\rho^*(M_2) \in co(M_1) = [x,z]$ and $z \in co(M_2)$ should imply $\neg(z \gg \rho^*(M_1))$, but actually for any $q \in (\rho^*(M_1), z]$ one has $q \gg \rho^*(M_1)$.

5.2 Characterization of SVC

In this section, first I am going to talk about the identification of the primitives for svc, which is $<\succeq, u, v>$. Then, I will impose conditions on the choice data which provides an empirical characterization for satisficing visual choice. First, I start by identifying the preferences of the DM. For the identification of preferences, binary menus are sufficient.

Definition 10. *x is revealed to dominate y if $\rho_x(xy) = 1$ and revealed to be incomparable if $\rho_x(xy) \in (0,1)$.*

Furthermore, observe that even though the possibility of multiple representations is open, preferences are *uniquely identified* in any case. It cannot be the case that x is *revealed to dominate* y under some representation, while the opposite is true under another representation, because these imply $\rho_x(xy) = 1$ for the first representation, and $\rho_y(xy) = 1$ for the latter, impossible to observe in the choice data. Let me simply denote the revealed preferences as $\succeq$ with $\bowtie$ denoting the incomparable part (slightly abusing the notation by using the same notation as with the primitives), and define $D(y,M) = \{x \in M : x \succ y\}$.

The identification is more subtle for the fixation and saccade. Let g_{xy} denote *relative fixation value* of x compared to y. Note that this information cannot be identified always. For instance, assume that $\rho_x(M) = 1$ for some menu M such that $x \in M$. Under svc, this is possible iff x dominates the rest of the alternatives in M. In this case, one cannot identify the relative fixation value of x compared to the other alternatives in M. If this holds for all menus, then this rules out the possibility of identifying the relative fixation value of x compared to another alternative (which can be revealed possibly using other menus). However, if $\rho_x(xy) \in (0,1)$, then the relative fixation value of x to y can be derived by defining $g_{xy} = \frac{\rho_x(xy)}{\rho_y(xy)}$.

Now consider any trinary menu $\{x,y,z\}$. There are several cases that should be considered. If $\rho_s(M) > 0$ for all $s \in M$, then this implies that all alternatives are incomparable to each other, because if there was a comparable pair then the alternative that is worst according to comparison is chosen with zero probability. Therefore, one can define $g_{st} = \frac{\rho_s(xyz)}{\rho_t(xyz)}$ for any $s,t \in \{x,y,z\}$. Another case is when one of the alternatives is chosen with probability 1. I already noted that if either $\rho_x(xyz)$ or $\rho_y(xyz)$ is equal to 1, then the relative fixation values cannot be identified. However, if $\rho_z(xyz) = 1$, this implies that z dominates the rest of the alternatives, while x and y are revealed to be incomparable to each other, so $\rho_x(xy) \in (0,1)$

and same reasoning for binary menu applies. If $\rho_x(xyz), \rho_z(xyz) > 0$ and $\rho_y(xyz) = 0$, then y is dominated either by x or z. This has several subcases. If for example all binary probabilities are deterministic, then this means that $\succeq$ is complete when restricted to $\{x,y,z\}$, and fixation values cannot be identified. However, if $\rho_y(yz) \in (0,1)$ or $\rho_x(xz) \in (0,1)$, then one can define $g_{xy} = g_{xz} \cdot g_{zy}$ where $g_{xz} = \frac{\rho_x(xz)}{\rho_z(xz)}$ and $g_{zy} = \frac{\rho_z(yz)}{\rho_y(yz)}$. In general, one can find the relative fixation value of two alternatives using an *incomparability path*: an incomparability path can be defined as a sequence of alternatives $\{x_i\}_{i=1}^n$ such that $x_i \bowtie x_{i+1}$ for all $i \leq n-1$. Assume that $\{x_i\}_{i=1}^n$ is an incomparability path that connects x and y, that is, $x_1 \equiv x$ and $x_n \equiv y$. Define g_{xy} as follows:

$$g_{xy} := \Pi_{i \in \{1,\dots,n-1\}} \frac{\rho_{x_i}(x_i x_{i+1})}{\rho_{x_{i+1}}(x_i x_{i+1})}$$

where $\rho_{x_i}(x_i x_{i+1}) \in (0,1)$ for all $i \in \{1,\dots,n-1\}$. In the next section, I show that there exists a well-defined u such that $g_x(M)$ can be defined as $\frac{u(x)}{u(M)}$ using the cyclical independence condition stated in the upcoming part. Thus, one can define u and hence $g_x(M)$ using the choice probabilities, which I will use next in defining the saccade value and probabilities.

To identify the saccade probabilities, I will rely on the choice data and the fixation probabilities identified above. Let $\phi_x(y,M)$ denote the revealed saccade probability from y to x in menu M. One cannot always identify the saccade probabilities. Consider a binary menu $M = \{x,y\}$. If $\rho_x(xy) = 1$, then $x \succ y$, so $\phi_x(x,M) = \phi_x(y,M) = 1$ and $\phi_y(x,M) = \phi_y(y,M) = 0$. If $\rho_x(xy) \in (0,1)$, then they are incomparable, so $\phi_x(x,M) = 1 = \phi_y(y,M)$ and $\phi_x(y,M) = 0 = \phi_y(x,M)$. If one considers a trinary menu $M = \{x,y,z\}$, then this is more complicated. Two cases where it is easier are when one of the alternatives is chosen with probability 1 and the case when all alternatives chosen with positive probability. In the former case, if say $\rho_x(M) = 1$, then $\phi_x(y,M) = 1$ for any $y \in M$, while $\phi_s(t,M) = 0$ for any $s \neq x$ and $t \in M$. In the latter, $\phi_s(t,M) = 1$ iff $s = t$, and otherwise it is equal to 0.

In general, one can consider two cases. A menu M can either satisfy positivity or not, and in the latter case there always exists at least one alternative with zero choice probability. In the former case, it is easy to define saccade probabilities, since every alternative is revealed to be incomparable to each other. Let $\phi_x(y,M) = 1$ iff $x = y$ when this is the case, and it is equal to 0 otherwise. This is also true when x is the unique alternative that dominates y in M, without the need for the positivity, or when x is incomparable to the rest of the alternatives. Similarly, $\phi_x(y,M) = 0$ whenever $y \succ x$ for any menu that includes both. So:

$$\phi_x(y,M) = \begin{cases} 1 & if \quad x \succeq y \;\&\; D(y,M) \subseteq \{x\} \\ 0 & if \quad y \succ x \quad \text{or} \quad y \bowtie x \end{cases}$$

Observe that the only excluded case is when there are multiple alternatives that dominate y, i.e. when $|D(y,M)| > 1$. Assume that positivity is not satisfied, and let $\rho_z(M) = 0$. Define the odds ratio of g as $o_g(x,M) := \frac{1-g_x(M)}{g_x(M)}$ and the following function:

$$h_x(z,M) := (1 + o_g(z,M))\rho_x(M) - o_g(z,M)\rho_x(M \setminus \{z\})$$

Note that the final case covers the situations when either $z \succ x$ or $z \bowtie x$. The value function v can be constructed using the relative saccade probabilities. In particular, one can do this for alternatives x,y if there is z is such that $x,y \succ z$ and $\rho_z(M) = 0$ in some menu $M \supseteq \{x,y,z\}$. Then, one can define the relative saccade probability ϕ_{xy} of x to y as:

$$\phi_{xy} := \frac{h_{t_1}(z_1,M_1)}{h_{t_2}(z_1,M_1)} \cdot \frac{h_{t_2}(z_2,M_2)}{h_{t_3}(z_2,M_2)} \cdot \ldots \frac{h_{t_{k-1}}(z_{k-1},M_{k-1})}{h_{t_k}(z_{k-1},M_{k-1})}$$

From this, it is possible to derive the saccade value v and define the saccade probabilities as in the case of fixation probabilities. Since D is also revealed, one can define the saccade probabilities as:

$$\phi_x(z,M) = \begin{cases} \frac{v(x)}{v(\hat{D}(z,M))} & if \quad x \in \hat{D}(z,M) \\ 0 & o.w. \end{cases}$$

In the appendix, I show that this formula is equivalent to the following if furthermore $\rho_z(M) = 0$:

$$\phi_x(z,M) = \begin{cases} 1 & if \quad x \succeq z \ \& \ D(z,M) \subseteq \{x\} \\ h_x(z,M) & if \quad \{x\} \subset D(z,M) \\ 0 & o.w. \end{cases}$$

The previous part provided an analysis about the identification of underlying parameters of satisficing visual choice. Now, I will provide the conditions used to empirically characterize a choice data generated by the proposed choice rule. These conditions can be grouped into three categories. The first category consists of three conditions which are mainly related to the underlying preferences of the DM. The final condition of this group is a weakening of Luce's IIA. The second category consists of three conditions which are related to the revealed fixation probabilities, while the last category is a single condition related to the revealed saccade probabilities.

The first condition I am going to use to characterize svc is dominance transitivity. This condition is a relaxation of previously discussed moderate stochastic transitivity, which requires also the compatibility of u.

Condition 4. ***Dominance Transitivity***
$\rho_x(xy), \rho_y(yz) = 1$ *implies* $\rho_x(xz) = 1$ *for all* $x,y,z \in X$.

It is straightforward to see the necessity of this condition: If ρ is represented by svc and $\rho_x(xy) = 1$, then x strictly dominates y, and similary $\rho_y(yz) = 1$ implies y strictly dominates z. By transitivity, x strictly dominates z, which implies that $\rho_x(xz) = 1$.

Consider any alternative x which is chosen with positive probability in menu M. The following condition states that x cannot be dominated without dominating another alternative, unless it is incomparable to the rest of the alternatives.

Condition 5. ***Rationalizability***
$\rho_x(M) > 0$ *iff* $\max_{y\in M} \rho_x(xy) = 1$ *or* $min_{y\in M} \rho_x(xy) > 0$.

To see why this is necessary, first consider the if part. This says that if either $\max_{y\in M} \rho_x(xy) = 1$ or $min_{y\in M} \rho_x(xy) > 0$, then x is chosen from M with strictly positive probability. Assume that the former holds, which implies that $x \succ y$ for some $y \in M$, and therefore $\rho_x(M) > 0$ because it has svc representation. If the latter is true, then even in the case $\max_{y\in M} \rho_x(xy) < 1$, it is revealed that x is incomparable to the rest of the alternatives and it is chosen only when the DM fixates on it. Because ρ is svc and visual value functions are strictly positive, the conclusion again follows. Now consider the only if part, and take the contrapositive: if $\max_{y\in M} \rho_x(xy) < 1$ and $min_{y\in M} \rho_x(xy) = 0$, then $\rho_x(M) = 0$. The if part holds iff x does not dominate any other alternative, but it is dominated itself, which implies by ρ being svc that it should be chosen with zero probability.

Now consider a menu M and let $s \in M$ be an alternative which is revealed to be incomparable to $M \setminus \{s\}$, which holds iff $\rho_s(xs) \in (0,1)$ for all $x \in M \setminus \{s\}$. Consider the relative probability ratio of x compared to y in menu M where x, y are distinct from s, $\frac{\rho_x(M)}{\rho_y(M)}$. The removal of s does not affect the domination sets for the alternatives, and it only changes the denominator of the fixation probability for the alternatives dominated by the corresponding alternative. Since s is incomparable to any alternative in $M \setminus \{s\}$, the relative probability ratio remains the same. This is stated in the next axiom:

Condition 6. ***Independence from Incomparable Alternatives***
Consider a menu M and let $\rho_s(st) \in (0,1)$ for all $t \in M \setminus \{s\}$. Then, the following equality holds:

$$\frac{\rho_x(M)}{\rho_y(M)} = \frac{\rho_x(M \setminus \{s\})}{\rho_y(M \setminus \{s\})}$$

for any $x, y \in M \setminus \{s\}$ provided that the ratios are well-defined.

For the following condition, consider an incomparability path that connects y and z. Given that y and z are incomparable, one has $\frac{\rho_y(yz)}{\rho_z(yz)} = \frac{u(y)}{u(z)}$ in svc. Let $\{x_i\}_{i=1}^k$ be an incomparability path that connects y and z. Then, $\frac{\rho_{x_i}(x_i x_{i+1})}{\rho_{x_{i+1}}(x_i x_{i+1})} = \frac{u(x_i)}{u(x_{i+1})}$, which implies the following result.

Condition 7. ***Cyclical Independence***
Consider a sequence of alternatives $x_1, x_2, \ldots, x_k$ such that $\rho_{x_i}(x_i x_{i+1}) \in (0,1)$ for all $i \leq k$ where $x_{k+1} \equiv x_1$. Then:

$$\frac{\rho_{x_1}(x_1 x_k)}{\rho_{x_k}(x_1 x_k)} = \frac{\rho_{x_1}(x_1 x_2)}{\rho_{x_2}(x_1 x_2)} \cdot \frac{\rho_{x_2}(x_2 x_3)}{\rho_{x_3}(x_2 x_3)} \cdots \frac{\rho_{x_{k-1}}(x_{k-1} x_k)}{\rho_{x_k}(x_{k-1} x_k)}$$

These last two conditions are relaxations of Luce's IIA, even though the former is not directly defined on the observed choice probabilities.[14] This condition guarantees that fixation value u and hence fixation probabilities can be properly defined, which is stated in the Appendix. Equipped with the specific u, one can define the fixation probability of x in M as $g_x(M) := \frac{u(x)}{u(M)}$. Note that depending on u, g values change, so g should also depend on u in the notation, which I suppress here. The following condition relies on this construction from the relative fixation probabilities, which states the relative probability increase in fixating at x when one adds z to some menu is equal to the probability that the DM does not fixate at z in the final menu.

Condition 8. ***Fixation Ratio Consistency***
For any menu M and $z \in M$:

$$1 - g_z(M) = \frac{g_x(M)}{g_x(M \setminus \{z\})}$$

To see its necessity, note that $1 - g_z(M) = \frac{u(M \setminus \{z\})}{u(M)}$, and $\frac{g_x(M)}{g_x(M \setminus \{z\})} = \frac{u(M \setminus \{z\})}{u(M)}$, which implies the conclusion.

The final set of conditions guarantee that saccade probabilities are properly defined. The first one is as follows:

[14]In their working paper version, ES uses very similar conditions to Conditions 6-7, which are relaxations of Luce's IIA, while the remaining two conditions are violated by svc. Also, in the proof, they use the dominance transitivity condition, which is implied by two conditions that characterize their model, which are not satisfied by svc.

Condition 9. ***Bounded Fixation Ratio***
For any menu M and $z \in M$ such that $\rho_z(M) = 0$, the following holds:

$$\frac{g_x(M)}{g_x(M \setminus \{z\})} \leq \min\{\frac{\rho_x(M)}{\rho_x(M \setminus \{z\})}, \frac{1-\rho_x(M)}{1-\rho_x(M \setminus \{z\})}\}$$

If furthermore $\rho_x(xz) \in (0,1)$, then the second inequality becomes:

$$\frac{g_x(M)}{g_x(M \setminus \{z\})} = \frac{\rho_x(M)}{\rho_x(M \setminus \{z\})}$$

The necessity of this condition is not straightforward, and therefore it is established in the Appendix. For the final condition, I need the following definition:

Definition 11. *$\{t_i\}_{i=1}^k$ is said to be a h-path connecting x and y if $t_1 \equiv x$ and $t_k \equiv y$ such that for any t_i, t_{i+1} there is an associated menu M_i where $t_i, t_{i+1} \succ z_i$ for some $z_i \in M_i$ with $\rho_{z_i}(M_i) = 0$. Given this, define:*

$$h(\{t_i\}_{i=1}^k) := \Pi_{i \leq k-1} \frac{h_{t_i}(z_i, M_i)}{h_{t_{i+1}}(z_i, M_i)}$$

Similarly to the case of fixation probabilities, $h(\{t_i\}_{i=1}^k) = \frac{v(t_1)}{v(t_2)} \cdot \dots \cdot \frac{v(t_{k-1})}{v(t_k)} = \frac{v(t_1)}{v(t_k)}$, and the last term is equal to $\frac{v(x)}{v(y)}$ by definition. The same ratio is reached if one considers another such path that connects x and y. This results in the following condition.

Condition 10. ***Cyclically Independent Saccade***
For any two h-paths $\{t_i\}_{i=1}^k$ and $\{t'_i\}_{i=1}^{k'}$ connecting x and y:

$$h(\{t_i\}_{i=1}^k) = h(\{t'_i\}_{i=1}^{k'})$$

The following result provides the characterization of svc.

Theorem 4. *A rcf ρ can be represented by svc iff ρ satisfies Conditions 4-10.*

Even though preferences are uniquely identified, it can be the case that relative fixation probabilities are not unique. To guarantee the uniqueness of these, a certain richness condition can be imposed on X. Next, I state this condition.

Definition 12. *X is said to be dense if for all $x, y \in X$ such that $\rho_x(xy) \in \{0,1\}$ there is an incomparability path $\{z_i\}_{i=1}^k$ such that $z_1 \equiv x$ and $z_k \equiv y$ with $\Pi_{i=1}^{k-1} \rho_{z_i}(z_i z_{i+1}) \cdot \in (0,1)$.*

The density assumption guarantees that there is an incomparability path for any two alternatives that are comparable, and hence the relative fixation probabilities are uniquely identified for any two alternatives. Note that this is true for relative fixation probabilities, and not for the fixation probabilities themselves. So, two different value functions u and u' can represent the same svc if the ratios remain the same for any two alternatives. The denseness condition is satisfied by the *feature spaces*. Assume that X is the feature space, that is, $X \subseteq \mathbb{R}^k$ for $k \geq 1$. For example, consider $k = 2$. The preferences $\succeq$ on X can be defined as a vector ordering over these two dimensions. Observe that when this is the case, two alternatives x and y are comparable iff either x or y dominates the other alternative in both dimensions. This implies that for any $x \in X$, there is $y \neq x$ in X such that $x \bowtie y$ provided that alternatives cannot beat each other in all dimensions. The denseness condition is satisfied for all feature spaces whenever $k > 1$ and X is not finite. On the other hand, when $k = 1$, $\succeq$ defined above is always complete, and hence the condition is violated. A similar condition can be stated for the saccade probabilities with the same logic:

Definition 13. *X is said to be dense* if for all $x,y \in X$ there is a menu M such that $x,y \succ z$ and $\rho_z(M) = 0$ for some $z \in M$.*

This is also plausible in the setting of feature spaces. For this, consider two distinct alternatives x and y. If x and y are incomparable to each other, then one can define a third alternative $z = (\min\{x_1,y_1\},\ldots,\min\{x_k,y_k\})$ which is also in the feature space. If they are comparable and say x dominates y, then one can directly consider $z = (y_i - \varepsilon, y_{-i})$ where $\varepsilon > 0$. In menu $M = \{x,y,z\}$, $x,y \succ z$ and $\rho_z(M) = 0$.[15]

Proposition 4. *If X is dense (dense*), then two satisficing visual choice models $(u,v,\succeq)$ and $(u',v',\succeq')$ represent the same rcv iff there is $\lambda > 0$ such that $u = \lambda u'$ ($v = \lambda v'$) and $\succeq = \succeq'$.*

5.3 Comparative Statics

In this section, I ask the question of how changes in primitive values affect the properties of the model. Assume that $\succeq$ is fixed. This leaves two parameters that can change: u and v. Take an alternative x and a menu M that includes x. I will first consider how changes in fixation value affects choice probabilities. Assume an increase in $u(x)$. The change of $\rho_x(M)$ with respect to $u(x)$ is given by:

$$\frac{\partial \rho_x(M)}{\partial u(x)} = \begin{cases} \frac{1}{u(M)} - \frac{A_x}{u(M)^2} & \text{if} \quad D(x,M) = \emptyset \\ -\frac{A_x}{u(M)^2} & \text{if} \quad D(x,M) \neq \emptyset \end{cases}$$

where $A_x := \sum_{x \succ y} u(y) \cdot \frac{v(x)}{v(D(y,M))}$. One can show that $1 - \frac{A_x}{u(M)} > 0$, which implies that $\frac{\partial \rho_x(M)}{\partial u(x)} > 0$ when $D(x,M) = \emptyset$. So, increasing $u(x)$ leads to an increase in the choice probability of $\rho_x(M)$ when there is no alternative that strictly dominates x, while it leads to a decrease otherwise.

What if the fixation value of a distinct alternative y, $u(y)$, increases? Consider any menu M that includes both x and y. If $y \succ x$ or $y \bowtie x$, the only effect is through increasing the denominator of the fixation probability, while this is not the case if $x \succ y$. In particular:

$$\frac{\partial \rho_x(M)}{\partial u(y)} = \begin{cases} \frac{1}{u(M)} \frac{v(x)}{v(D(y,M))} - \frac{A_x}{u(M)^2} & \text{if} \quad x \succ y \\ -\frac{A_x}{u(M)^2} & \text{o.w.} \end{cases}$$

The effect of a increase in $u(y)$ results in the same conclusion when $y \succ x$ or $y \bowtie x$, leading to a decrease in $\rho_x(M)$. The situation when $x \succ y$ depends on the comparison between $\frac{v(x)}{v(D(y,M))}$ and $\frac{A_x}{u(M)}$. Through some algebra, one can show that increasing $u(y)$ increases $\rho_x(M)$ iff:

$$\frac{v(x)}{v(D(y,M))} \geq \frac{A_x}{u(M)} \equiv \sum_{x \succ y} \frac{u(y)}{u(M)} \cdot \frac{v(x)}{v(D(y,M))}$$

This shows that the increase in choice probability depends on the comparison of two effects: on the left-hand side, choice probability of x might increase due to saccades, while on the right hand side it can decrease due to the decrease in fixation probabilities.

[15] If one assumes that X is finite, then it might admit a *maximum* or *minimum* alternative for which there is no incomparable alternative. Still, finite feature spaces can be dense under suitable conditions. For example, assume that each property is binary and $x_i = 1$ iff x has the i^{th} property. Let $\mathbf{1} = (1,1,\ldots,1)$ and $\mathbf{0} = (0,0,\ldots,0)$. Then, $X = \{0,1\}^k \setminus \{\mathbf{0},\mathbf{1}\}$ is dense. For any $x,y \in X$, one can find z incomparable to both by only changing a property that is not present in both alternatives. If there is no such property, then do this by changing two distinct properties which are not present in x and y. The second denseness condition is harder to satisfy without assuming further structure.

Now, I will look at how the saccade values affect the choice probabilities. First, consider an increase in $v(x)$. This has a positive impact by increasing the contribution coming from the saccade probability's numerator, while it leads to a decrease if the denominator is affected more. In particular:

$$\frac{\partial \rho_x(M)}{\partial v(x)} = \begin{cases} \sum_{x \succeq y} \frac{u(y)}{u(M)} \frac{v(D(y,M)) - v(x)}{v(D(y,M))^2} & \text{if there is y s.t.} \quad x \succ y \\ 0 & \text{o.w.} \end{cases}$$

Since $v(D(y,M)) \geq v(x)$, the derivative is always positive, and therefore $\rho_x(M)$ increases as $v(x)$ increases. On the other hand, an increase in $v(y)$ leads to a decrease in $\rho_x(M)$ if there is $z \in M$ such that $x, y \in D(z,M)$. If there is no such z, then an increase in $v(y)$ does not affect $\rho_x(M)$. These observations can be summed up as follows:

- An increase in the fixation value for an alternative x leads to a decrease in $\rho_x(M)$ unless $D(x,M) = \emptyset$, while it leads to a decrease in $\rho_y(M)$ unless $y \succ x$ and $\frac{u(M)}{v(D(y,M))} \geq \sum_{x \succeq z} \frac{u(z)}{v(D(z,M))}$.
- An increase in the saccade value for an alternative x increases $\rho_x(M)$ only when $x \in \hat{D}(y,M)$ for some $y \in M$, and has no effect otherwise. Provided that $x, y \in \hat{D}(z,M)$ for some distinct z, $\rho_y(M)$ decreases, and otherwise it has no impact on $\rho_y(M)$.

Say (u,v) is *coherent* if an increase in $u(x)$ occurs iff there is an increase in $v(x)$ for any $x \in X$, which is trivially the case if $u = v$. If $u(x)$ increases and (u,v) is coherent, then a decrease in $\rho_x(M)$ occurs only when $D(x,M) \neq \emptyset$ and the decrease due to the increase in $u(x)$ (fixation effect) dominates the increase due to the increase in $v(x)$ (saccade effect). For a distinct alternative y, there are several cases to consider. First, assume that x and y do not share an alternative that both dominate. This can occur only if $x \succ y$ or $x \bowtie y$, since otherwise when $y \succ x$, they need to share at least one alternative provided $x \succ z$ for some $z \in M \setminus \{x,y\}$. In both cases, $\rho_y(M)$ decreases only due to the fixation effect. On the other hand, if they share an alternative both dominate, then $y \succ x$ (in addition to $x \succ y$ or $x \bowtie y$) is possible. If it is not the case that $y \succ x$, then $\rho_y(M)$ decreases from both fixation and saccade effects. If $y \succ x$ and the condition $\frac{u(M)}{v(D(y,M))} \geq \sum_{x \succeq z} \frac{u(z)}{v(D(z,M))}$ is not satisfied, then the same conclusion is arrived. The only case when $\rho_y(M)$ increases is when this condition holds and the fixation effect dominates the saccade effect. Hence, the coherency case presents an interesting situation where the salience value of x increases for both u and v, but this leads to an increase in the choice probability of y. Let $\tilde{u}(x)$ and $\tilde{v}(x)$ denote the transformed values for x. The following proposition provides a sufficient condition for this situation to happen:

Proposition 5. *Assume that (u,v) is coherent and the saliency of x improves in both. If $\rho_y(xy) = 1$ and $\sum_{x \succ z} \frac{v(D(y,M))}{v(D(z,M))} \leq 1$, then $\rho_y(M)$ increases. If furthermore $u(M \setminus \{x\}) < \frac{v(x)\tilde{u}(x) - \tilde{v}(x)u(x)}{\tilde{v}(x) - v(x)}$, then $\rho_x(M)$ decreases.*

Proof. The proof of the first claim is provided in the above discussion. For the latter, note that $\rho_x(M)$ strictly decreases after the transformation iff $\frac{1}{u(M)} A_x > \frac{1}{u(M \setminus \{x\}) + \tilde{u}(x)} \tilde{A}_x$ where the latter denotes the $\sum_{x \succ z} u(z) \cdot \frac{v(x)}{v(D(z,M))}$ with transformed values for x. This is equivalent to the following claim:

$$\frac{v(x)}{u(M)} \sum_{x \succ z} \frac{u(z)}{v(D(z,M))} > \frac{\tilde{v}(x)}{u(M \setminus \{x\}) + \tilde{u}(x)} \sum_{x \succ z} \frac{u(z)}{v(D(z,M) \setminus \{x\}) + \tilde{v}(x)}$$

Because $\sum_{x\succ z}\frac{u(z)}{v(D(z,M))} > \sum_{x\succ z}\frac{u(z)}{v(D(z,M)\setminus\{x\})+\tilde{v}(x)}$, the right hand side is strictly less than

$$\frac{\tilde{v}(x)}{u(M\setminus\{x\})+\tilde{u}(x)}\sum_{x\succ z}\frac{u(z)}{v(D(z,M))}$$

which implies coupled with the assumption $u(M\setminus\{x\}) < \frac{v(x)\tilde{u}(x)-\tilde{v}(x)u(x)}{\tilde{v}(x)-v(x)}$ the desired conclusion. $\square$

Another widely looked comparative static is the change in the probability of choosing the best product and its relative ratio compared to the second best product. Assume in this section that $\succeq$ is complete and fix the menu to be the grand set X. Assume that $X=\{x_1,x_2,\dots,x_n\}$ and $x_i \succ x_{i+1}$ for all $i\le n-1$. Let $\rho_i(X)$ denote the probability of choosing the i^{th} best product. I am specifically interested in understanding how does $\rho_1(X)$ and also $\frac{\rho_1(X)}{\rho_2(X)}$ change with respect to the changes in X. In particular, a widely asked question regarding the contextual effects is about the impact of adding a third alternative into a choice set consisting of two alternatives. First, note that:

$$\rho_1(X)=\frac{u(\{x_1,x_2\})}{u(X)}+\sum_{i>2}\frac{u(x_i)}{u(X)}\cdot\frac{v(x_1)}{v(\{x_1,\dots,x_{i-1}\})}$$

and

$$\rho_2(X)=\sum_{i>2}\frac{u(x_i)}{u(X)}\cdot\frac{v(x_2)}{v(\{x_1,\dots,x_{i-1}\})}$$

This implies that:

$$\frac{\rho_1(X)}{\rho_2(X)}=\frac{v(x_1)}{v(x_2)}+\frac{u(\{x_1,x_2\})}{v(x_2)}\cdot\frac{1}{\sum_{i>2}\frac{u(x_i)}{v(\{x_1,\dots,x_{i-1}\})}}$$

which simplifies to:

$$\frac{\rho_1(X)}{\rho_2(X)}=\frac{v(x_1)}{v(x_2)}+\frac{u(\{x_1,x_2\})v(\{x_1,x_2\})}{v(x_2)u(x_3)}$$

when X contains first three alternatives. Note that because $\rho_1(X)=1$ when X contains only two alternatives, so this ratio is equivalent to plus infinity. Thus, the addition of the third alternative necessarily decreases this ratio. Furthermore, an increase in the value of the third alternative also decreases this ratio, meaning that adding a better third alternative implies a relative gain for the second-best alternative.

5.4 Connections to the Literature

In this section, I discuss the most closely related models in the literature. Further discussion of other models and their relation to visual choice is postponed to the Appendix. The observation that svc satisfies MST has two immediate implications regarding the relation of svc to the literature. First, one can conclude that svc is not subsumed by the *scalability models*, characterized by a stronger condition compared to SST as shown in Tversky and Russo [49]. Second, it shows that svc and rum are not nested by each other, but they intersect nontrivially. This follows because svc might violate regularity, which is satisfied by rum, while rum might violate WST using Condorcet cycles. Luce rule lies in the intersection of both models as will be clear later on. Therefore, the following is concluded.

Many models in the literature including *Luce rule*, *random consideration set rule* of Manzini and Mariotti [33] (MM), and the *attribute rule* by Gul, Natenzon and Pesendorfer [20] (GNP) are special cases of rum. Therefore, this conclusion also implies that svc is not nested by any of these models.[16]

The situation regarding visual and rational visual choice is different. To explain this, consider the **random attention model (ram)** developed by Cattaneo et al. [6], which further generalizes rum. In ram, the DM is endowed with a strict preference relation $\succ$ and a random attention mapping $\Gamma : \mathbb{X} \times \mathbb{X} \to [0,1]$ such that for all $M \in \mathbb{X}$ one has $\Gamma(S|M) \geq 0$ if $S \subseteq M$, $\Gamma(S|M) = 0$ otherwise, and finally $\sum_{S \subseteq M} \Gamma(S|M) = 1$. They impose an additional *monotonicity* assumption which says that $\Gamma(S|M) \leq \Gamma(S|M \setminus \{x\})$ for all $x \in M \setminus S$. Given these, the ram is defined as:

$$\rho_x^{ram}(M) := \sum_{\{S \subseteq M : x = \max(S,\succ)\}} \Gamma(S|M)$$

The random attention model is characterized by the acyclicity imposed on a binary relation induced from violations of regularity. A violation of regularity implies for the visual choice the following:

$$\rho_x(M) \geq \rho_x(M \setminus \{y\}) \quad \Longrightarrow \quad x \in s(y,M)$$

Let $x\tilde{P}y$ iff $x \in s(y,M)$. The random attention model is characterized by the acyclicity of $\tilde{P}$. It is easy to show that visual choice can accommodate cycles of $\tilde{P}$, so vc is not nested by ram. On the other hand, rational visual choice satisfies the acyclicity condition, implying that rational visual choice can be 'framed' as a random attention model suggesting a specific attention process with a certain level of rationality. Thus, rvc is nested by ram, and hence svc.

Previously, visual choice is shown to to generalize the general Luce model. Specifically, fixation-independent visual choice generalizes glm, and coincides with it if furthermore the saccade correspondence is nonempty valued. The restrictions associated with svc makes visual choice closer to glm, although they are not nested by each other. Next, I show this:

Proposition 6. *Satisficing visual choice and general luce model are independent with nonempty intersection.*

Proof. To see that glm is not nested by svc, observe that glm might violate even WST. Consider three alternatives x,y,z with $x \geq_\rho y$ and $y \geq_\rho z$. Furthermore, let $c(xy) = xy$, $c(yz) = yz$, and $c(xz) = z$. It is straightforward to see by the definition of glm that $\rho_x(xz) = 0$, and WST is violated.

The other side is less trivial. I am going to show that svc might violate cyclical independence. Consider x_1,x_2,x_3,x_4 such that $x_1 \bowtie x_2 \bowtie x_3$, $x_1 \succ x_3 \succ x_4$ and assume that $u = v$. Let $M_1 = x_1x_2$, $M_2 = x_2x_3$ and $M_3 = x_1x_3x_4$. First, let me check whether $\rho_{x_i}(M_i), \rho_{x_i}(M_{i+1}) > 0$ for all $i \in \{1,\ldots,n\}$ holds. Note that $\rho_{x_1}(M_1) = \frac{u(x_1)}{u(x_1x_2)}$ and $\rho_{x_2}(M_1) = \frac{u(x_2)}{u(x_1x_2)}$, so these satisfy the condition. Similarly, the condition holds for the pairs (x_2,M_2) and (x_3,M_2). These imply that:

$$\frac{\rho_{x_1}(M_1)}{\rho_{x_2}(M_1)} = \frac{u(x_1)}{u(x_2)} \quad \& \quad \frac{\rho_{x_2}(M_2)}{\rho_{x_3}(M_2)} = \frac{u(x_2)}{u(x_3)}$$

[16] Luce is a special case of gnp as well, so they have nonempty intersection. Moreover, gnp shows that attribute rules approximate any rum under a further condition, which implies that gnp and svc are not nested by each other when this is the case. On the other hand, the issue about MM is more subtle because they assume that there is a *default option*. Horan [22] analyzed MM without the default option and showed that the version with and without default are almost the same from an empirical point of view. In particular, both versions are mainly separated by the difference in the *regularity* condition they impose on choice probabilities. It is shown that svc is able to accommodate violations of regularity. This shows that svc is not nested by mm. On the other hand, mm can accommodate violations of stochastic transitivity, and hence not nested by svc.

Cyclical independence condition implies that:

$$\frac{\rho_{x_1}(M_3)}{\rho_{x_3}(M_3)} = \frac{\rho_{x_1}(M_1)}{\rho_{x_2}(M_1)}\frac{\rho_{x_2}(M_2)}{\rho_{x_3}(M_2)} = \frac{u(x_1)}{u(x_3)}$$

For the left-hand side, first note that:

$$\rho_{x_1}(M_3) = \frac{u(x_4)}{u(M_3)}\frac{u(x_1)}{u(x_1x_3)} + \frac{u(x_1x_3)}{u(M_3)}$$

whereas:

$$\rho_{x_3}(M_3) = \frac{u(x_4)}{u(M_3)}\frac{u(x_3)}{u(x_1x_3)}$$

which implies that the left-hand side is equal to:

$$\frac{u(x_4)u(x_1) + u(x_1x_3)u(x_1x_3)}{u(x_4)u(x_3)}$$

Thus, the condition implied by cyclical independence is satisfied iff:

$$u(x_3)u(x_4)u(x_1) + u(x_3)u(x_1x_3)u(x_1x_3) = u(x_1)u(x_4)u(x_3)$$

which is equivalent to the following:

$$u(x_3)u(x_1x_3)u(x_1x_3) = u(x_4)[u(x_3)u(x_1) - u(x_1)u(x_3)]$$

The right-hand side is equal to 0, while the left-hand side is strictly positive, and hence equality cannot be satisfied.

□

ES defines two further special cases: *two-stage glm* and *threshold glm*. The former is a special case of glm, while the latter is a further special case. In two-stage glm, the constraint correspondence is defined as the set of undominated alternatives in the relevant menu with respect to a partial order. This makes the connection between svc and glm more apparent.

Another model closely related is the **random reference model (rrm)** recently proposed by Kibris et al [26]. Assume that the DM is endowed with a family of strict complete preference relations $(\succ_x)_{x \in X}$ associated with the set of alternatives in X. Furthermore, assume this family satisfies the condition that if $x \succ_z y$, then $x \succ_x y$. An alternative x in menu M becomes the reference point of the DM with probability $\alpha(M,x)$. Assume that $\alpha(M,x) > 0$ for all $x \in M$ and α is regular. Then, rrm is defined as:

$$\rho_y^{rrm}(M) := \sum_{x \in M} \alpha(M,x)\mathbb{1}\{y = \max(M, \succ_x)\}$$

It turns out that even svc is not nested by the rrm. To see this, it is sufficient to check the following characterizing condition of rrm:

$$\rho_x(xy) = 0 \implies \rho_x(M) = 0 \quad \forall M \ni y$$

It is easy to show that svc can violate this condition. Consider three distinct alternatives x, y, z such that $x \succ y \succ z$. While $\rho_y(xy) = 0$, $\rho_y(xyz) > 0$. Thus, the above condition is not satisfied. The main discussion evolves around the special case of rrm called *logit random reference model* *(lrrm)*. In lrrm, the reference-point distribution is specified using Luce rule as in visual choice, that is, $\alpha(M,x) = \frac{s_x}{\sum_{y\in M} s_y}$ where s_x is the *salience weight* associated with x. Thus, rational visual choice generalizes the *logit-random reference model (lrrm)*, and if s is single-valued everywhere and satisfies a condition corresponding to the status-quo bias assumption of lrrm, then it becomes logit rrm. [17] Thus, lrrm lies in the intersection of rational visual choice (and hence visual choice) and random reference model.

6 Discussion

6.1 Inhibition of Return and Fixation Dependence with a Stopping Point

For the deterministic version of visual choice, I focused on the case where the DM fixates on one alternative. However, the deterministic version (as opposed to the random version) does not account for the *inhibition of return* effect. The experimental studies show that individuals fixate usually on more than one alternative. Let $f_i(M)$ be the i^{th} alternative fixated by the individual, and define $f^i(M) := \{f_1(M),\ldots,f_i(M)\}$ where $f^0(M) = \emptyset$. The individual fixates on the most salient alternative according to u. So:

$$f_i(M) := \max(M \setminus f^{i-1}(M), \succ_u)$$

Let $k(M) \leq |M|$ be the number of alternatives exogenously given conditional on the menu. Assume for simplicity that this is a uniform number over all menus, so $k(M) = k$ for any menu M. If $k \geq |M|$, this implies necessarily that the DM fixates on $|M|$ alternatives, so the number of fixations in a menu M is given by $\min\{|M|,k\}$. The alternatives that are seen by the DM is given by:

$$S(M) := \bigcup_{i\leq k(M)} \hat{s}(f_i(M),M)$$

Note that this formulation of the 'seen set' (see Reutskaya et al. [41]) assumes implicitly that the DM has *perfect memory*. Still, the DM does not need to remember these seen alternatives provided that the brain does not forget them. Since $\hat{s}$ is assumed to be a singleton, this implies that for any menu M, the individual will saccade to $\min\{|M|,k\}$-alternatives, which are not necessarily different from each other. The deterministic visual choice can be defined as:

$$c(M) := \max(S(M), \succ_v)$$

The identification and characterization of this procedure is more complicated, and not analyzed in this paper. However, I will make several notes regarding satisficing visual choice case. The deterministic choice rule becomes:

$$c(M) = \max(\bigcup_{i\leq k} \hat{D}(f_i(M),M), \succ_v)$$

[17] lrar reduces to *deterministic choice* when $\succ_x = \succ$ for all $x \in X$, which is not the case with svc except binary menus. On the other hand, if $x \succ_x y$ for all x and y distinct, then lrar becomes the *Luce rule* with s. Thus, letting $u = s$ (the salience value which determines fixation in svc is equal to salience weight in lrar), the two models coincide under different scenarios. Recall that svc reduces to *Luce rule* when all alternatives are incomparable, while lrar reduces to Luce when all alternatives are strictly comparable but each alternative x favors itself provided that it is indeed the reference point. One important difference between these models (beyond the interpretation and their independence) is that the number of primitives rar has directly proportional to the number of alternatives. In particular, for each alternative, there is an associated preference relation (i.e. $|X|$ preferences), and in addition there is salience weight ($|X|$ salience weights). On the other hand, svc has 3 primitives, and further assuming $u = v$ makes this number 2.

If $\succeq$ is complete, then the DM always saccades to the set of maximal alternatives in the menu with respect to $\succeq$. Because c is a choice function, $c(M) \in \max(M, \succeq)$ when $\succeq$ is complete. In particular, $c(M)$ will be the most salient alternative with respect to v among the alternatives in $\max(M, \succeq)$.[18] If $\succeq$ is not complete, then the DM does not necessarily choose an alternative in $MAX(M, \succeq)$. However, note that this is guaranteed when the DM fixates on all alternatives, that is, when $k = |M|$. Another condition that guarantees this is satisfied is when top-down preferences affect the bottom-up salience functions. For this, assume that the order induced by u preserves the order of $\succeq$, that is, if $x \succeq y$, then $x \succ_u y$. This implies that even if $k = 1$, the first fixated alternative is an alternative in $MAX(M, \succeq)$. Therefore, $c(M) \in MAX(M, \succeq)$. This is in line with the observation that the choices of the DM might not be in line with her preferences, and only influenced by the bottom-up influenced fixations and eye movements of the individual (see Armel et al. [2]).

6.2 Allowing for a General Saccade Correspondence

Consider again the fixation-dependent saccade without assuming the satisficing structure. Assume that the following two conditions are satisfied, which are easily seen to be consistent with the satisficing case.

- $x \notin s(x, M) \quad \forall x \in M \;\&\; M \in \mathbb{X}$.
- $s(x, xy) = y$ implies that $s(y, xy) = \emptyset$.

This is plausible given the nature of saccades, because saccadic eye movements occur to different alternatives from the fixation point. However, the identification of the primitives $< s, u, v >$ is problematic.

Denote the *revealed saccade correspondence* by s^R. Consider a binary menu $\{x, y\}$. If $\rho_x(xy) = 1$, then this reveals that $\hat{s}^R(x, xy) = x = \hat{s}^R(y, xy)$. Otherwise, y should be chosen with positive probability either due to a saccade from x or fixation at y and no saccade to x. By definition, $x \notin s^R(x, M)$, so $\hat{s}^R(x, xy) = x$ implies that $s^R(x, xy) = \emptyset$. Now consider the interior probability case, that is, $\rho_x(xy) \in (0, 1)$. Because $\rho_x(xy) \neq 0$, it should be the case that either $\hat{s}^R(x, xy) \neq y$ or $\hat{s}^R(y, xy) \neq y$, which is equivalent to $\hat{s}^R(x, xy) = x$ or $\hat{s}^R(y, xy) = x$. If the former holds, then $\hat{s}^R(y, xy) = y$, which is implied by $\rho_x(xy) \neq 1$. Similarly, in the latter case, $\hat{s}^R(x, xy) = y$.

Also, regularity violations provide information about the saccade correspondence. For example, consider the larger menu $\{x, y, z\}$ and assume that $\rho_x(xyz) > \rho_x(xyz)$. This is possible only if $x \in s^R(z, xyz)$, because otherwise the choice probability of x should decrease due to the increase in the denominator of fixation probability and possibly saccade probability. This can be generalized to any menu M, so $\rho_x(M \cup \{z\}) > \rho_x(M)$ implies that $x \in s^R(z, M)$.

Still, identifying the saccade correspondence fully is problematic. Let me restrict my attention to rational visual choice. Assume that $\rho_x(xy) = 1$, which implies that $s^R(y, xy) = x$. Because I assume that the visual choice is rational, this reveals that x is strictly better compared to y when y is the fixation point. If there is no other alternative z such that $\rho_z(yz) = 1$, then x is the only alternative that dominates y with respect to y. This implies that for any M that includes y, the saccade correspondence will include only x if $x \in M$, and otherwise there would be no saccade and hence the (induced) saccade correspondence will be equal to x. However, if there is z such that $\rho_z(yz) = 1$, then z also dominates y with respect to y. This arises an issue when

[18] Later on, I define the notion of compatibility. If compatibility holds here for v, then this implies that for any $x, y \in max(M, \succeq)$ one has $v(x) = v(y)$. This is inconsistent with the assumption that $\succ_v$ is asymmetric.

one considers the menu $M = \{x, y, z\}$ because one needs to reveal the situation with respect to y between x and z. Assume furthermore that either $\rho_x(xz) = 1$ or $\rho_z(xz) = 1$, and wlog the former. The following is an assumption widely used in the literature that deals with status-quo bias, which is extended to the case of possible indifference here:

- $x \succsim_y z$ implies that $x \succsim_x z$.

Observe that the choice pattern described above is equivalent to the *compromise effect*, in which case one would expect that $\rho_y(xyz) < 1$. If $\rho_x(xyz) < 1$ and the status-quo bias assumption hold, then this is possible only if $z \succsim_y x$, because otherwise saccades only occur to x. Therefore, this would imply that x and z are indifferent with respect to each other. Since the one of the goals of the model is to capture such regularities, I viewed two such alternatives as indifferent to each other with respect to y, which is the case in the satisficing visual choice.

7 Further Discussion and Limitations

7.1 Different Saccade Correspondences

The literature suggests different versions of saccade correspondence. Two interesting versions inspired by Koch and Ullman [27] are as follows:

- Proximity-based saccade
- Similarity-based saccade

Both can be modeled using the feature spaces assuming that features contain the spatial coordinates. Let x be the fixation point. According to the former correspondence, the DM will saccade to any alternative that is sufficiently close to the fixation point. The most straightforward way to capture this is using the Euclidean metric to measure the distance and determining a certain threshold under which two alternatives are deemed to be close to each other. For determining the threshold, knowledge coming from the vision literature can be used. Alternatively, one can also assume that this threshold is unknown, and try to infer the threshold from the choice probabilities. For the latter, one needs to be careful about the choice of distance function. Although Euclidean metric (or any other metric) can be plausible in certain contexts, it is established in the literature that a distance function aimed at measuring similarity might violate the properties of a metric (see for example Tversky [50]). Note that one can also define a metric that is aimed to measure the contrast rather than the symmetry, and in fact this can be defined using the similarity metric provided that one assumes two things that are not similar are different enough.

Alternatively, one can model this by making v fixation-dependent and assuming that saccade correspondence is equal to the full menu for any fixation point. This implies the following choice procedure:

$$\rho_x(M) := \sum_{y \in M} \frac{u(y)}{u(M)} \cdot \begin{cases} \frac{v_y(x)}{v_y(M)} & if \quad x \in \hat{s}(y, M) \\ 0 & o.w. \end{cases}$$

Consider the proximity-based saccade case. Assume for simplicity that each alternative has a one-dimensional spatial representation (feature s). One can define a threshold-based v_y as follows:

$$v_y(x) := \begin{cases} |x_s - y_s|^{-1} & if \quad 0 < |x_s - y_s| \leq \varepsilon \\ 0 & o.w. \end{cases}$$

for some $\varepsilon > 0$ given exogenously. Thus, the choice probabilities conditional on fixation is determined through the relative distance of alternatives that are sufficiently close to the fixation point.

7.2 Menu Dependence

An issue concerning the model discussed in this paper is that both the salience and Luce values are defined in absolute terms. However, it is possible that both depend on the values other alternatives takes, so u and v can be functions of other alternatives in the menu. This makes visual choice menu dependent:

$$\rho_x(M) := \sum_{y \in M} \frac{u(y,M)}{\sum_{z \in M} u(z,M)} \cdot \begin{cases} \frac{v(x,M)}{\sum_{z \in \hat{s}(y,M)} v(z,M)} & \text{if} \quad x \in \hat{s}(y,M) \\ 0 & o.w. \end{cases} \tag{3}$$

The *divisive normalization* method developed in neuroscience of value-based decision making (see Carandini and Heeger [5]) can be used for determining the choice probabilities conditional on the fixation. The simplest version of the normalization takes the following form:

$$v(x,M) = \frac{v(x)}{\sqrt[\beta]{\sum_{y \in M} v(y)^{\beta}}}$$

without taking into account the so-called saturation parameter and the noise. If $\beta = 1$, applying divisive normalization is equivalent to the model in this paper. A recent suggestion by Laundry and Webb [30] connected to this is pairwise normalization defined as:

$$v(x,M) = \frac{v(x)}{\sum_{y \in M \setminus \{x\}} v(xy)}$$

Regarding the fixation probability, the salience measure developed by Bordalo et al. [3] can be used. They define a salience function σ such that $\sigma(x_i, \overline{M}_i)$ measures the salience of feature i for alternative with respect to $\overline{M}_i := \frac{1}{|M|} \sum_{x \in M} x_i$. This function is assumed to satisfy the properties called ordering and diminishing sensitivity. According to ordering, feature i is more salient for an alternative when x_i is more different from the average value $\overline{M}_i$ in the menu, and diminishing sensitivity says that the impact of feature differences decreases as the levels of the feature increases (Weber-Fechner law of sensory perception). Let $\omega_i(x,M)$ be the weight of feature i for alternative x in M, defined as a function of $\sigma_i(x_i, \overline{M}_i)$ and $\sigma_j(x_j, \overline{M}_j)$ for any feature $j \neq i$. The adjusted visual value function according to this salience can be defined as:

$$u_{\sigma}(x,M) := \sum_{i \leq k} \omega_i(x,M) \cdot x_i$$

Note that this is differentiated from the u proposed in the definition for visual choice by making the weight of a feature dependent on other alternatives in the menu using σ.

7.3 Complexity

Reutskaya et al. [41] provide a complexity-constrained visual choice model which is closely connected to visual choice. Let $\kappa(M)$ be the probability that menu M is complex, which depends on the complexity of the menu M. Given the menu is complex, the DM chooses each alternative with equal probability, and otherwise chooses the best alternative in the seen set. Thus, the conditional choice probability of x given fixation at y in menu M is given by:

$$\mathbf{p}^s_{(M,y)}(x) = \kappa(M)\frac{1}{|M|} + (1-\kappa(M))\mathbb{1}[\max(M,\succeq) = x]$$

To see how this is related to visual choice, consider the fixation-independent saccade correspondence with $s(x,M) = M$ for any $x \in M$. If furthermore $v(x) = v(y)$ for all x, y, then the choice probability given saccade is uniformly distributed in the menu, so $\mathbf{p}^s_{(M,y)}(x) = \frac{1}{|M|}$. If on the other hand s is still fixation-independent but $s(x,M) = \max(M,\succeq)$, then $\mathbf{p}^s_{(M,y)}(x) = 1$ iff $\max(M,\succeq) = x$, and 0 otherwise. Observe that both correspondences satisfy warp, so these two cases are examples of rational visual choice. However, the visual choice model presented here does not take into account the complexity of the choice problem.

References

[1] Ahumada, A., and Ülkü, L. (2018). Luce rule with limited consideration. Mathematical Social Sciences, 93, 52-56.

[2] Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. Nature neuroscience, 13(10), 1292-1298.

[3] Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. Journal of Political Economy, 121(5), 803-843.

[4] Bundesen, C. (1990). A theory of visual attention. Psychological review, 97(4), 523.

[5] Carandini, M., and Heeger, D. J. (2012). Normalization as a canonical neural computation. Nature Reviews Neuroscience, 13(1), 51-62. Chicago

[6] Cattaneo, M. D., Ma, X., Masatlioglu, Y., and Suleymanov, E. (2020). A random attention model. Journal of Political Economy, 128(7), 2796-2836.

[7] Li, X., and Camerer, C. F. (2022). Predictable effects of visual salience in experimental decisions and games. The Quarterly Journal of Economics, 137(3), 1849-1900.

[8] Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., and Riella, G. (2019). Deliberately stochastic. American Economic Review, 109(7), 2425-2445.

[9] Debreu, G. (1960). Individual choice behavior: A theoretical analysis.

[10] Demirkan, Yusufcan. (2017). Essays on decision theory (Master's thesis, Sosyal Bilimler Enstitüsü).

[11] Doğan, S., and Yıldız, K. (2021). Odds supermodularity and the Luce rule. Games and Economic Behavior, 126, 443-452.

[12] Dutta, R. (2020). Gradual pairwise comparison and stochastic choice. Theoretical Economics, 15(4), 1335-1364.

[13] Echenique, F., and Saito, K. (2019). General luce model. Economic Theory, 68(4), 811-826.

[14] Echenique, F, Saito, K., and Tserenjigmid, G. (2018). The perception-adjusted Luce model. Mathematical Social Sciences, 93, 67-76. Chicago

[15] Fechner, Gustav, T. (1860): Elemente der psychophysik.

[16] Fishburn, P., Choice probabilities and choice functions, J. Math. Psych. 18 (1978), 201-219.

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

[18] Giarlotta, A., Petralia, A., and Watson, S. (2022). Semantics meets attractiveness: Choice by salience. arXiv preprint arXiv:2204.08798.

[19] Webb, R., Glimcher, P. W., and Louie, K. (2020). Divisive normalization does influence decisions with multiple alternatives. Nature human behaviour, 4(11), 1118-1120.

[20] Gul, F., Natenzon, P., and Pesendorfer, W. (2014). Random choice as behavioral optimization. Econometrica, 82(5), 1873-1912.

[21] He, J., and Natenzon, P. (2019). Moderate expected utility. Available at SSRN 3243657.

[22] Horan, S. (2019). Random consideration and choice: A case study of "default" options. Mathematical Social Sciences, 102, 73-84.

[23] Huber, J., Payne, J. W., and Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. Journal of consumer research, 9(1), 90-98. Chicago

[24] Itti, L., and Koch, C. (2001). Computational modelling of visual attention. Nature reviews neuroscience, 2(3), 194-203.

[25] Kıbrıs, Ö., Masatlioglu, Y., and Suleymanov, E. (2018). A theory of reference point formation. Economic Theory, 1-30.

[26] Kibris, Ö., Masatlioglu, Y., and Suleymanov, E. A Random Reference Model. American Economic Journal: Microeconomics. Chicago

[27] Koch, C., and Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. Human neurobiology, 4(4), 219-227.

[28] Kovach, M., and Tserenjigmid, G. (2022). Behavioral foundations of nested stochastic choice and nested logit. Journal of Political Economy, 130(9), 2411-2461.

[29] Kovach, M., and Tserenjigmid, G. (2022). The Focal Luce model. American Economic Journal: Microeconomics, 14(3), 378-413.

[30] Landry, P., and Webb, R. (2021). Pairwise normalization: A neuroeconomic theory of multi-attribute choice. Journal of Economic Theory, 193, 105221.Chicago

[31] Louie, K., Khaw, M. W., and Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making. Proceedings of the National Academy of Sciences, 110(15), 6139-6144. Chicago

[32] Luce, R. D. (1959): Individual choice behavior, Wiley, New York

[33] Manzini, P., and Mariotti, M. (2014). Stochastic choice and consideration sets. Econometrica, 82(3), 1153-1176.

[34] Masatlioglu, Y., and Ok, E. A. (2014). A canonical model of choice with initial endowments. Review of Economic Studies, 81(2), 851-883.

[35] Masatlioglu, Yusufcan and Efe A. Ok (2005). Rational choice with status quo bias. Journal of Economic Theory, 121, 1– 29.

[36] Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. American Economic Review, 102(5), 2183-2205. Chicago

[37] Milosavljevic, M., Navalpakkam, V., Koch, C., and Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. Journal of consumer psychology, 22(1), 67-74.

[38] Ok, E. A., and Tserenjigmid, G. (2023). Measuring Stochastic Rationality. arXiv preprint arXiv:2303.08202.

[39] Kıbrıs, Ö., Masatlioglu, Y., and Suleymanov, E. (2018). A theory of reference point formation. Economic Theory, 1-30.

[40] Ravid, D., and Steverson, K. (2019). Focus, then compare.

[41] Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. American Economic Review, 101(2), 900-926.

[42] Rieskamp, J., Busemeyer, J. R., and Mellers, B. A. (2006). Extending the bounds of rationality: Evidence and theories of preferential choice. Journal of Economic Literature, 44(3), 631-661.

[43] Rubinstein, Ariel, and Lin Zhou. 1999. “Choice problems with a ‘reference’ point” Mathematical Social Sciences 37 (1999): 205–209

[44] Rubinstein, Ariel and Yuval Salant, 2006. "Two Comments on the Principle of Revealed Preference," Levine’s Bibliography 321307000000000272, UCLA Department of Economics.

[45] Towal, R. B., Mormann, M., and Koch, C. (2013). Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. Proceedings of the National Academy of Sciences, 110(40), E3858-E3867.

[46] Treisman, A. (2006). How the deployment of attention determines what we see. Visual cognition, 14(4-8), 411-443.

[47] Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. Cognitive psychology, 12(1), 97-136.

[48] Tversky, A. (1972). Choice by elimination. Journal of mathematical psychology, 9(4), 341-367.

[49] Tversky, A., and Russo, J. E. (1969). Substitutability and similarity in binary choices. Journal of Mathematical psychology, 6(1), 1-12.

[50] Tversky, A. (1977). Features of similarity. Psychological review, 84(4), 327.

[51] Valkanova, K. (2021). Markov stochastic choice. Mimeo.

[52] Wolfe, J. M., and Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it?. Nature reviews neuroscience, 5(6), 495-501.

[53] Yang, X., and Krajbich, I. (2023). A dynamic computational model of fixation and choice in multi-attribute decisions. Psychological Review, 130(1), 52–70

Appendix A Proofs

A.1 Proofs for Deterministic Visual Choice

Proof.
Neccessity is already shown. Assume that c satisfies the Condition 1, 2 and 3.

Step 1: $\succ_u^R$ is acyclic.

In the main body, I defined $z \succ_u^R x$ if there exists a set M such that $z \neq c(M) \neq c(M \setminus \{z\})$. If there is no choice reversal, then Condition 1 implies that $c(M) = c(M')$ for any $M' \subseteq M$ with $c(M) \in M'$. If this is the case, let $c(M) \succ_u^R x$ for $x \neq c(M)$ and $x \in M$. Assume to the contrary that $\succ_u^R$ is cyclic. Then there exist $x_1, x_2, ..., x_n$ such that $x_i \neq x_j$ for all $i \neq j$, and $x_1 \succ_u^R x_2 \succ_u^R .... \succ_u^R x_n \succ_u^R x_1$. By definition and cyclicity, for each $i \in \{1, ..., n\}$ there is $M_i \in \mathbb{X}$ such that $x_i \neq c(M_i) \neq c(M_i \setminus x_i)$, where $x_{i+1} \in M_i$ and $x_i \neq x_{i+1}$. Let $M^* = \bigcup_{i=1}^n M_i$ and $M_{-J} = M^* \setminus \bigcup_{i \in J} M_i$ where $J \subseteq \{1, ..., n\}$. Assume wlog that $c(M^*) \in M_k$ for some $k \in \{1, 2, ..., n\}$. I will apply Condition 1 repeatedly and reach a contradiction. Let $M = \{x_1, ..., x_n\}$. If $k \neq 1$, start from the first menu and since $x_1 \neq c(M_1) \neq c(M_1 \setminus x_1)$, by Condition 1:

$$c(M^*) = c(M_{-1} \cup M \cup \{c(M^*)\})$$

To see this, note that by the condition $c(M^*) = c(M^* \setminus z)$ for any $z \notin \{c(M^*), x_1\}$. Then note that $x_1 \in M^* \setminus \{z\}$, so again by the same condition, $c(M^* \setminus \{z\}) = c((M^* \setminus \{z\}) \setminus \{t\})$ for any $t \notin \{x_1, c(M^* \setminus \{z\})\}$ and so $c(M^*) = c(M^* \setminus \{z\}) = c(M^* \setminus \{z, t\})$. Repeating this, $c(M^*) = c(M^* \setminus (M_1 \setminus (M \cup \{c(M^*)\})))$, since $M_1 \setminus (M \cup \{c(M^*)\}) \subseteq M_1 \setminus \{c(M^*), x_1\}$. Let $M^1 = M_{-1} \cup M \cup \{c(M^*)\}$. If $k \neq 2$, since $x_2 \neq c(M_2) \neq c(M_2 \setminus x_2)$, we have $c(M^1) = c(M_{-\{1,2\}} \cup M \cup c(M^*))))$. Continuing this way, we have $c(M^*) = c(M_k \cup M)$. [19] Let x_l be the largest indexed element in $M \cap M_k$. If $x_{l-1} \notin M_k$, using condition D we have $c(M^*) = c(M_k \cup M) = c(M_k \cup (M \setminus \{x_{l-1}\}))$, since $x_{l-2} \neq c(M_{l-2}) \neq c(M_{l-2} \setminus x_{l-2})$. We can repeat this process until we have $c(M^*) = c(M_k \cup M) = c(M_k)$ [20] and by assumption $c(M_k) \neq x_k$. Thus, we know that $c(M^*) \in M_k \setminus x_k$, but we can apply the same procedure one more time [21] and get $c(M^*) = c(M_k \setminus \{x_k\})$ which implies $c(M_k) = c(M_k \setminus \{x_k\})$, a contradiction. Therefore, $\succ_u^R$ is acylic and we can extend $\succ_u^R$ into a complete preference relation, $\overline{\succ}_u^R$. The fixation point is defined as:

$$f^R(M) := \max(M, \overline{\succ}_u^R)$$

Note that by construction $f^R(M)$ satisfies the weak axiom.

Step 2: $\succeq^R$ is a partial order.

---

[19] Note that we are not doing this for M_k and recall $c(M^*) \in M_k$.

[20] We start the process by x_l, since this gives an algorithm that terminates at the desired result, that is $c(M^*) = c(M_k)$. The algorithm works in general, we can start with any element $x_j \in M_k \cap M$ and remove first the element that is not in S_k and has the largest index smaller than j. If we start with an element that is not in $M_k \cap M$, since we continue eliminating one-by-one using elements next to each other, the last element remaining in the process cannot be eliminated. But since we start the cycle with an element that we are not going to eliminate, the process reaches the desired point.

[21] We started with the assumption $c(M^*) \in M_k$ and using condition D reached to the conclusion that $c(M^*) = c(M_k)$, which furthermore implies that $c(M^*) \in M_k \setminus \{x_k\}$. So, starting from the beginning, we can derive $c(M^*) = c(M_k \setminus \{x_k\})$ again using D repeatedly, since this time we can directly drop $x_k \in M_k$, because we know it is not chosen from M^*.

Recall that $\succeq^R$ is defined as follows $x \succeq^R y$ if $y = f^R(M)$ for some $M \supseteq \{x,y\}$ and $c(\{x,y\}) = x$. Reflexivity is trivially satisfied since $f^R(\{x\}) = \max(\{x\}, \overline{\succ}_u^R) = x$. Similarly, antisymmetry follows easily, because both $x \succeq^R y$ and $y \succeq^R x$ cannot happen at the same time. So, only transitivity is left to prove. This follows by Condition 2. This can be easily seen because it amounts to $\succeq^R$ being transitive. More precisely, because $x \succeq^R y \succeq^R z$, it is known that $y = f^R(M)$ for some $M \supseteq \{x,y\}$ and $c(\{x,y\}) = x$, and also $z = f^R(M')$ for some $M' \supseteq \{y,z\}$ and $c(\{y,z\}) = y$. Because f^R satisfies the weak axiom, $y = f^R(\{x,y\})$ and $z = f^R(\{y,z\})$, and hence $x = f^R(\{x,z\})$. By Condition 2, it should be the case that $c(\{x,z\}) = x$, which implies by definition that $x \succeq^R z$.

Step 3: $\succ_v^R$ is acyclic.

Let $x \succ_v^R y$ if there is a menu M s.t. $\{x,y\} \subseteq M$ and $y \in D^R(M)$ such that $c(M) = x$. Assume to the contrary $\succ_v^R$ is cyclic, i.e. $x_1 \succ_v^R x_2 \succ_v^R ... \succ_v^R x_n \succ_v^R x_1$ for distinct x_i's from 1 to k and let $x_{k+1} \equiv x_1$. By cyclicity and the definition above, there exists M_i's such that $c(M_i) = x_i$ and $x_{i+1} \in D^R(M_i)$ for all $i \leq k$. However, this is directly ruled out by the Condition 3, so $\succ_v^R$ is acylic.

Step 3: Sufficiency

There are two cases to consider. If $z \neq c(M) \neq c(M \setminus \{z\})$, then by definition $z \overline{\succ}_u^R x$ for all $x \in M \setminus \{z\}$. Also, by definition $D^R(M) = \{y \in M : y = c(\{y,z\})\}$. By Condition 1, $c(M) = c(\{c(M),z,y\})$ and hence by definition $c(M) \overline{\succ}_v^R y$ for all $y \in M \setminus \{c(M)\}$. Thus, $c(M) = \max(D^R(M), \overline{\succ}_v^R)$, and hence it is visual choice. If there is no $z \in M \setminus \{c(M)\}$ such that $c(M) \neq c(M \setminus \{z\})$, then by Condition 1, $c(M) = c(M')$ for all $M' \subseteq M$ with $c(M) \in M'$. In this case, $f^R(M) = c(M)$ and hence $c(M) \overline{\succ}_u^R y$ for all $y \in M \setminus \{c(M)\}$. By Condition 1, there is no z such that $c(\{c(M),z\}) = z$, so $D^R(M) = \{c(M)\}$. This shows that $c(M) = \max(D^R(M), \overline{\succ}_v^R)$, and hence it is visual choice also in this case. This concludes the proof.

$\square$

A.2 Proofs for Rationality of SVC

The following is a result shown in Ok and Tserenjigmid (henceforth OT) that is going to be used in the following results.

Proposition 7. *ρ is maximally (minimally) rational iff ρ is λ-rational (not λ-rational) $\forall \lambda \in [0,1]$.*

Now, I am going to provide the characterization presented by OT for $\trianglerighteq_r$ in order to operationalize the comparative rationality ordering.

$$\begin{aligned}
\mathrm{Ch}(\rho) &:= \bigcup_{\{S,M \in \mathbb{X}: S \subseteq M, x \in S\}} (\rho_x^*(S), \rho_x^*(M)] \\
\mathrm{Con}(\rho) &:= \bigcup_{\{M \in \mathbb{X}: x \in M\}} (\rho_x^*(M), min_{y \in M} \rho_x^*(xy)] \\
\mathrm{ST}(\rho) &:= \bigcup_{x,y,z \in \mathbb{X}} (\max\{\rho_y^*(xy), \rho_z^*(yz)\}, \rho_z^*(xz)]
\end{aligned}$$

For example, if $\lambda \in \mathrm{Ch}(\rho)$ defined above, then this means that *Chernoff axiom* is violated by rcf ρ and therefore ρ is not λ-rational. From these three sets, one can define the following:

$$\Lambda(\rho) := \mathrm{Ch}(\rho) \cup \mathrm{Con}(\rho) \cup \mathrm{ST}(\rho)$$

OT shows that for any pair of rcfs ρ and $\tilde{\rho}$:

$$\rho \trianglerighteq_r \tilde{\rho} \iff \Lambda(\rho) \subseteq \Lambda(\tilde{\rho})$$

The *rationality index* which enables one to compare all rcfs as opposed to the partial order $\trianglerighteq_r$ is defined as:

$$\iota_r(\rho) := 1 - l(\Lambda(\rho))$$

where l denotes the *Lebesgue measure*. This index is consistent with the partial order, and attaches a value of 1 (0) iff a rcf is maximally (minimally) rational. Two simplifications with regard to computing the rationality index are the following.

Lemma 1. *$ST(\rho) = \emptyset$.*

Proof. In Proposition 4 it is shown that svc satisfies MST. OT shows that $\mathrm{ST}(\rho) = 0$ for any rcf that satisfies MST. □

Lemma 2. *If $\succeq$ is complete, then $Con(\rho) = \emptyset$.*

Proof. Consider $x, y \in X$ and assume wlog $y \succ x$. When this is the case, $\rho_y(xy) = 1$ and $\rho_x(xy) = 0$, which further implies that $\rho^*_y(xy) = 1$ and $\rho^*_x(xy) = 0$. Thus, $min_{y\in M}\rho^*_x(xy) = 0$ for any M such that $y \in M$ and $y \succ x$. This results in $\emptyset$ for all such menus x is in because $(\rho^*_x(M), 0] = \emptyset$. Therefore, consider any menu M in which x is undominated, i.e. there is no $y \in M$ such that $y \succ x$. By completeness of $\succeq$, this is possible only when $x \succ y$ for all $y \in M \setminus \{x\}$, which implies $min_{y\in M}\rho^*_x(xy) = 1$. Furthermore, $\rho^*_x(M) = 1$ for all such menus by monotonicity shown in Proposition **??**. This also induces $\emptyset$ since $(1, 1] = \emptyset$. Therefore, $\mathrm{Con}(\rho) = \emptyset$. □

Proof of Theorem 2

Proof. The previous observations show that I only need to check the Chernoff set $\mathrm{Ch}(\rho)$. Also, I showed that svc satisfies monotonicity in Proposition **??**. Let $\rho_i(M)$ denote the choice probability of the i^{th} best product in M with respect to $\succeq$. Compatibility of v implies monotonicity, and this in turn implies that $\rho_i(M) \geq \rho_{i+1}(M)$ for all $i \in \{1, 2, \ldots, |M| - 1\}$. $\max(X, \succeq)$ is chosen in every menu it is present with probability 1, and hence induces $\emptyset$. The same holds for $\min(X, \succeq)$, but now because it is chosen with 0 probability in any menu it is present. In general, any alternative that is worst with respect to $\succeq$ in a menu is chosen with 0 probability. Consider i_X for some $i \notin \{1, |X|\}$. If say $i = 2$, then 2_X is chosen with 0 probability in a menu M iff $M = \{1_X, 2_X\}$. If $i = 3$, then M should contain either 1_X or 2_X without containing any j_X such that $j > i$. In general, $\rho^*_{i_X}(M) = 1$ if $j_X \notin M$ for any $j < i$. This implies that i_X induces the empty set if $\rho^*_{i_X}(M) = 1$, because this implies $\rho^*_{i_X}(S) = 1$ for any $S \subseteq M$ that has i_X. Therefore, consider $M \ni j_X$ for some $j < i$. The lower bound takes the smallest value when i_X is the $\succeq$-worst product in S, which is equal to 0. Letting $\overline{\rho}^*_{i_X} \equiv \max_{\{M\in\mathbb{X}:\exists j<i:\{i_X, j_X\}\subset M\}} \rho^*_{i_X}(M)$, this shows that the set induced by i_X is equal to $(0, \overline{\rho}^*_{i_X}]$. This implies that $\mathrm{Ch}(\rho) = \bigcup_{i\in\{2,\ldots,|X|-1\}}(0, \overline{\rho}^*_{i_X}]$, which proves that any two rcfs are comparable with each other. To see that any svc is not maximally/minimally rational, note that the upper bound $\overline{\rho}^*_{i_X}$ is always strictly positive, because any i_X such that $i \in \{2, \ldots, |X| - 1\}$ contains at least one alternative that it is strictly preferred, and hence chosen with strictly positive probability. This shows that $\iota_r(\rho) \in (0, 1)$ for any ρ whenever $\succeq$ is complete. This shows that svc is not maximally/minimally rational by Proposition 7.

Proof of Theorem 3

Proof. I start by showing the if part. For this, I will use the fact that Luce rule is maximally rational as shown in OT. I am going to show that svc is Luce rule iff all alternatives are comparable. For the if part, assume no alternatives are (strictly) comparable. When this is the case, $D(x,M) = \emptyset$ for any $x \in X$, which implies that:

$$\mathbf{p}^{s}_{(M,x)}(y) = \begin{cases} 1 & \text{if} \quad y = x \\ 0 & \text{if} \quad y \neq x \end{cases}$$

So, the choice probabilities are fully determined by the fixation probabilities, and hence:

$$\rho_y(M) = \frac{u(y)}{u(M)}$$

for any $M \ni y$. For the only if part, assume that svc is Luce, i.e. $\rho_y(M) = \frac{u(y)}{u(M)}$ for any $y \in M$ and $M \in \mathbb{X}$. In particular, consider a binary menu $\{x,y\}$. By definition of svc, $\rho_y(xy) > 0$ implies that $x \bowtie y$. This implies that all alternatives are incomparable.

For the second case, assume that there is a unique alternative that dominates the rest of the alternatives, say x, and the rest of the alternatives are incomparable. First, I show that this is maximally rational. Because $x \succ y$ for all $y \neq x$ and there are no paths, neither $y \succ z$ nor $z \succ y$ for any $y,z \in X \setminus \{x\}$. Thus, $y \bowtie z$ for any such y,z. Because MST holds, only Chernoff and Condorcet axioms need to be checked. In particular, I am going to check the stochastic analogues of these conditions shown in OT, which they operationalize using the set-theoretic approach provided above.

Consider any menu S,T such that $S \subseteq T$ and $x \notin T$. So, S and T consists of incomparable pairs, which implies that both Chernoff and Condorcet are satisfied because the rcf restricted to such menus reduces to Luce rule as shown before, which is maximally rational. Now assume that $x \in T$. I will first check Chernoff axiom. Since x dominates all other alternatives and otherwise alternatives are incomparable, $\rho_x(T) = 1$ and $\rho_y(T) = 0$ for any $y \neq x$, so the choice probabilities are deterministic. For any S such that $x \in S$, the same deterministic choice holds, and therefore Chernoff is not violated. If $x \notin S$, then Chernoff is not violated since $\rho_y(T) = 0$ and $\rho_y(S) > 0$ for any $y \neq x$.

Now I need to check the Condorcet axiom. Take a menu S with $x \in S$: $\rho^*_s(st) > 0$ iff $s = x$ or $x \notin \{s,t\}$. In the former case, $\rho^*_x(xy) = 1$ for any $y \in S$, and also $\rho^*_x(S) = 1$, which shows Condorcet holds. In the latter case:

$$\rho^*_y(yz) = \begin{cases} \frac{u(y)}{u(z)} & \text{if} \quad u(y) < u(z) \\ 1 & \text{if} \quad u(y) > u(z) \end{cases}$$

because the remaining alternatives excluding x are incomparable. If S includes an alternative z such that $u(y) < u(z)$, then $\rho^*_y(S) = \frac{u(y)}{u(z^*)}$ where z^* is the alternative with highest u value in S. Note that $\rho^*_y(yz^*) = \frac{u(y)}{u(z^*)}$ and this is the lowest value y's choice probability takes in a binary menu with alternatives from S. This shows that Condorcet is satisfied. If there is no such z, then $\rho^*_y(S) = 1$. Thus, Condorcet is also satisfied in this case. Therefore, svc is maximally rational also in this case.

Now, I show the equivalence between this case and almost Luce rule. Since x dominates all other alternatives, $\rho_x(T) = 1$ and $\rho_y(T) = 0$ for any $y \neq x$, so the choice probabilities are deterministic for any menu that contains x. Because the remaining alternatives excluding x are incomparable, for any S such that $x \notin S$, the choice probabilities are $\rho_y(S) = \frac{u(y)}{u(S)}$ for any $y \in S$. This shows that this case results in almost Luce. For the converse direction, assume svc is almost Luce, and consider any binary menu including x, say $\{x,y\}$. Since $\rho_x(xy) = 1$ for any such binary menu and by definition of svc, it follows that $x \succ y$. The same holds for all such menus, and hence $x \succ y$ for any $y \neq x$. For any $\{y,z\}$ distinct from x, $\rho_y(yz), \rho_z(yz) \in (0,1)$, so $y \bowtie z$.

For the only if part, I am going to show the contrapositive. Let $\mathscr{P}_{\succ}$ be the set of all (strictly) comparable pairs. The contrapositive statement corresponds to the following: if there is at least one (strictly) comparable pair ($|\mathscr{P}_{\succ}| \geq 1$) and either there is no unique dominating alternative x such that $x \succ y$ for all $y \neq x$ (because there cannot be more than one dominating alternative) or even if there is such an x there exists a distinct pair (y,z) of alternatives that are comparable, then svc is not maximally rational. Thus, assume that $|\mathscr{P}_{\succ}| \geq 1$.

I start by showing if there is at least one path of the form $x \succ y \succ z$ for some x,y,z, then svc is not maximally rational (which also implies that if $\succeq$ is complete, then svc is not maximally rational, already shown in the previous result). Consider the path $x \succ y \succ z$. Let $S = \{x,y\}$ and $T = \{x,y,z\}$. This implies that $\rho_y(T) = \frac{u(z)}{u(T)} \cdot \frac{v(y)}{v(xy)}$ and hence $\rho^*_y(T) > 0$. On the other hand, $\rho_y(S) = 0$ because $x \succ y$, which implies that $\rho^*_y(S) = 0$. Thus, $y \in C_{\rho,0}(T) \cap S$ but $y \notin C_{\rho,0}(S)$, which violates the Chernoff axiom for $\lambda = 0$. This also implies that if there is a unique dominating alternative with at least one distinct comparable pair, then svc is not maximally rational.

Given the last observation, assume that there is no path of the form above, and also that there is no alternative dominating the rest of the alternatives. Because there is at least a pair that is comparable and there is no dominating alternative, X should contain at least three alternatives. Thus, for any $x \in X$, if $x \succ y$ for some $y \neq x$, then there is some z such that $x \bowtie z$ for which it cannot be the case that $z \succ x$ or $y \succ z$. So, for any z, $x \bowtie z$ and either $y \bowtie z$ or $z \succ y$. Since paths are excluded, it cannot be the case that $x \succ z$ and $z \succ y$. This implies that the following are the possible paths involving these three alternatives:

- **P1** $x \succ y$, $x \bowtie z$ and $y \bowtie z$.
- **P2** $x \succ y$, $x \bowtie z$ and $z \succ y$.

Let me start by demonstrating that **P1** is not maximally rational. Assume first $u(x) \geq u(z)$. This implies that $\rho_x(xz) = \frac{u(x)}{u(xz)} \geq \frac{u(z)}{u(xz)} = \rho_z(xz)$, so $\rho^*_z(xz) = \frac{u(z)}{u(x)}$. Assessing $\rho^*_z(yz)$ depends on the comparison between $u(y)$ and $u(z)$:

$$\rho^*_z(yz) = \begin{cases} 1 & \text{if} \quad u(z) > u(y) \\ \frac{u(z)}{u(y)} & \text{o.w.} \end{cases}$$

For the trinary menu $\{x,y,z\}$, $\rho_x(xyz) = \frac{u(xy)}{u(xyz)}$ and $\rho_z(xyz) = \frac{u(z)}{u(xyz)}$. Since $u(x) > u(z)$ and $u(xy) > u(z)$, $\rho^*_z(xyz) = \frac{u(z)}{u(xy)}$. Condorcet condition says that $\rho^*_z(xyz) \geq min\{\rho^*_z(xz), \rho^*_z(xz)\}$, which implies $\rho^*_z(xyz) \geq min\{\frac{u(z)}{u(x)}, \frac{u(z)}{u(y)}\}$. However, this contradicts with the above, and therefore Condorcet is violated when $u(x) > u(z)$.

Now assume that $u(z) > u(x)$. Then:

$$\rho_x^*(xyz) = \begin{cases} 1 & \text{if} \quad u(z) < u(xy) \\ \frac{u(xy)}{u(z)} & \text{o.w.} \end{cases}$$

Since $u(z) > u(x)$, $\rho_x^*(xz) = \frac{u(x)}{u(z)}$. If furthermore $u(z) > u(xy)$, this violates the Chernoff axiom by choosing $\lambda \in (\frac{u(x)}{u(z)}, \frac{u(xy)}{u(z)})$. Otherwise, $\rho_z^*(xyz) < 1$ because $\rho_x(xyz) > \rho_z(xyz)$. On the other hand, $\rho_z^*(xz) = 1 = \rho_z^*(yz)$, following because z is incomparable to x and y with $u(z) > u(x) > u(y)$, where the latter inequality holds by the consistency between $\succ$ and u. This violates the Condorcet axiom. Thus, for the case of **P1**, svc is not maximally rational.

Consider the case **P2**. Assume first $u(x) > u(z)$. Note that this and $x \succ y$ implies $\rho_x^*(xz) = 1 = \rho_x^*(xy)$. If furthermore $\rho_x^*(xyz) < 1$, this violates the Condorcet axiom. Assume that this is not the case, so $\rho_x^*(xyz) = 1$, which implies that $\rho_z^*(xyz) < 1$. In particular:

$$\rho_z^*(xyz) = \frac{u(z)v(xz) + u(y)v(z)}{u(x)v(xz) + u(y)v(x)}$$

This expression is strictly larger than $\frac{u(z)}{u(x)}$, which is equal to $\rho_z^*(xz)$. To see this, note that the former is greater than or equal to the latter iff $u(x)v(z) \geq v(x)u(z)$. Since $x \bowtie z$ and (u, v) is strongly compatible, $u(x) > u(z)$ implies that $v(z) > v(x)$. Thus, the inequality holds strictly. But then there is λ that lies strictly between $\frac{u(z)}{u(x)}$ and $\frac{u(z)v(xz)+u(y)v(z)}{u(x)v(xz)+u(y)v(x)}$, which contradicts the Chernoff axiom since $z \in C_\lambda(xyz) \cap \{xz\}$ but $z \notin C_\lambda(xz)$. If $u(z) > u(x)$, then one can show that Chernoff is violated in the same way. This implies that $\rho_x^*(xz) = \frac{u(x)}{u(z)}$ and $\rho_z^*(xz) = 1$. If $\rho_x^*(xyz) = 1$, this contradicts Chernoff because $x \in C_\lambda(xyz) \cap \{xz\}$ but $x \notin C_\lambda(xz)$ for some $\lambda > \frac{u(x)}{u(z)}$. Otherwise, $\rho_x^*(xyz) < 1$, that is, $\rho_x(xyz) < \rho_z(xyz)$. This implies that:

$$\rho_x^*(xyz) = \frac{u(x)v(xz) + u(y)v(x)}{u(z)v(xz) + u(y)v(z)}$$

Similar to the above, one can show that this expression is strictly larger than $\rho_x^*(xz) = \frac{u(x)}{u(z)}$, which follows because this is true iff $u(z)v(x) \geq v(z)u(x)$, and this is implied since $x \bowtie z$ and (u, v) is strongly compatible. This concludes the proof. □

□

A.3 Proof for the Characterization of SVC

Necessity of the Conditions:

Lemma 3. *If ρ can be represented by svc, then it satisfies Condition 9.*

Proof. If adding z violates regularity, then $\frac{\rho_x(M)}{\rho_x(M \setminus \{z\})} > 1 > \frac{1-\rho_x(M)}{1-\rho_x(M \setminus \{z\})}$, and therefore the minimum value is equal to $\frac{1-\rho_x(M)}{1-\rho_x(M \setminus \{z\})}$. Note that

$$\begin{aligned} &\frac{g_x(M)}{g_x(M \setminus \{z\})} \cdot (1 - \rho_x(M \setminus \{z\})) \\ &= \frac{u(M \setminus \{z\})}{u(M)} \left(1 - \sum_{y \in M \setminus \{z\}: x \succeq y} \frac{u(y)}{u(M \setminus \{z\})} \frac{v(x)}{v(D(y, M))}\right) \end{aligned} \tag{4}$$

because $\frac{g_x(M)}{g_x(M\setminus\{z\})} = \frac{u(M\setminus\{z\})}{u(M)}$ and ρ has svc representation. The right-hand side of this equality simplifies to:

$$\frac{u(M\setminus\{z\})}{u(M)} - \sum_{y\in M\setminus\{z\}:x\succeq y} \frac{u(y)}{u(M)}\frac{v(x)}{v(D(y,M))}$$

The second term of this equation is equal to $\rho_x(M) - \frac{u(z)}{u(M)}\cdot\frac{v(x)}{v(D(z,M))}$ because regularity violation implies that x strictly dominates z, which implies the following final equality:

$$\frac{g_x(M)}{g_x(M\setminus\{z\})}\cdot(1-\rho_x(M\setminus\{z\})) = \frac{u(M\setminus\{z\})}{u(M)} + \frac{u(z)}{u(M)}\cdot\frac{v(x)}{v(D(z,M))} - \rho_x(M)$$

This expression is less than or equal to $1-\rho_x(M)$. On the other hand, when regularity is not violated, the minimum value is equal to $\frac{\rho_x(M)}{\rho_x(M\setminus\{z\})}$ and $\frac{g_x(M)}{g_x(M\setminus\{z\})} \leq \frac{\rho_x(M)}{\rho_x(M\setminus\{z\})}$ needs to be shown. Note that by Equation 4, the following equality holds:

$$\frac{g_x(M)}{g_x(M\setminus\{z\})}\cdot\rho_x(M\setminus\{z\}) = \frac{u(M\setminus\{z\})}{u(M)}\cdot\sum_{y\in M\setminus\{z\}:x\succeq y}\frac{u(y)}{u(M\setminus\{z\})}\frac{v(x)}{v(D(y,M))}$$

which is easily seen to be strictly less than $\rho_x(M)$ if $x\succ z$, and equal if $x\bowtie z$. □

The proof will proceed in several steps.

1: Revealed dominance relation is a partial order.

I start by showing that the revealed preferences, denoted $\succeq$, is a partial order. By definition, $x\succeq x$ because $\rho_x(x)=1$. To show transitivity, assume that $x\succeq y\succeq z$. By definition, $\rho_x(xy)=1=\rho_y(yz)$. Because of Condition 4, this implies that $\rho_x(xz)=1$, so $x\succeq z$ and hence $\succeq$ is transitive. Finally, to show antisymmetry, let $x\succeq y\succeq x$, which implies that $\rho_x(xy)=1=\rho_y(xy)$. This is only possible when $x=y$, and hence $\succeq$ is antisymmetric. This shows that $\succeq$ is a partial order.

Given $\succeq$, define $D(y,M)=\{x\in M : x\succ y\}$.

2: Useful Lemmas

Condition 11. ***Strong Rationalizability***
$\rho_x(M)=1$ *iff* $\rho_x(xy)=1$ *and* $\rho_y(yz)\in(0,1)$ *for all* $y,z\in M\setminus\{x\}$.

Lemma 4. *Rationalizability and Dominance Transitivity implies Strong Rationalizability.*

Proof. Assume first $\rho_x(M)=1$. This holds iff $\rho_y(M)=0$ for all $y\neq x$ in M. The contrapositive of the if part of the rationalizability condition says that for any $y,z\in M\setminus\{x\}$ such that $y\neq z$, $\rho_y(yz)<1$ and there is a such that $\rho_y(ya)=0$. The latter implies that for any such y, there is a that dominates y. Since the latter holds for any $y\in M\setminus\{x\}$, this leads to a cycle unless x dominates all, which cannot be the case by dominance transitivity and the fact that it implies $\succeq$ is transitive. But then it should be the case that $\rho_x(xy)=1$ for all $y\in M\setminus\{x\}$, that is, $\rho_y(xy)=0$ for all $y\in M\setminus\{x\}$. For the former, note that $\rho_y(yz)=0$ implies that $\rho_z(yz)=1$, but then $\rho_z(M)>0$ by rationalizability leading to a contradiction because $\rho_x(M)=1$. Therefore, $\rho_y(yz)\in(0,1)$, concluding the only if part of the proof. For the if part, assume that $\rho_x(xy)=1$ and $\rho_y(yz)\in(0,1)$ for all $y,z\in M\setminus\{x\}$. The former holds iff $\rho_y(xy)=0$ for any $x\neq y$, so none of the conditions for the rationalizability is satisfied. The contrapositive of the if part then implies that $\rho_y(M)=0$ for all $y\neq x$ in M, and therefore $\rho_x(M)=1$.

□

Condition 12. ***Weak Positivity***
$\rho_x(xy) \in (0,1)$ for any $\{x,y\} \subseteq M$ iff $\rho_x(M) \in (0,1)$ for all $x \in M$.

Lemma 5. *Rationalizability and Dominance Transitivity implies Weak Positivity.*

Proof. Assume first $\rho_x(xy) \in (0,1)$ for all $\forall x,y \in M$. This directly implies $\rho_x(M) > 0$ by rationalizability for all x. For the other side, assume $\rho_x(M) \in (0,1)$ for all $x \in M$. By rationalizability, either there is an alternative $y \in M \setminus \{x\}$ such that $\rho_x(xy) = 1$ or $\rho_x(xy) \in (0,1)$ for all $y \in M \setminus \{x\}$. Consider a chain of alternatives that satisfy the former condition, say $\{m_1, m_2, \ldots, m_k\} \subseteq M$ with $\rho_{m_i}(m_i m_{i+1}) = 1$ for any $i \leq k-1$. Assume that this is the largest chain that includes $\{m_1, m_2, \ldots, m_k\}$, so there cannot be a subset S such that $\{m_1, m_2, \ldots, m_k\} \subset S \subseteq M$ and the alternatives in S can be ordered in a chain. By definition, $\rho_{m_k}(m_k m_i) = 0$ for any $i \in \{1, \ldots, k\}$. Since this is a largest chain, m_k should be incomparable to the rest of the alternatives, so $\rho_{m_k}(x m_k) \in (0,1)$ for any $x \in M \setminus \{m_1, m_2, \ldots, m_k\}$. Thus, m_k does not satisfy any features in the rationalizability condition, so $\rho_{m_k}(M) = 0$. This is a contradiction to positivity on M, and therefore there cannot be such a chain. This concludes the proof because only the second condition for rationalizability can be satisfied for all alternatives in M.

□

Lemma 6. *Condition 8 and Condition 9 imply that if $\rho_z(M) = 0$ and $\rho_y(yz) \in (0,1)$, then the following holds:*

$$\rho_y(M \setminus \{z\}) - \rho_y(M) = g_z(M) \cdot \rho_y(M \setminus \{z\})$$

Proof. Because $\rho_z(M) = 0$ and $\rho_y(yz) \in (0,1)$, Condition 9 implies that:

$$\frac{g_y(M)}{g_y(M \setminus \{z\})} = \frac{\rho_y(M)}{\rho_y(M \setminus \{z\})}$$

Also, Condition 8 says that:

$$1 - g_z(M) = \frac{g_y(M)}{g_y(M \setminus \{z\})}$$

for any distinct $y, z \in M$. These two equalities imply:

$$1 - g_z(M) = \frac{\rho_y(M)}{\rho_y(M \setminus \{z\})}$$

which implies the conclusion. □

3: Constructing Fixation Value

Lemma 7. *Assume that ρ satisfies Conditions 4 and 7. Then, there exists a function u which can be derived from binary choice probabilities.*

Proof of Lemma 7

I start by constructing the fixation probabilities. Let $\mathscr{P}$ be a partition of X such that x and y are in the same cell of the partition iff there is an incomparability path $\{z_i\}_1^k$ from x to y such that $z_1 \equiv x$ and $z_k = y$ with x and y being incomparable too. Take a cell in the partition and choose some x, and let $u(x) = 1$. Since z_i and z_{i+1} are incomparable to each other, one can define:

$$g_{z_i}(z_i z_{i+1}) := \rho_{z_i}(z_i z_{i+1})$$

Given these, one can define $u(y)$ for y in the same cell of the partition as:

$$u(y) = \frac{g_{z_2}(z_1 z_2)}{g_{z_1}(z_1 z_2)} \cdot \frac{g_{z_3}(z_2 z_3)}{g_{z_2}(z_2 z_3)} \cdot \dots \cdot \frac{g_{z_k}(z_{k-1} z_k)}{g_{z_{k-1}}(z_{k-1} z_k)}$$

This can be shown to be well-defined. First of all, $u(y)$ is a strictly positive real number because $\{z_i\}_1^k$ is an incomparability path. Secondly, different incomparability paths lead to the same result by Condition 7. Consider another incomparability path $\{z'_i\}_1^L$ such that $z'_1 \equiv x$ and $z'_l = y$. Condition 7 implies the following:

$$1 = \frac{g_{z_2}(z_1 z_2)}{g_{z_1}(z_1 z_2)} \cdot \frac{g_{z_3}(z_2 z_3)}{g_{z_2}(z_2 z_3)} \cdot \dots \cdot \frac{g_{z_k}(z_{k-1} z_k)}{g_{z_{k-1}}(z_{k-1} z_k)} \cdot \frac{g_{z'_{l-1}}(z'_{l-1} z'_l)}{g_{z'_l}(z'_{l-1} z'_l)} \cdot \frac{g_{z'_{l-2}}(z'_{l-2} z'_{l-1})}{g_{z'_{l-1}}(z'_{l-2} z'_{l-1})} \cdot \dots \cdot \frac{g_{z'_1}(z'_1 z'_2)}{g_{z'_2}(z'_1 z'_2)}$$

which shows that

$$\frac{g_{z_1}(z_1 z_2)}{g_{z_2}(z_1 z_2)} \cdot \frac{g_{z_2}(z_2 z_3)}{g_{z_3}(z_2 z_3)} \cdot \dots \cdot \frac{g_{z_{k-1}}(z_{k-1} z_k)}{g_{z_k}(z_{k-1} z_k)} = \frac{g_{z'_1}(z'_1 z'_2)}{g_{z'_2}(z'_1 z'_2)} \cdot \frac{g_{z'_2}(z'_2 z'_3)}{g_{z'_3}(z'_2 z'_3)} \dots \frac{g_{z'_{l-1}}(z'_{l-1} z'_l)}{g_{z'_l}(z'_{l-1} z'_l)}$$

and hence u is well-defined. Let $g_x(M)$ be defined as:

$$g_x(M) := \frac{u(x)}{u(M)}$$

for any menu M and $x \in M$.

4: Constructing Saccade Probabilities

Consider the following connectedness relation: x, y are said to be connected to each other if both strictly dominates some distinct z in some menu M with $\rho_z(M) = 0$. Take the transitive closure of this relation, and collect all alternatives that are connected to each other according to this relation. It is easy to see that this relation is reflexive and symmetric, and furthermore transitive because transitive closure is taken. This constitutes a partition. Consider any alternative x in a cell of this partition, and let $v(x) = 1$. $v(y)$ can be defined for any y in the same cell by considering the path that connects x and y. Let $\{t_i\}_{i=1}^k$ be a path that connects these such that $t_1 \equiv x$ and $t_k \equiv y$, and M_i be the corresponding menu in which one can find a strictly dominated alternative by t_i and t_{i+1}. Define the following function:

$$h_x(z, M) := \left(\frac{1}{g_z(M)}\right) \rho_x(M) - \left(\frac{1 - g_z(M)}{g_z(M)}\right) \rho_x(M \setminus \{z\})$$

Then, one can define:

$$v(y) := \frac{h_{t_1}(z_1, M_1)}{h_{t_2}(z_1, M_1)} \cdot \frac{h_{t_2}(z_2, M_2)}{h_{t_3}(z_2, M_2)} \cdot \dots \frac{h_{t_{k-1}}(z_{k-1}, M_{k-1})}{h_{t_k}(z_{k-1}, M_{k-1})}$$

The construction of the saccade values are like the construction of fixation value u, and proof proceeds along similar lines with the difference that now Condition 10 is used to show it. This condition shows only that two different paths lead to the same result, however, one needs to show also that h is well-defined in order to have a well-defined v. First, note that:

$$\left(\frac{1}{g_z(M)}\right) \rho_x(M) - \left(\frac{1 - g_z(M)}{g_z(M)}\right) \rho_x(M \setminus \{z\}) \geq 0$$

iff:

$$\left(\frac{1}{g_z(M)}\right)\rho_x(M) \geq \left(\frac{1-g_z(M)}{g_z(M)}\right)\rho_x(M\setminus\{z\})$$

which is equivalent to:

$$\frac{\rho_x(M)}{\rho_x(M\setminus\{z\})} \geq 1-g_z(M)$$

This is implied by Condition 8 and Condition 9. On the other hand:

$$\left(\frac{1}{g_z(M)}\right)\rho_x(M) - \left(\frac{1-g_z(M)}{g_z(M)}\right)\rho_x(M\setminus\{z\}) \leq 1$$

iff:

$$\left(\frac{1}{g_z(M)}\right)\rho_x(M) \leq \frac{g_z(M)+(1-g_z(M))\rho_x(M\setminus\{z\})}{g_z(M)}$$

equivalent to:

$$\rho_x(M) \leq \rho_x(M\setminus\{z\}) + g_z(M)(1-\rho_x(M\setminus\{z\}))$$

This holds true iff:

$$\frac{\rho_x(M)-\rho_x(M\setminus\{z\})}{1-\rho_x(M\setminus\{z\})} \leq g_z(M)$$

By Condition 8, $g_z(M) = 1 - \frac{g_x(M)}{g_x(M\setminus\{z\})}$, which further implies that this condition holds when:

$$\frac{1-\rho_x(M)}{1-\rho_x(M\setminus\{z\})} \geq \frac{g_x(M)}{g_x(M\setminus\{z\})}$$

which is also implied by the same condition. So, h is well-defined. From now on, define the saccade probabilities as:

$$\phi_x(y,M) := \begin{cases} \frac{v(x)}{v(\hat{D}(y,M))} & if \quad x \in \hat{D}(y,M) \\ 0 & o.w. \end{cases}$$

The following lemma will be used in the following parts.

Lemma 8. *If $|D(y,M)| > 1$ and $\rho_y(M) = 0$, then $\phi_x(y,M) = h_x(y,M)$.*

Proof. To define $\phi_x(y,M)$, the construction of v is used. Since $|D(y,M)| > 1$, the cell that contains x also contains some z that dominates y. By definition, $\phi_x(y,M) = \frac{v(x)}{v(D(y,M))}$. For the construction, assume wlog that z is the fixed alternative in the cell and let $v(x) = \frac{h_x(y,M)}{h_z(y,M)}$. This implies that:

$$\phi_x(y,M) = \frac{\frac{h_x(y,M)}{h_z(y,M)}}{\sum_{x'\in D(y,M)} \frac{h_{x'}(y,M)}{h_z(y,M)}}$$

One can directly cancel $h_z(y,M)$ from both sides, implying that showing the claim $\phi_x(y,M) = h_x(y,M)$ is equivalent to showing $\sum_{x'\in D(y,M)} h_{x'}(y,M) = 1$. By definition of h, this is equivalent to showing:

$$1 = \left(\frac{1}{g_y(M)}\right)\sum_{x\in D(y,M)} \rho_x(M) - \left(\frac{1-g_y(M)}{g_y(M)}\right)\sum_{x\in D(y,M)} \rho_x(M\setminus\{y\})$$

The right-hand side of this can written in the following way:

$$\left(\frac{1}{g_y(M)}\right)\sum_{x\in D(y,M)}[\rho_x(M)-\rho_x(M\setminus\{y\})]+\sum_{x\in D(y,M)}\rho_x(M\setminus\{y\})$$

Since $\rho_y(M)=0$, y is either dominated by or incomparable to another alternative in M. Therefore, the change in probabilities due to $D(y,M)$ is balanced by the change in probabilities of alternatives not in $D(y,M)$, and these alternatives are the alternatives that are incomparable to y. Thus, it follows that:

$$\left(\frac{1}{g_y(M)}\right)\sum_{x\in M:x\bowtie y}[\rho_x(M\setminus\{y\})-\rho_x(M)]+\sum_{x\in D(y,M)}\rho_x(M\setminus\{y\})$$

Since $\rho_y(M)=0$ and $\rho_x(xy)\in(0,1)$ for any $x\bowtie y$ by the definition of $\bowtie$, Lemma 6 implies that $\rho_x(M\setminus\{y\})-\rho_x(M)=g_y(M)\cdot\rho_x(M\setminus\{y\})$. Therefore, the above expression becomes:

$$\left(\frac{1}{g_y(M)}\right)\sum_{x\in M:x\bowtie y}g_y(M)\cdot\rho_x(M\setminus\{y\})+\sum_{x\in D(y,M)}\rho_x(M\setminus\{y\})$$

which is equal to:

$$\sum_{x\in M:x\bowtie y}\rho_x(M\setminus\{y\})+\sum_{x\in D(y,M)}\rho_x(M\setminus\{y\})$$

and this is equal to 1 because $\rho_y(M)=0$ and all alternatives are either incomparable to or dominates y. $\square$

This shows that whenever $|D(y,M)|>1$ and $x\in D(y,M)$, one can use directly the h-function to define the saccade probability of x:

$$\phi_x(y,M)=\left(\frac{1}{g_z(M)}\right)\rho_x(M)-\left(\frac{1-g_z(M)}{g_z(M)}\right)\rho_x(M\setminus\{z\})$$

Representation if ρ satisfies positivity

Let me start by considering the case at which $\rho_x(M)>0$ for all $x\in M$. Note that by Condition 12, positivity on M implies positivity is satisfied in binary menus. Thus, one has:

$$\frac{g_x(xy)}{g_y(xy)}=\frac{\rho_x(xy)}{\rho_y(xy)}$$

and by what is shown before:

$$\frac{g_x(xy)}{g_y(xy)}=\frac{u(x)}{u(y)}$$

Consider $z\in M$, by positivity on binary menus, Condition 6 implies that $\frac{\rho_x(M)}{\rho_y(M)}=\frac{\rho_x(M\setminus\{z\})}{\rho_y(M\setminus\{z\})}$. One can repeatedly apply this condition to reach:

$$\frac{\rho_x(M)}{\rho_y(M)}=\frac{\rho_x(xy)}{\rho_y(xy)}=\frac{u(x)}{u(y)}$$

which implies that:

$$\frac{1-\rho_x(M)}{\rho_x(M)}=\frac{\sum_{y\in M\setminus\{x\}}u(y)}{u(x)}$$

concluding the proof.

Representation for menus with at most 3 alternatives

Consider any $M = \{x,y,z\}$ which does not necessarily involve distinct alternatives but at least two alternatives, since otherwise the characterization is trivial. Recall that $\phi_x(y,M)$ denote the revealed saccade probability from y to x in menu M. For binary menus, either the choice probabilities are deterministic or both interior. Consider menu $M = \{x,y\}$ and assume wlog $\rho_x(xy) = 1$. This shows $x \succ y$, and the representation holds also in this case. Finally, assume $\rho_x(xy) \in (0,1)$. In this case, by the definition of saccade probabilities, $\phi_s(s,xy) = 1$ and $\phi_s(t,xy) = 0$ when $s \neq t$. The fixation probability can be directly identified by comparing x and y as it is done previously, so it is given by $g_s(xy) = \rho_s(xy)$. But then $\rho_s(xy) = g_s(xy) \cdot \phi_s(s,xy)$, so the representation holds in this case too. Now assume $|M| = 3$. The case of $\rho_s(M) > 0$ for all $s \in M$ is implied by the proof of the representation for ρ satisfying positivity, so I will assume that there is at least one alternative chosen with probability 0.

Case 1: $\rho_s(M) = 1$**.**

This implies that $\rho_t(M) = 0$ for any $t \in M \setminus \{s\}$. By Condition 11, $\rho_s(sy) = 1$ for any $y \in M \setminus \{s\}$, which implies that $s \succ y$ for all such y. Therefore, the definition of saccade probability implies that $\phi_s(t,M) = 1$ for any $t \in M$, and equals to 0 for all alternatives except 0 also by definition. This implies that the representation holds in this case.

Case 2: $\rho_s(M) = 0$ **for a unique** $s \in M$**, positive otherwise.**

Condition 12 shows that not all alternatives are incomparable. This implies that there is at least one comparable pair. Assume that $s = z$. It cannot be the case that z dominates one of the remaining alternatives by Condition 5, so $\rho_z(xz), \rho_z(yz) < 1$. Furthermore, it cannot be the case that $\rho_z(xz), \rho_z(yz) > 0$ again by Condition 12. Hence, at least one of the remaining alternatives dominates z and z dominates none. Assume wlog x dominates z, that is, $\rho_x(xz) = 1$. There are several subcases.

- (Subcase 1: $\rho_y(yz) = 1$)

  There are three subcases given $\rho_y(yz) = 1$. It is possible that $\rho_x(xy) = 1$, $\rho_y(xy) = 1$ or $\rho_x(xy), \rho_y(xy) \in (0,1)$. First two cases are symmetric to each other, so assume wlog that $\rho_x(xy) = 1$. So, $\rho_y(xy) = 0$. This reveals that $x \succ y \succ z$. Applying the definition ϕ for x results in:

$$\begin{aligned}\phi_x(z,M) &= \left(\frac{1}{g_z(M)}\right)\rho_x(M) - \left(\frac{1-g_z(M)}{g_z(M)}\right)\rho_x(M\setminus\{z\}) \\ &= \frac{\rho_x(M)}{g_z(M)}\end{aligned}$$

  because $\rho_y(xz) = 0$. Similarly, one can get:

$$\phi_y(z,M) = \frac{\rho_y(M)}{g_z(M)}$$

  The definition directly implies that $\phi_x(y,M) = 1$, $\phi_y(x,M) = 0 = \phi_y(y,M)$ and finally $\phi_z(s,M) = 0$ for $s \in M$. Since $\rho_z(M) = 0$, $\rho_x(M) = 1 - \rho_y(M)$. By definition, $1 = g_x(M) + g_y(M) + g_z(M)$, which implies that $\rho_x(M) = g_x(M) + g_y(M) + g_z(M) - \rho_y(M)$. Since $\phi_x(s,M) = 1$ for $s \neq z$ and $\phi_x(z,M) = \frac{g_z(M)-\rho_y(M)}{g_z(M)}$, this is equivalent to:

$$\rho_x(M) = g_x(M)\cdot\phi_x(x,M) + g_y(M)\cdot\phi_x(y,M) + g_z(M)\cdot\phi_x(z,M)$$

which shows that $\rho_x(M)$ can be represented by svc. It is straightforward to see the same for $\rho_y(M)$, because $\rho_y(M) = g_z(M) \cdot \frac{\rho_y(M)}{g_z(M)} = g_z(M) \cdot \phi_y(z,M)$. Finally, $\rho_z(M) = 0$ by assumption, and it can be represented by svc. These show that the representation holds for this case.

Now assume that $\rho_x(xy) \in (0,1)$. Since by assumption $\rho_x(xz) = 1$, this implies $x, y \succ z$ and $x \bowtie y$. Since $x \bowtie y$, $\rho_x(xy) = g_x(xy)$. Applying the definition:

$$\begin{aligned}
\phi_x(z,M) &= (\frac{1}{g_z(M)})\rho_x(M) - (\frac{1-g_z(M)}{g_z(M)})\rho_x(M \setminus \{z\}) \\
&= (\frac{1}{g_z(M)})\rho_x(M) - (\frac{1-g_z(M)}{g_z(M)})g_x(xy) \\
&= \frac{\rho_x(M) - g_x(xy)(1-g_z(M))}{g_z(M)}
\end{aligned}$$

By Condition 8, $1 - g_z(M) = \frac{g_x(M)}{g_x(M \setminus \{z\})}$. This implies that $g_x(xy)(1-g_z(M)) = g_x(M)$, and therefore:

$$\phi_x(z,M) = \frac{\rho_x(M) - g_x(M)}{g_z(M)}$$

The same arguments apply to y, so:

$$\phi_y(z,M) = \frac{\rho_y(M) - g_y(M)}{g_z(M)}$$

This implies:

$$\begin{aligned}
&g_x(M) + g_y(M) \cdot \phi_x(y,M) \\
&= g_x(M) + g_y(M) \cdot \frac{\rho_x(M) - g_x(M)}{g_y(M)} \\
&= \rho_x(M)
\end{aligned}$$

and similarly for $\rho_y(M)$. The case for $\rho_z(M) = 0$ is trivial. So, the representation also holds in this case.

- (Subcase 2: $\rho_y(yz) \in (0,1)$)

  This implies that y and z revealed to be incomparable to each other. Furthermore, observe that $\rho_x(xy) \in (0,1)$. To see, note that if $\rho_y(xy) = 1$, then $\rho_y(yz) = 1$ by Condition 4, a contradiction. If $\rho_x(xy) = 1$, then $\rho_y(M) = 0$ by Condition 5. Since $\rho_z(M) = 0$ and it is revealed that it is dominated by x with being incomparable to y, the representation holds for z. Because y is incomparable to the rest, it can be represented as in Case 2. Finally, $\rho_x(M) = 1 - \rho_y(M)$ because $\rho_z(M) = 0$, and since $\rho_y(M) = g_y(M)$,

  $$\rho_x(M) = g_x(M) + g_z(M) = g_x(M) \cdot \phi_x(x,M) + g_z(M) \cdot \phi_x(z,M)$$

  where both $\phi_x(x,M)$ and $\phi_x(z,M)$ are equal to 1 by definition, concluding the proof.

Representation for any menu

The svc representation is shown for all menus of size at most 3. I am going to show that this holds generally using induction. So, assume that the representation holds for all menus with at most k cardinality for some $k \geq 3$, and consider now a menu M such that $|M| = k+1$. There are several cases to consider.

Since the proof for ρ that satisfies positivity is shown, let me assume that $\rho_s(M) = 0$ for some $s \in M$. First, consider the case when $\rho_x(M) = 1$ for some $x \in M$. By Condition 11, $\rho_x(xy) = 1$ and $\rho_y(yz) \in (0,1)$ for all $y, z \in M \setminus \{x\}$. The former implies that $x \succ y$ for all $y \in M \setminus \{x\}$, while the latter implies that $y \bowtie z$ for all $y, z \in M \setminus \{x\}$. These show that the svc representation holds. Any $y \in M \setminus \{x\}$ is dominated by x and does not dominate any products, so $\rho_y(M) = 0$ is represented by svc. Similarly, x dominates all products, so $\rho_x(M) = 1$ is represented by svc.

Therefore, assume further that $\rho_x(M) < 1$ for all $x \in M$ with the previous assumption that $\rho_s(M) = 0$ for some $s \in M$. Let me start by assuming there is a unique alternative s for which $\rho_s(M) = 0$. This can be represented using svc. To see, note that the contrapositive of the if part of Condition 5 implies that s is dominated by at least one product and does not dominate any products. Hence, the representation follows trivially. Note that Condition 5 implies that $\rho_x(M) > 0$ for some x only if x dominates another alternative ($\rho_x(xy) = 1$ for some $y \in M$) or x is incomparable to all the remaining alternatives in M ($\rho_x(xy) \in (0,1)$ for all $y \in M \setminus \{x\}$). This implies that M either is a complete chain or consists of a chain and alternatives that are incomparable to the rest. First, assume that the latter exists. Let z be an alternative that is incomparable to the rest in M. By Condition 6:

$$\frac{\rho_x(M)}{\rho_y(M)} = \frac{\rho_x(M \setminus \{z\})}{\rho_y(M \setminus \{z\})}$$

for any x, y such that $\rho_x(M), \rho_y(M) > 0$. Assume wlog that $\rho_x(xy) = 1$, so x is revealed to dominate y. By the induction assumption, $\rho_x(M \setminus \{z\})$ and $\rho_y(M \setminus \{z\})$ have svc representations. Furthermore, note that $D(a, M) = D(a, M \setminus \{z\})$ given z is incomparable to other products. Thus:

$$\begin{aligned}
\frac{\rho_x(M)}{\rho_y(M)} &= \frac{\sum_{x \succeq a} \frac{u(a)}{u(M \setminus \{z\})} \cdot \frac{v(x)}{v(D(a, M \setminus \{z\})}}{\sum_{y \succeq a} \frac{u(a)}{u(M \setminus \{z\})} \cdot \frac{v(y)}{v(D(a, M \setminus \{z\})}} \\
&= \frac{v(x)}{v(y)} \cdot \frac{\sum_{x \succeq a} \frac{u(a)}{v(D(a, M \setminus \{z\}))}}{\sum_{y \succeq a} \frac{u(a)}{v(D(a, M \setminus \{z\}))}} \\
&= \frac{\frac{v(x)}{u(M)}}{\frac{v(y)}{u(M)}} \cdot \frac{\sum_{x \succeq a} \frac{u(a)}{v(D(a, M))}}{\sum_{y \succeq a} \frac{u(a)}{v(D(a, M))}} \\
&= \frac{\sum_{x \succeq a} \frac{u(a)}{u(M)} \cdot \frac{v(x)}{v(D(a, M))}}{\sum_{y \succeq a} \frac{u(a)}{u(M)} \cdot \frac{v(y)}{v(D(a, M))}}
\end{aligned}$$

Summing over all $x \neq y$ in M, one gets:

$$\frac{1 - \rho_y(M)}{\rho_y(M)} = \frac{1 - \sum_{y \succeq a} \frac{u(a)}{u(M)} \cdot \frac{v(y)}{v(D(a, M))}}{\sum_{y \succeq a} \frac{u(a)}{u(M)} \cdot \frac{v(y)}{v(D(a, M))}}$$

which implies that $\rho_y(M) = \sum_{y \succeq a} \frac{u(a)}{u(M)} \cdot \frac{v(y)}{v(D(a, M))}$. Since $\rho_s(M) = 0$ for a unique alternative, there is no alternative that s dominates and s is dominated at least by one alternative in M by Condition 5.

Now assume that there is no alternative that is incomparable to the rest. Consider the menu $M \setminus \{s\}$. By the inductive step, ρ here has a representation in terms of pvc. Because $\rho_s(M) = 0$, there is no alternative that is dominated by s in M: so, any alternative $t \neq s$ in M either dominates s or incomparable to it. Recall the definition of ϕ:

$$\phi_x(s,M) = \begin{cases} 1 & if \quad x \succeq s \ \& \ D(s,M) \subseteq \{x\} \\ (\frac{1}{g_s(M)})\rho_x(M) - (\frac{1-g_s(M)}{g_s(M)})\rho_x(M \setminus \{s\}) & if \quad \{x\} \subset D(s,M) \\ 0 & o.w. \end{cases}$$

By the inductive step, $\rho_x(M \setminus \{s\})$ has a pvc representation:

$$\rho_x(M \setminus \{s\}) = \sum_{x \succeq y: y \in M \setminus \{s\}} g_y(M \setminus \{s\}) \cdot \phi_x(y, M \setminus \{s\})$$

First, consider the case when x is incomparable to s. By Condition 9:

$$\frac{g_x(M)}{g_x(M \setminus \{s\})} = \frac{\rho_x(M)}{\rho_x(M \setminus \{s\})}$$

Also by the same condition:

$$1 - g_s(M) = \frac{g_y(M)}{g_y(M \setminus \{s\})}$$

for any $y \in M \setminus \{s\}$. Note that these imply:

$$(1 - g_s(M))\rho_x(M \setminus \{s\}) = \frac{g_x(M)}{g_x(M \setminus \{s\})}$$

$$\begin{aligned} (1 - g_s(M))\rho_x(M \setminus \{s\}) &= \frac{g_x(M)}{g_x(M \setminus \{s\})}\rho_x(M \setminus \{s\}) \\ &= \frac{\rho_x(M)}{\rho_x(M \setminus \{s\})} \cdot \rho_x(M \setminus \{s\}) \\ &= \rho_x(M) \end{aligned}$$

Also:

$$\begin{aligned} (1 - g_s(M))\rho_x(M \setminus \{s\}) &= (1 - g_s(M)) \sum_{x \succeq y: y \in M \setminus \{s\}} g_y(M \setminus \{s\}) \cdot \phi_x(y, M \setminus \{s\}) \\ &= \sum_{x \succeq y: y \in M \setminus \{s\}} \frac{g_y(M)}{g_y(M \setminus \{s\})} \cdot g_y(M \setminus \{s\}) \cdot \phi_x(y, M \setminus \{s\}) \\ &= \sum_{x \succeq y: y \in M \setminus \{s\}} g_y(M) \cdot \phi_x(y, M \setminus \{s\}) \\ &= \sum_{x \succeq y: y \in M} g_y(M) \cdot \phi_x(y, M) \end{aligned}$$

where the last line follows because s does not dominate any $y \in M \setminus \{s\}$. These show that:

$$\rho_x(M) = \sum_{x \succeq y: y \in M} g_y(M) \cdot \phi_x(y, M)$$

which shows that the representation holds in this case x is incomparable to s.

Now assume that x dominates s. Consider:

$$(1-g_s(M))\rho_x(M\setminus\{s\})+g_s(M)\phi_x(s,M)$$

By the definition of ϕ in this case:

$$\begin{aligned}
&(1-g_s(M))\rho_x(M\setminus\{s\})+g_s(M)\phi_x(s,M)\\
&=(1-g_s(M))\rho_x(M\setminus\{s\})+g_s(M)(\frac{1}{g_s(M)})\rho_x(M)-(\frac{1-g_s(M)}{g_s(M)})\rho_x(M\setminus\{s\})\\
&=(1-g_s(M))\rho_x(M\setminus\{s\})+\rho_x(M)-(1-g_s(M))\rho_x(M\setminus\{s\})\\
&=\rho_x(M)
\end{aligned}$$

Again using the inductive step, $\rho_x(M\setminus\{s\})$ has a representation in terms of pvc. I also showed above that:

$$(1-g_s(M))\rho_x(M\setminus\{s\})=\sum_{x\succeq y:y\in M} g_y(M)\cdot\phi_x(y,M)$$

So:

$$\begin{aligned}
\rho_x(M)&=(1-g_s(M))\rho_x(M\setminus\{s\})+g_s(M)\phi_x(s,M)\\
&=\sum_{x\succeq y:y\in M\setminus\{s\}} g_y(M)\cdot\phi_x(y,M)+g_s(M)\phi_x(s,M)\\
&=\sum_{x\succeq y:y\in M} g_y(M)\cdot\phi_x(y,M)
\end{aligned}$$

which shows that ρ is represented by pvc also in this case. The same logic applies to multiple alternatives with zero choice probability. Therefore, one can apply the same reasoning above and reach the conclusion using inductive step. This concludes the proof.

□

A.4 Further Related Literature

Dutta [12] defined a procedural model which results in stochastic choice relying on studies in eyetracking. The DM is able to do binary comparisons only, and the order of these comparisons depend on the difficulty of the comparisons. Let $\succ$ be the strict preference relation of the DM, and $\mathbb{P}$ be an ordered partition of $\succ$ such that $\mathbb{P}=\{P_i\}_{i=1}^{K}$ and $P_0=\emptyset$. Faced with a menu M, the DM makes all comparisons in the order imposed by $\mathbb{P}$ relevant to M. The following is the recursive formulation used by Dutta:

$$\begin{aligned}
r_0^{\mathbb{P}}(M)&:=M\\
&\vdots\\
r_{i+1}^{\mathbb{P}}(M)&:=\{x\in r_i^{\mathbb{P}}(M):\forall y\in r_i^{\mathbb{P}}(M)\quad (y,x)\notin P_i\}
\end{aligned}$$

Let π be a *stopping function* which models the inability of the DM to complete all binary comparisons. If this was the case, then the procedure would reduce to the deterministic choice. Let $K^*(M)$ be the final cell of the partition that reduces the set of surviving options to a singleton. Define the stopping function as a mapping $\pi:(\mathbb{P}\cup P_0)\to(0,1)$ which is a probability distribution such that $\pi(P_{K^*(M)},M)>0$ for all M. The DM chooses each alternative that survives the iterated elimination with equal probability. The choice rule **gradual pairwise stochastic choice (gpsc)** is then defined as:

$$\rho_x^{gpsc}(M) := \begin{cases} \sum_{\{i: x \in r_i^{\mathbb{P}}(M)\}} \pi(P_i, M) \cdot \frac{1}{|r_i^{\mathbb{P}}(M)|} & \text{if} \quad x \in M \\ 0 & \text{o.w.} \end{cases}$$

To see that svc is not nested by gpsc, I am going to use an axiom characterizing this model, called the *unique best*. According to this, for all menus $M \in \mathbb{X}$, there is a unique best alternative x such that $\rho_x(M) > \rho_y(M)$ for all $y \neq x$. To see that this is violated, just consider x, y, z such that $x \succ y \bowtie z$. Note that $\rho_x(xyz) = \frac{u(xy)}{u(xyz)}$ and $\rho_z(xyz) = \frac{u(z)}{u(xyz)}$. These choice probabilities are equal to each other when $u(xy) = u(z)$, so svc can violate this condition. The second axiom is called sWARP, which is formulated using the notion of *stochastic revealed preference* defined in the model. Let x be *stochastically revealed preferred* to y (denoted $x \gg_\rho y$) if there is a menu $M \ni x, y$ such that $\rho_x(M) \geq \rho_z(M)$ for all $z \neq x$ in M. sWARP says that if $x \gg_\rho y$, then $\neg(y \gg_\rho x)$. Consider the same example, but now assume that $u(xy) > u(z) > u(x)$. This implies that $\rho_x(xz) < \rho_z(xz)$ and $\rho_z(xyz) < \rho_x(xyz)$, and hence sWARP is violated. These show that svc is not nested by gpsc. Because gpsc can accommodate violations of MST as shown in Dutta, svc does not nest gpsc.

Valkanova [51] developed a model of Markov choice in which the DM compares the alternatives sequentially in discrete time. The DM starts with alternative x randomly with probability $\pi_x(M)$. Then, with some *transition probability* $\theta_{xy}(M)$ she checks the alternative y. The DM chooses y from $\{x, y\}$ with probability q_{xy} and x with probability $1 - q_{xy}$, which implies that $\tau_{xy}(M)$, the choice probability of y given x is the starting point, is equal to $\pi_x(M) \cdot q_{xy}$. The probability that x is chosen in M is equal to $1 - \sum_{y \in M \setminus \{x\}} \tau_{xy}$. Importantly, it is assumed that transition probabilities satisfy *IIA*:

$$\tau_{xy}(xy) \cdot \tau_{yx}(M) = \tau_{yx}(xy) \cdot \tau_{xy}(M)$$

This process repeats consecutively and terminates at each period with some positive probability $a > 0$. The choice procedure generates by the following stochastic choice rule:

$$\rho^{msc}(a, \pi(M), \theta(M)) = a\pi(M)(1 - (1 - a)\tau(M))^{-1}$$

which is called the *baseline msc*. In the limit as $a \to 0$, provided that the markov chain is ergodic, the *limiting msc* as the unique stationary distribution exists and it can be only characterized by the transition matrix τ. There are two types of a limiting msc which are called *reversible* and *nonreversible*, which are defined using the Kolomogorov's necessary and sufficient conditions for reversible Markov chains. Finally, a limiting msc for which IIA condition does not hold, but the DM is able to explore the whole choice set and all alternatives are comparable is called *ergodic* limiting msc. I am going to show that svc is not nested by the limiting msc model, since this is the only model for which a characterization is provided. To demonstrate this, I need to provide the main definition used for characterizing both models:

Definition 14. *$x \succ_M y$ iff $\rho_x(xy) \cdot \rho_y(M) > \rho_y(xy) \cdot \rho_x(M)$ for any $M \supseteq \{x, y\}$.*

It is shown that *reversible* limiting MSC is characterized by the *acyclicity* of $\succ_M$, while *nonreversible* limiting MSC is characterized by the cyclicity of the same binary relation. svc can accommodate both types of behavior for $\succ_M$. For the acyclicity, recall that Luce rule is a special case of svc when all alternatives are incomparable. It is also shown in msc that Luce rule is a special case of reversible MSC with positive transition probabilities. Thus, it is straightforward to see that acyclicity is satisfied for Luce rule. For an example of a cyclic $\succ_M$, let $M = \{x,y,z,t\}$ such that $x \succ y \succ t$ and $z \bowtie x,y$. I am going to show that $x \succ_M y \succ_M z \succ_M x$. First, observe that $\rho_x(xy) = 1$ and $\rho_y(xy) = 0$, which shows $x \succ_M y$. On the other hand, $\rho_y(yz) = \frac{u(y)}{u(yz)}$ and $\rho_z(yz) = \frac{u(z)}{u(yz)}$ implies that $\rho_y(yz) \cdot \rho_z(M) > \rho_z(yz) \cdot \rho_y(M)$ iff $\frac{u(y)}{u(z)} > \frac{\rho_y(M)}{\rho_z(M)}$. The latter is equal to:

$$\frac{\rho_y(M)}{\rho_z(M)} = \frac{\frac{u(t)}{u(M)} \cdot \frac{u(y)}{u(xy)}}{\frac{u(z)}{u(M)}} = \frac{u(t)u(y)}{u(z)u(xy)}$$

which implies that the required inequality holds iff $1 > \frac{u(t)}{u(xy)}$, which follows by the compatibility of u with $\succeq$. Finally, I need to show that $z \succ_M x$. To see this, note that this inequality holds iff $\frac{u(z)}{u(x)} > \frac{\rho_z(M)}{\rho_x(M)}$. The latter is equal to:

$$\frac{\rho_z(M)}{\rho_x(M)} = \frac{\frac{u(z)}{u(M)}}{\frac{u(xy)}{u(M)} + \frac{u(t)}{u(M)} \cdot \frac{u(x)}{u(xy)}} = \frac{u(z)u(xy)}{u^2(xy) + u(x)u(t)}$$

which implies that the inequality holds iff:

$$\frac{1}{u(x)} > \frac{u(xy)}{u(x)u(t) + u^2(xy)}$$

This inequality holds because $u^2(xy) > u^2(x) + u(xy)$, and therefore $\succ_M$ has a cycle. The final conclusion about ergodic msc follows because it says that all binary choice probabilities should be interior.

Finally, I will discuss the model of Ravid and Steverson [40]. In their model, the DM chooses a *focal option* with uniform probability, and then compares this option with other option in pairs sequentially, randomly and independently. Let $\pi(x,y)$ denote the probability that x passes a comparison with y, which does not need to be symmetric. For convenience, they also assume that $\pi \in [0,1)$. The choice probability of x in M is defined as:

$$\rho_x^{ftc}(M) := \frac{\prod_{y \in M \setminus \{x\}} \pi(x,y)}{\sum_{y \in M} [\prod_{z \in M \setminus \{y\}} \pi(y,z)]}$$

This choice rule is characterized by three features called *strong expansion* (SE), *independence of shared alternatives* (ISA), and finally *cancellation* without the *full support assumption* on the comparison mapping π, from which I will only use SE:

$$\forall x \in M \cap M' : \quad \rho_x(M) > 0 \;\&\; \rho_x(M') > 0 \iff \rho_x(M \cup M') > 0$$

To see svc is not nested by ftc, consider $x \succ y \succ z$. Observe that $\rho_y(xyz) > 0$ since y strictly dominates z, while $\rho_y(xy) = 0$. This violates the if part of strong expansion. These models have nonempty intersection since Luce is a special case of ftc. To see that svc does not nest ftc, it is enough to note that MM is nested by ftc, which is not true for svc.