

# *Choice by Iterative Search\**

Yusufcan Masatlioglu<sup>†</sup>  
University of Michigan

Daisuke Nakajima<sup>‡</sup>  
University of Michigan

Aug, 13, 2012

## **Abstract**

When making choices, decision makers often either lack information about alternatives or lack the cognitive capacity to analyze every alternative. To capture these situations, we formulate a framework to study behavioral search by utilizing the idea of consideration sets. Consumers engage in a dynamic search process. At each stage, they consider only those options in the current consideration set. We provide behavioral postulates characterizing this model. We illustrate how one can identify both search paths and preferences.

JEL Classification: D11, D81.

Keywords: Search, Bounded Rationality, Consideration Set, Reference-Dependent Choice, Revealed Preference

---

\*We would like to thank Tilman Börgers, Kfir Eliaz, Larry Epstein, Fred Feinberg, Michihiro Kandori, Miles Kimball, Kai-Uwe Kühn, Stephan Laueremann, Marco Mariotti, Efe Ok, Emre Özdenören, Doug Smith, Lones Smith, Elena Spatoulas, Ran Spiegler, Neslihan Üler, Gabor Virag and Frank Yates for their helpful comments. Special thanks are due to Steve Salant, Collin Raymond and Ben Meiselman who commented on a large portion of the manuscript. We also thank the National Science Foundation for financial support under Grant SES-1024544.

<sup>†</sup>Department of Economics, yusufcan@umich.edu.

<sup>‡</sup>Department of Economics, ndaisuke@umich.edu.

# 1 Introduction

Classical choice theory assumes that a decision maker (DM) chooses the best option among *all* available alternatives. This might not be practical, especially in situations where (i) a DM cannot easily view all the alternatives, and she must actively seek out alternatives, as when she buys a house or a car (incomplete information), or (ii) there is a menu in front of her but either the menu is too long or her time is too short (limited cognitive capacity).<sup>1</sup> In these situations, the budget set must be explored by the DM. As Herbert Simon pointed out many years ago, exploration of the budget set is one of the most important aspects of real world decision making but is neglected in the standard theory where the whole budget set is assumed to be evaluated simultaneously.

This paper proposes a new descriptive model of decision making in which a DM explores her budget set (unlike the standard model) and has a stable preference (as in the standard model). Our procedure is dynamic and incorporates the idea of limited search into a model of decision making. We utilize consideration sets, which have been extensively studied in marketing.<sup>2</sup> A *consideration set* is the subset of all available options to which the DM pays attention.

While the notion of consideration sets is consistent with standard search theory (Stigler [1961], McCall [1970], and Mortensen [1970]), there is ample evidence that consumers use heuristic decision rules to determine their consideration sets (Hauser [2010] and Gigerenzer and Selten [2001]).<sup>3</sup> We do not take a stand on which explanations for consideration sets are most plausible. Instead we treat the consideration sets as latent variables and infer both consideration sets and preferences from observed choice behavior.

The novelty of this paper is to model explicitly how the consideration set of a DM evolves during the course of search. The economics literature has only recently begun to study the effects of limited consideration in market environments (Chioveanu and Zhou [2009], Eliaz and Spiegler [2011], Spiegler and Eliaz [2011], Goeree [2007], Hendricks et al. [2012], Jeziorski and Segal [2010], and Piccione and Spiegler [2012]). No one yet has provided a general model of how a DM forms consideration sets over time while choosing among options. This paper provides such a model by linking the formation of consideration sets to a search process.

Our decision procedure is dynamic because interim decisions about what items to pay

---

<sup>1</sup>Although choice overload is usually attributed to the number of options presented to a DM, another source of choice overload could be the number of attributes. Therefore, even with small number of alternatives, one may not compare all available alternatives.

<sup>2</sup>See Hauser and Wernerfelt [1990], Roberts and Lattin [1991], and Nedungadi [1990].

<sup>3</sup>There is also ample experimental evidence that people do not search the way theory predicts (Hey [1982, 1987], Kogut [1990], Moorthy et al. [1997], Zwick et al. [2003], Gabaix et al. [2006], and Brown et al. [2011]). Instead of optimally searching, subjects use heuristics to decide whether to continue or not.

attention to affect the evolution of the consideration set. Consider a consumer who is searching for a digital camera. She begins with a tiny amount of knowledge relative to the vast array of products available. Suppose she utilizes an e-commerce site to explore alternatives. First, she looks up a particular camera that she has heard about. Then the website recommends several other cameras. The recommendations influence her choice of what to examine next and when to stop searching. When she stops searching, she chooses the best alternative from the entire consideration set. We call such behavior *choice by iterative search* (CIS). In this model, the consideration set evolves during the course of search because the items observed during the search (history) trigger additional search.

There are two plausible interpretations of our model. The first is that the evolution of the consideration set reflects a mental process of the DM. Although there is no explicit search process, the consumer is only able to consider a subset of all available options at any given time. For instance, she may consider only options that are similar to the best item she has already investigated. Under this interpretation, our model becomes a model of an internal process of decision making, even when all options appear to be available.

The second interpretation is that consideration sets are shaped by the environment faced by the DM. For example, a consumer buying items online is often presented with options related to what she is examining. These related options, which are provided by the seller, might form the consideration set of the consumer at any given moment. As the consumer clicks through to look at different items, the consideration set evolves. Although we will discuss our model in terms of search, the reader should keep in mind that the search may be entirely a mental process.

In Section 2 we introduce and characterize a general CIS model, where the history and the feasible set affect the consideration set in an arbitrary manner. Because of the generality of the model, inference about consideration sets and preferences is limited. However, this model is still useful both because it identifies the largest class of choice behaviors that might result from *limited search* rather than changing preferences and because it acts as a benchmark to compare with models where additional structure is imposed on consideration sets.<sup>4</sup>

In Section 3 we study a special class of CIS models where each consideration set depends only on the best alternative previously observed. This model matches important real world situations, such as Internet shopping websites that recommend alternatives to a product being considered. For example, amazon.com provides information about alternatives in

---

<sup>4</sup>In addition, this model provides a new tool for understanding the market interactions between profit-maximizing firms and boundedly rational consumers. Using two special cases of our model, Eliaz and Spiegler [2011] and Spiegler and Eliaz [2011] illustrate the usefulness of the CIS model in explaining observed economic behavior which is inconsistent with conventional models.

a section called “What Do Customers Ultimately Buy After Viewing This Item?” The restricted CIS model can be seen as a model of dynamic status quo bias, where the status quo is always the best element in the current consideration set. The current status quo restricts the set of options the DM is willing to consider. The DM will continue to change her status quo until the current status quo is better than all elements in the status quo’s consideration set.

We characterize the model and identify behavioral patterns that are consistent with this consideration set formation process. The restriction improves the predictive power of the CIS model and still allows for choice reversals and cyclical behavior. The restricted CIS model provides richer information about the process leading up to the final choice and about the DM’s preferences. The restricted CIS model can be used to pin down uniquely the path a DM follows during her search from her choice data.

This analysis assumes that the available data about the DM’s choices are slightly richer than the standard choice data: we observe not only her choice but also her starting point. The *starting point* of a search is the alternative that the DM initially pays attention to. There are many potential reasons for a particular starting point. For example, a starting point could be (i) the default option, (ii) the last purchase or status quo, (iii) a product advertised to the DM, or (iv) a recommendation from someone in the DM’s social network.<sup>5</sup> With the explosion of data mining technologies, observability of such data is more plausible. Nevertheless, it is conceivable that starting points will not be observable in some situations. In Section 4 we identify both the general CIS model and the restricted CIS model using only standard choice data.

Section 5 discusses the relation between CIS and three branches of decision theory: search, limited attention, and reference-dependent choice. Section 6 concludes.

## 2 Model

Throughout the paper  $X$  denotes an arbitrary non-empty finite set, with each element of  $X$  a potential choice alternative. Let  $K(X)$  denote the set of all nonempty subsets of  $X$ . An *extended choice problem* is a pair  $(S, x)$  where  $S \in K(X)$  is a feasible set (budget set) and  $x \in S$  is a starting point. We interpret this to mean that the individual is confronted with the problem of choosing an alternative from a feasible set  $S$  and her search starts with  $x$ .

An *extended choice function* assigns a single chosen element to each extended choice problem. That is,  $c(S, x) \in S$  for every extended choice problem. Before we introduce

---

<sup>5</sup>Masatlioglu and Ok [2005], Salant and Rubinstein [2008], and Bernheim and Rangel [2009] use similar domains in which  $x_0$  is interpreted as a status quo, as a frame, and as ancillary conditions, respectively.

our main concept, *choice by iterative search*, we formally define its two basic components: preference relations and consideration sets.

A preference relation, which is typically denoted by  $\succ$ , is a strict order over  $X$  as in the standard theory.<sup>6</sup> If  $x \succ y$  for all  $y \in S \setminus \{x\}$ , we call  $x$  the  $\succ$ -best in  $S$ .

Suppose, at the beginning of particular stage of search in extended choice problem  $(S, x)$ , a DM has already considered a set of alternatives  $A \subseteq S$ , which we call a *history*. Given a history, the *consideration set* after  $A$  under choice problem  $S$  consists of all alternatives which the DM will have considered by the end of this stage, which is denoted by  $\Omega(A, S)$ . The set  $\Omega(A, S)$  is defined for each  $A$  and  $S$  ( $A \subseteq S$ ) and satisfies  $A \subseteq \Omega(A, S) \subseteq S$ . Notice that  $\Omega(A, S)$  represents the entire set of items considered in the search process. Thus, the set of new items searched in this stage is  $\Omega(A, S) \setminus A$ . We call  $\Omega$  a *consideration set mapping*.

**Definition.** An extended choice function  $c$  is a choice by iterative search (CIS) if there exist a preference  $\succ$  and a consideration set mapping  $\Omega$  such that, for every extended choice problem  $(S, x)$ , there exists a sequence of histories  $A_0 (= \{x\}), A_1, \dots, A_n$  with  $n \geq 0$  and

- $\Omega(A_k, S) = A_{k+1}$  for  $k = 0, \dots, n - 1$ , and  $\Omega(A_n, S) = A_n$
- $c(S, x)$  is the  $\succ$ -best element in  $A_n$ .

In this case, we say  $c$  is represented by  $(\succ, \Omega)$ . We may also say that  $\succ$  represents  $c$ , which means that there exists some  $\Omega$  such that  $(\succ, \Omega)$  represents  $c$ .

The consideration set in the CIS model evolves depending on a starting point. Given  $(S, x)$ , the DM first considers all elements in  $\Omega(\{x\}, S) = A_1$ . Given history  $A_1$ , she further searches and considers all elements in  $\Omega(A_1, S) = A_2$  (in this stage she finds  $A_2 \setminus A_1$ ). The search process ends when the decision maker is convinced there is no need for additional search, so  $\Omega(A_n, S) = A_n$ . At this point, she chooses the best element in  $A_n$ . Mathematically,  $A_n$  is a fixed point of the consideration set mapping given  $S$ . Since  $\{A_n\}$  is an expanding sequence in  $K(S)$ , which is finite, a fixed point is eventually reached regardless of the starting point.

## Characterization

We first characterize the most general version of CIS where the history and the feasible set affect the consideration set in an arbitrary manner. Not surprisingly, the general CIS model does not provide a very strong prediction. Nevertheless, we provide the characterization in order to draw the boundaries of the current framework. The general CIS model identifies

---

<sup>6</sup>A binary relation  $\succ$  on  $X$  is a strict order over  $X$  if it is asymmetric ( $x \succ y$  implies not  $y \succ x$ ) and negatively transitive (not  $x \succ y$  and not  $y \succ z$  imply not  $x \succ z$ ).

choice behavior that might result from limited search. Within this class, seemingly irrational choice behavior can be generated even with stable preferences. In other words, context-dependent preferences are not necessary to explain choice anomalies. We will use the general CIS model as a benchmark for comparison with models where we make additional assumptions about consideration sets.

In this benchmark case, we do not impose any structure on consideration sets. The restrictions imposed by the model come only from revealed preference. Indeed, the necessary and sufficient condition for the model is equivalent to acyclicity of revealed preference. We next illustrate how to elicit preference in our framework.

Suppose we observe  $c(S, x) = x$ . The standard theory concludes that  $x$  is preferred to any other alternative in  $S$ . To justify such an inference, one must implicitly assume that the consideration set is equal to  $S$ . Without this assumption, we cannot make any inference because it is possible that  $x$  is the worst alternative but the final consideration set consists of only  $x$  (i.e. the decision maker does not search at all). Therefore, the choice does not reveal any information about preferences.

However, if a decision maker chooses  $x$  from some set for which  $y$  is the starting point, we must conclude that  $x$  is better than  $y$ . Formally, for any two distinct  $x$  and  $y$ , define  $x \succ_c y$  if  $c(S, y) = x$ . We now state a postulate in terms of observables, then show that it implies that  $\succ_c$  is acyclical.

**DOMINATING ANCHOR:** *For any budget set  $S$ , there is some alternative  $x^* \in S$  such that  $c(T, x^*) \notin S \setminus \{x^*\}$  for all  $T \ni x^*$ .*

By contradiction, we show the Dominating Anchor axiom implies  $\succ_c$  being acyclical. Assume  $c$  satisfies the Dominating Anchor axiom and  $\succ_c$  is cyclic. If elements of  $S$  form a cycle in terms of  $\succ_c$ , this means that for every alternative  $x \in S$ , there exists another item  $y \in S \setminus \{x\}$  such that  $y = c(T, x)$  for some  $T$ , which violates the Dominating Anchor axiom. Hence,  $\succ_c$  cannot have a cycle. We now state our characterization theorem.

**Theorem 1.** *An extended choice function  $c$  obeys the Dominating Anchor Axiom iff  $c$  is a CIS. Moreover,  $\succ$  represents  $c$  iff  $\succ$  includes the transitive closure of  $\succ_c$ .*

*Proof.* Suppose  $c$  satisfies the Dominating Anchor axiom.<sup>7</sup> As we show above,  $\succ_c$  is acyclical. Therefore, there exists a preference that includes  $\succ_c$ . Take one of these preferences arbitrarily, say  $\succ$ , and define  $\Omega(\{x\}, S) := \{x, c(S, x)\}$ , and for all other  $A$ ,  $\Omega(A, S) := A$ . Now we show that  $(\succ, \Omega)$  represents  $c$ . If  $x = c(S, x)$ , by construction,  $\Omega(\{x\}, S) = \{x\}$

---

<sup>7</sup>The proof of the “if” part of the first claim in Theorem 1 is left to the reader.

so the decision maker stops searching immediately. Suppose  $c(S, x) = y \neq x$ . Then  $\Omega(\{x\}, S) = \{x, y\} = \Omega(\{x, y\}, S)$ . By the construction of  $\succ_c$ , we have  $y \succ_c x$  so it is  $y \succ x$ . So,  $y$  is the  $\succ$ -best element of  $\Omega(\{x, y\}, S)$ . Therefore,  $(\succ, \Omega)$  represents  $c$ . Moreover, since  $\succ$  can be any preference including  $\succ_c$ , this proves the “if” part of the second claim. The “only if” part is given above.  $\square$

The second claim in the theorem illustrates the extent to which we can identify the DM’s preference through the choice data. To answer this question, we use the notion of revealed preference introduced by Masatlioglu et al. [2012] where multiple preferences may represent a single choice function.<sup>8</sup> In Masatlioglu et al. [2012]’s definition,  $x$  is revealed preferred to  $y$  if and only if every preference representing the choice behavior ranks  $x$  above  $y$ . Adopting this notion, the second claim of Theorem 1 implies that the transitive closure of  $\succ_c$  fully characterizes the revealed preference of the CIS model because every preference representing  $c$  must respect  $\succ_c$  and any preference doing so represents  $c$ . In other words, the necessary and sufficient condition for concluding  $x$  is preferred to  $y$  is that  $x$  is ranked above  $y$  according to the transitive closure of  $\succ_c$ .

As a final word, we want to highlight one important limitation of the general CIS model. As the proof of Theorem 1 illustrates, it is possible to explain any choice function generated by the general CIS model as if the DM always ends her search after the first iteration. In other words, the general CIS model is behaviorally equivalent to the static model where the DM simply chooses the best element of  $\Omega(\{x\}, S)$ . The general CIS model is therefore not informative about search behavior.<sup>9</sup> In the next section, we impose a realistic and reasonable restriction on consideration sets. We illustrate that this restriction not only delivers more information about true preference but also provides a full identification result about the search path.

### 3 Markovian CIS

Consider a consumer using an internet commerce site to buy a cell phone. On the page of each phone, the internet commerce site suggests several alternatives under a heading such as “What Do Customers Ultimately Buy After Viewing This Item?” (amazon.com). The

---

<sup>8</sup>Revealed preference with multiple representations is also studied by Ambrus and Rozen [2008], Caplin and Dean [2011], Cherepanov et al. [2008], Green and Hojman [2008], Manzini and Mariotti [2009], and Manzini and Mariotti [2012].

<sup>9</sup>Even though there is no difference between the dynamic and the static model for our characterization, the difference could be substantial when one utilizes this model in applications. For example, one might be concerned about implications of a policy that affects the DM’s consideration set mapping. A policy that expands her consideration set mapping (i.e.  $\Omega(A, S)$  gets bigger for any  $(A, S)$ ) will always make the DM weakly better off under the static model but not necessarily under the dynamic model.

consumer first visits the page of her starting point and considers the alternatives suggested by the site. She then jumps to the page of the phone she prefers among those alternatives, where she finds a new set of suggested phones. She looks for the best phone within the new suggestions and repeats the process. She stops searching when she finds a phone for which there is no better alternative among those suggested on its page.

Under this strategy, the range of the next search depends only on the best alternative already encountered. We call such a procedure *Markovian* CIS. Formally, the decision maker's consideration set mapping has the following structure:

$$\Omega(A, S) = (\Omega^*(x_A) \cap S) \cup A \tag{1}$$

where  $x_A$  is the best item in  $A$  and  $\Omega^*$  is a mapping from  $X$  to  $K(X)$  and  $x \in \Omega^*(x)$ .  $\Omega^*(x)$  represents the set of suggested alternatives at the page of  $x$  (the consideration set of  $x$ ). Sometimes, a suggested item may not be available. Thus, what the consumer can actually compare are those included in the intersection:  $\Omega^*(x_A) \cap S$ . We call mapping  $\Omega^*$  the *Markovian consideration set*. We also say that  $(\succ, \Omega^*)$  represents a Markovian CIS  $c$  if  $c$  is represented by  $(\succ, \Omega)$  where  $\Omega$  is defined by (1).

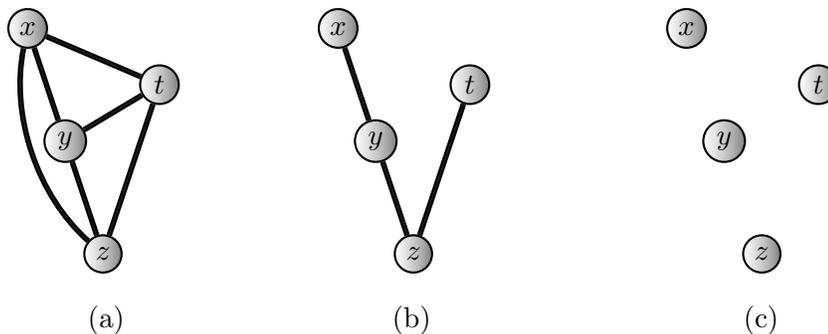


Figure 1: Graphical Illustrations: (a) Full Search, (b) Partial Search and (c) No Search.

Under the assumption that consideration sets are Markovian, we can provide a graphical illustration for the CIS model. In Figure 1, each node corresponds to a particular alternative, and the positions of nodes represent her preferences: the higher the position is, the better the alternative is,  $(x \succ t \succ y \succ z)$ . A connection between two alternatives indicates that each of them belongs to the consideration set of the other.<sup>10</sup> We illustrate three cases from one extreme to the other with four alternatives in Figure 1(a)-(c): while (a) represents a standard decision maker (independent of starting point, the decision maker considers all

<sup>10</sup>A directed graph can be used to illustrate asymmetric membership of consideration sets: only one of them is the consideration set of the other.

feasible alternatives, hence always chooses the best alternative), in (c), she does not consider anything other than the starting point, hence she sticks to the starting point no matter what the budget set is. The interesting case is Figure 1(b): Her best option  $x$  is visible from  $y$  but not from  $z$ . Therefore, she cannot reach  $x$  when  $z$  is the starting point unless  $y$  is available. In other words, to choose  $x$ , she needs to follow the path,  $z \rightarrow y \rightarrow x$ .

Under the Markovian structure, the search stops whenever the best item so far is also best within its own consideration set. To see this, suppose the decision maker has history  $A$  in which  $x$  is the best element and it is also best in  $\Omega^*(x) \cap S$ . Then her next consideration set will be  $A' = (\Omega^*(x) \cap S) \cup A$ , so  $x$  is still the best item. Therefore, her next consideration set is also  $A'$  so she stops searching and chooses  $x$ .

Markovian CIS is a plausible model of search not only for those searching the internet, but also for many other scenarios:

- The set  $\Omega^*(x)$  includes alternatives that have similar attributes or overlapping features with  $x$ , for example, “comps” of a house. It is natural to assume that decision makers, when considering whether to purchase a house, only compare it to houses in a similar price range, rather than all available houses.
- Depending on the best item a decision maker has observed so far, some cues (or aspects) of items might look more important than others. The decision maker only considers items that dominate the current one in *these* cues. Hence  $\Omega^*(x)$  excludes alternatives which incurs loss relative to  $x$  in any of these cues.
- Depending on trade-offs, some binary comparisons could be difficult. To avoid these difficult trade-offs, the decision maker only considers items which are easy to compare to  $x$ .

## Characterization

Now, we draw the boundaries for the empirical scope of the Markovian CIS model. To do this we present a set of postulates that are both necessary and sufficient for Markovian CIS.

The first one is a stronger version of the Dominating Anchor axiom, which ensures the acyclicity of the revealed preference. Given the additional structure of the consideration set mapping, the Markovian CIS model generates more revealed preference. For example, if  $x = c(\{x, y\}, y)$  but  $x \neq c(S, y)$  implies that  $x$  must be inferior to  $c(S, y)$  according to the Markovian CIS model but it may not be in the general CIS model. Thus, we need the following axiom:

**A 1.** (STRONG DOMINATING ANCHOR) *For any  $S$ , there is some alternative  $x^* \in S$  such that for any  $y \in X$  if  $x^* = c(\{x^*, y\}, y)$  then  $c(T, y) \notin S \setminus \{x^*\}$  for all  $T$  including  $x^*$ .*

Notice that the Dominating Anchor axiom is the special case of A 1 when  $y = x^*$ . If the axiom is violated, then for every alternative in  $S$  there exists another alternative in  $S$  which is superior. Therefore, the violation of this axiom implies that  $S$  does not have a most preferred alternative, which is a contradiction. We shall discuss the revealed preference of the Markovian CIS model in detail in the next subsection.

The next postulate states conditions under which an alternative is irrelevant for a fixed starting point. An alternative is irrelevant if adding it to a choice set does not affect the choice. First, fix the starting point, say  $x$ . If  $y$  is irrelevant in  $\{x, z\}$ , i.e.,  $c(\{x, z\}, x) = c(\{x, y, z\}, x)$ , then in the classical choice theory,  $y$  is also irrelevant in any decision problem whenever  $z$  is present. In other words, the presence of  $z$  makes  $y$  globally irrelevant. Our next postulate is a weaker version of this idea. For an alternative to be irrelevant globally in the presence of  $z$ , it must be irrelevant in  $\{x, z\}$  and it must also dominate the starting point, ( $y = c(\{x, y\}, x)$ ).

**A 2.** (IRRELEVANCE) *Assume  $y = c(\{x, y\}, x)$ . Then  $c(\{x, y, z\}, x) = c(\{x, z\}, x)$  implies  $c(S, x) = c(S \cup y, x)$  for any  $S$  including  $z$ .*

We now illustrate that this is a necessary condition of Markovian CIS. The assumptions of the postulate imply that both  $y$  and  $z$  belong to the consideration set of  $x$ . Since  $y$  is not chosen from  $(\{x, y, z\}, x)$ ,  $y$  will never be the best element in the first consideration set if  $z$  is available. This means the alternatives encountered in the next stage will be the same whether  $y$  is present or not. Hence this postulate is also necessary.

Next, consider a DM who chooses the starting point  $x$  in two different decision problems. The next postulate requires that she also chooses  $x$  when the two budget sets are combined and  $x$  is the starting point.

**A 3.** (EXPANSION) *If  $x = c(S, x) = c(T, x)$ , then  $x = c(S \cup T, x)$ .*

It is easy to see that this is a necessary condition of Markovian CIS. If  $x = c(S, x) = c(T, x)$  then there exists no element in  $S$  or  $T$  such that it is in  $\Omega^*(x)$  and better than  $x$ . In this case, the DM sticks to her starting point under the decision problem  $(S \cup T, x)$ .

Our final postulate states that if there is only one alternative, say  $y$ , that makes the DM move away from the starting point, say  $x$ , then it does not matter whether  $x$  or  $y$  is the starting point. In other words, one can replace the starting point with this dominating alternative.

**A 4.** (REPLACEMENT) *If  $c(S, x) = x$  and  $c(\{x, y\}, x) = y$  then  $c(S \cup y, x) = c(S \cup y, y)$ .*

Since  $c(\{x, y\}, x) = y$ , we must have  $c(S \cup y, x) \neq x$ . This means the decision maker must encounter a better alternative than  $x$  in the first stage of search. Since  $c(S, x) = x$ , there is only one such alternative, which is  $y$ . Then she continues to search as if she started from  $y$  because of the Markovian structure. Thus, in both decision problems  $(S \cup y, x)$  and  $(S \cup y, y)$ , the DM will follow the same path and reach the same final choice, i.e.  $c(S \cup y, x) = c(S \cup y, y)$ .<sup>11</sup>

**Theorem 2.** *An extended choice function  $c$  satisfies A1-A4 iff  $c$  is a Markovian CIS.*

The proof is provided in the Appendix.

## Revealed Preference, Revealed Consideration and Search Path

Given that Markovian CIS imposes more structure on the consideration set mapping, we can infer more about the preference of the DM. We now list three choice patterns which reveal the DM prefers  $x$  over  $y$ .

1. Abandonment of the starting point:  $x = c(S, y)$ ,
2. Choice reversal:  $x = c(S \cup y, z) \neq c(S, z)$ ,
3. Chosen over an dominating alternative:  $x = c(S \cup y, z)$  and  $y = c(\{y, z\}, z)$ .

The first choice pattern is straightforward. Indeed, the inference that  $x$  is preferred to  $y$  is possible when this choice pattern is observed even when no structure is imposed on the consideration set mapping in the general CIS model.

The second choice pattern is a choice reversal where  $x$  is unchosen when a seemingly irrelevant alternative  $y$  is removed. Then in the decision problem  $(S \cup y, z)$ ,  $y$  must be the best element at some step of the search process because that is the only way it can affect the DM's choice. We can conclude that  $x$  is preferred to  $y$  because we know that  $y$  was considered and  $x$  was chosen. Notice that this inference is not possible with the general CIS model, in which an element that is not the best at any step may affect the consideration set and thereby influence the final choice.

In the third choice pattern,  $x$  is chosen over  $y$ , which dominates the starting point  $z$ .  $y = c(\{y, z\}, z)$  reveals that  $y$  is in  $z$ 's consideration set. Therefore, anything chosen over  $y$  when  $z$  is the starting point must be better than  $y$ .

---

<sup>11</sup>In the Appendix, we provide examples illustrating that these four axioms are independent.

Any preference that can represent  $c$  must be consistent with the above revelations. Formally, given a Markovian CIS  $c$ , let

$$x \gg_c y \text{ if one of above three choice patterns is observed}$$

and  $\gg_c^{TC}$  be the transitive closure of  $\gg_c$ . If  $x \gg_c^{TC} y$ , then we can conclude the DM prefers  $x$  over  $y$ . The converse is also true: any preference that respects  $\gg_c^{TC}$  represents  $c$  as well. Thus,  $\gg_c^{TC}$  fully characterizes the revealed preference of the Markovian CIS model. The next proposition states these facts formally.

**Proposition 1.** (*Revealed Preference*) *Suppose  $c$  is a Markovian CIS. Then  $x$  is revealed to be preferred to  $y$  iff  $x \gg_c^{TC} y$ .*<sup>12</sup>

We would like to note that the transitive closure of the binary relation generated by the third choice pattern (chosen over an dominating alternative) itself fully captures the revealed preference of the Markovian CIS model. Thus, the first two patterns (abandoning the starting point and choice reversal) may look redundant. For instance, when a choice reversal is observed, one can prove that  $x$  is preferred to  $y$  in the sense of the third condition by seeking more choice data and/or applying the axioms to fill in the missing choice data. We provide these two conditions explicitly because it eliminates such a cumbersome need and makes Proposition 1 more useful for empirical studies.

Now we illustrate when we can conclude that  $x$  is in the consideration set of  $y$ , which we call the revealed consideration set. As in our definition of the revealed preference with multiple representations, we say  $x$  is revealed (not) to be in the consideration set of  $y$  if every possible Markovian representation of an extended choice function agrees that  $x$  belongs (does not belong) to the consideration set of  $y$ .

In order to conclude that  $x$  is in  $y$ 's consideration set, we need to observe  $x = c(\{x, y\}, y)$ . Notice that it is not enough to observe  $x = c(S, y)$  because  $x$  may be considered only after the DM switches from  $y$  to something else. On the other hand, to infer  $x$  is not in  $y$ 's consideration set, we need to observe  $y = c(S, y)$  for some  $S$  and to know  $x$  is better than  $y$ . If  $x$  were in  $y$ 's consideration set, the DM would have abandoned  $y$ . The following remarks summarize these observations.

**Remark 1.** (*Revealed Consideration*) *Suppose  $c$  is a Markovian CIS. Then  $x$  is revealed to be in the consideration set of  $y$  iff  $y = c(\{x, y\}, x)$ .*

**Remark 2.** (*Revealed Inconsideration*) *Suppose  $c$  is a Markovian CIS. Then,  $x$  is revealed not to be in the consideration set of  $y$  iff  $y = c(S, y)$  for some  $S$  including  $x$  and  $x \gg_c^{TC} y$ .*

---

<sup>12</sup>All results in this section can be easily verified by the proof of Theorem 2.

Finally, we illustrate what we can infer about the decision making process. For instance, suppose  $c(\{x, y, z\}, x) = y$ . We may wonder if the DM finds  $y$  immediately (in the consideration set of  $x$ ) or finds  $y$  only after inspecting  $z$ . If we also observe  $c(\{x, y\}, x) = x$ , we can conclude that the DM finds  $y$  only after inspecting  $z$  because this information reveals that  $y$  is not included in  $x$ 's consideration set. In this way we can identify how the DM reaches  $y$ , which we call her search path, as  $x \rightarrow z \rightarrow y$ .

This idea is generalizable. Suppose we would like to know how the DM reaches the final choice in  $(S, x_0)$ . We can do so by citing her choice from the set of alternatives that are known to be in  $x_0$ 's consideration set. If we let  $S_0 = \{y \in S : y = c(\{x_0, y\}, x_0)\}$ , her next point must be  $c(S_0, x_0)$ . Repeating the process, we can fully identify how she reaches her final choice ( $c(S, x_0)$ ) from her starting point  $x_0$ . Thus, her observed choices fully identify her search path. The following remark formalizes this.

**Remark 3.** *Suppose  $c$  is a Markovian CIS. Every Markovian CIS representation of  $c$  generates the same path for every extended decision problem.*

The novelty of Remark 3 is that it provides choice-based insights about the process of search. While the standard search theory informs us of how a DM *should* behave, Remark 3 tells us how the DM actually *does* behave during the search. If the data on the search path is available, our identification result of the search path (Remark 3) can be used to test the model in addition to our axiom.

## Some Behavioral Implications of the Markovian CIS Model

The Markovian CIS model retains several consistency properties even though it accommodates choice reversal. We highlight the following properties, which are implied by the set of postulates that characterize the Markovian CIS model (Theorem 2).

- Starting Point Bias: If  $y = c(S, x)$ , then  $y = c(S, y)$ .
- Starting Point Contraction: If  $c(S, x) = x$ , then  $c(T, x) = x$  for all  $T \subset S$ .
- Sandwich Property: If  $c(S, x) = c(\{x, y\}, x) = y$  and  $\{x, y\} \subset T \subset S$ , then  $c(T, x) = y$ .
- Irrelevance of Revealed Inferior Alternatives: If  $x \gg_c^{TC} y$ , then  $c(S, x) = c(S \cup y, x)$  for all  $S$ .

The starting point bias property says that if a decision maker ever chooses  $y$  from  $(S, x)$ , then her search outcome will be the same when  $y$  is the starting point. This property appears

as the status quo bias axiom in Masatlioglu and Ok [2005], which interprets the starting point as a status quo. In the Markovian CIS model, the final choice must be best within its own consideration set, and if the DM starts from such an alternative, she will be unwilling to move away.

The starting point contraction property is a version of the independence of irrelevant alternatives (IIA) axiom. This version of the axiom is applicable in the Markovian CIS model when the DM sticks with her starting point, which she does only if she cannot find anything better in its consideration set. If she is happy with her starting point facing a large menu, she will still be happy with her starting point when the menu shrinks.

In general, IIA does not apply if the DM does not stick with her starting point. This is because removing an alternative that is on the path of search can change the DM's path and cause a choice reversal (Example 1(b)). Thus, the IIA property is preserved only when the removed alternative is not on her path of search. For instance, if  $c(S, x) = c(\{x, y\}, x) = y$ , then we learn that  $y$  is reached in the initial iteration of search, so all other alternatives in  $S$  are off the search path. In this case, we can safely state that removing an alternative (other than  $x$  and  $y$ ) does not affect the DM's choice, which is the sandwich property.

The final property, the irrelevance of revealed inferior alternatives property, is based on the observation that an alternative that is worse than the starting point never affects the DM's search behavior. Taking the unobservability of her preferences into account, this property states that removing an alternative which is revealed to be worse than the starting point does not change her choice.

### **Additional Knowledge about Preference and Consideration Sets**

Although we have characterized the Markovian CIS model assuming that both  $\succ$  and  $\Omega^*$  are unobservable, one can imagine circumstances where these factors are partially or fully observable. In our example of a DM using a website to explore digital cameras, we may learn about consideration sets and thereby observe  $\Omega^*$  by inspecting how the website makes suggestions. In a market with identical products, we may make the reasonable assumption that consumers only care about prices and thereby observe  $\succ$ .

In order to represent a certain extended choice function by the Markovian CIS model, we must use certain pairs of a preference and a consideration set mapping although they may not be unique. Thus an extended choice function which admits a Markovian CIS representation may fail to have a representation consistent with extra information about the DM's preference and consideration sets. We are now addressing this issue.<sup>13</sup>

---

<sup>13</sup>Cherepanov et al. [2008] and Manzini et al. [2011] also study a similar question in different contexts.

First, consider the case where we fully observe both the DM's preference  $\succ$  and the DM's behavior  $c$ . If she follows the Markovian CIS model, the true preference will not contradict the revealed preference generated by  $c$  ( $\gg_c^{TC}$  by Proposition 1). Conversely, if there is no contradiction, we can find a consideration set mapping  $\Omega^*$  so that  $c$  is represented by  $(\succ, \Omega^*)$ .<sup>14</sup> Thus, as long as there is no contradiction between her (known) preference and the revealed preference, her behavior can be represented by the Markovian CIS model.

Indeed, this logic is applicable even when we only have partial information about her preference. For instance, we know she prefers  $x$  over  $y$  but do not know how she ranks other elements. In the abstract, we know that her preference must include a partial order (asymmetric and transitive)  $\succ'$ . The special case of full information about the DM's preference is when  $\succ'$  is complete. The following remark summarizes these findings and it can be easily verified by the proof of Theorem 2.

**Corollary 3.** *Let  $c$  be an extended choice function and  $\succ'$  be a partial order over  $X$ . Then there exists a pair of preference  $\succ$  and Markovian consideration set mapping  $\Omega^*$  that represents  $c$ , and  $x \succ y$  whenever  $x \succ' y$ , iff (i)  $c$  satisfies A1-4, and (ii)  $\succ'$  does not contradict  $\gg_c^{TC}$  (there is no  $x$  and  $y$  such that  $x \succ' y$  but  $y \gg_c^{TC} x$ ).*

Next, we study the case where the consideration set is fully observable to be  $\Omega^*$ . Unlike the known preference case, the consistency between revealed (in)consideration and  $\Omega^*$  is not sufficient for a valid representation. The first extra condition we must impose is that the DM never moves away from her starting point if we know that she does not consider anything other than her starting point.

The second condition is about the consistency of the revealed preference generated by this extra information. Suppose we observe  $x = c(S, z)$  and it is known that  $y \in \Omega^*(z)$ .<sup>15</sup> Then clearly the DM's final choice is made after considering  $y$ , so we learn that  $x$  is preferred to  $y$ . Formally, for any distinct alternatives  $x$  and  $y$

$$x \gg'_c y \text{ if and only if } c(S, z) = x \text{ for some } z \text{ and } S \ni y \text{ with } y \in \Omega^*(x).$$

where the possibility of  $z = y$  is not excluded. Notice that the revealed preference  $\gg'_c$  due to the extra information about  $\Omega^*(x)$  contains more information than the revealed preference  $\succ_c$  generated without this information. It happens  $z = c(\{y, z\}, z)$  and  $y \in \Omega^*(z)$  is known. Thus, the slightly stronger revealed preference must not have a contradiction.

It turns out that these two extra conditions, along with all of the axioms in Theorem

---

<sup>14</sup>We showed this in the proof of Theorem 2. We can take any completion of  $(\succ_c)$  as the preference to represent a given choice function.

<sup>15</sup>As we have seen, if  $y = c(\{y, z\}, z)$ , then  $x \succ y$  is revealed even without the extra information about  $\Omega^*$  because that choice itself reveals  $y \in \Omega^*(z)$ .

2, are sufficient for an extended choice function to be compatible with the Markovian CIS model. The proof is given in the appendix.

**Proposition 2.** *Let  $c$  be an extended choice function and  $\Omega^*$  be a Markovian consideration set mapping. Then there exists a preference  $\succ$  such that  $(\succ, \Omega^*)$  represents  $c$  iff*

- (i)  $c$  satisfies A1-4,
- (ii) if  $\Omega^*(x) \cap S = \{x\}$ , then  $c(S, x) = x$ ,
- (iii)  $\succ'_c$  has no cycle.

## 4 Unobservable Starting Points

Our revealed preference approach is based on the assumption that we can observe not only what the decision maker chooses from a budget set but also which alternative she initially contemplates. However, we can also imagine situations where we observe only standard choice data and do not observe the starting point. For example, if the starting point is what the DM expects to buy in the market, it can be difficult to elicit such information. Given possible limitations in the data, we investigate how to identify DMs who follow an underlying CIS model (general or Markovian).

Suppose the DM actually follows a CIS model denoted by  $c$ , but we do not observe her starting point. We observe her choosing  $y$  from  $S$  — the only available data. Then there must exist a starting point which leads her to pick  $y$ . Of course, such a starting point may not be unique; several distinct starting points might result in the same final choice,  $y = c(S, x) = c(S, z)$ . At the same time, given a fixed budget set, observed final choices might depend on the starting points,  $y = c(S, x)$  and  $w = c(S, z)$ . We define *the choice correspondence induced* by  $c$ , denoted by  $C_c : K(X) \rightarrow K(X)$ , as follows:

$$y \in C_c(S) \text{ if there exists } x \text{ such that } y = c(S, x).$$

In other words,  $y \in C_c(S)$  means  $y$  is chosen from  $S$  for some starting point. This is in line with Sen [1993]: “It may be useful to interpret  $C(S)$  as the set of ‘choosable’ elements — the alternatives that may be chosen.”<sup>16</sup> Note that  $x, y \in C_c(S)$  does not imply that  $x$  is indifferent to  $y$  in our framework. It simply means that both  $x$  and  $y$  may be chosen from  $S$  depending on the starting point.

---

<sup>16</sup>Salant and Rubinstein [2008] also follow this idea in defining a choice correspondence from a frame-dependent choice function.

The induced choice correspondence,  $C_c$ , is what one can observe when the DM follows a particular CIS,  $c$ , but her starting point is unobservable. In other words,  $C_c$  are the only available data to an outside observer who knows that the choices of the DM are affected by the starting point but lacks information about the starting point.

## General Model

We first start with the general model. Imagine that we observe a choice correspondence  $C$ . We would like to determine if there exists a general CIS model, say  $c$ , that induces the observed choice correspondence, i.e.  $C = C_c$ . Such a general CIS model exists if and only if the DM's choice correspondence satisfies a simple axiom called the Bliss Point Axiom. This result makes it possible to identify DMs following a CIS even with standard choice data.

THE BLISS POINT AXIOM (BP): *For any set  $S$ , some may-be-chosen alternative from  $S$  must always be may-be-chosen from any smaller decision problem: There exists  $x \in C(S)$  such that  $x \in C(T)$  if  $x \in T \subset S$ .<sup>17</sup>*

We call the alternative that satisfies the condition of BP a bliss point of  $S$ . By contrast, IIA dictates that *all* may-be-chosen alternatives should be may be chosen from any smaller choice problem (whenever they are available). Therefore, BP relaxes IIA by requiring that *one* may-be-chosen element fulfills the condition rather than all may-be-chosen elements. Even though BP is a weaker condition than IIA, it still does not allow choice cycles.

We next show that BP guarantees the existence of underlying choice by iterative search.

**Proposition 3.** *A choice correspondence  $C$  satisfies the Bliss Point axiom iff there exists an underlying CIS that induces  $C$ .*

The power of Proposition 3 is that it connects choice patterns that are considered “irrational” in traditional choice theory to our choice by iterative search model, which captures bounded rationality. It provides a very simple and intuitive postulate to identify CIS decision makers using only choice data.

## Markovian

We would like to reconsider the case where our data are limited to the choices from a choice set — there is no knowledge about starting point. As before, we would like to

---

<sup>17</sup>This axiom appears in Agaev and Aleskerov [1993] as “the fixed-point condition” to characterize interval choice models.

identify DMs who follow an underlying Markovian CIS from such restricted data. If a choice correspondence  $C$  is generated by an acyclical binary relation,<sup>18</sup> then there exists an underlying Markovian CIS that induces  $C$ .

**Proposition 4.** *A choice correspondence  $C$  can be rationalized by an acyclical binary relation iff there exists an underlying Markovian CIS that induces  $C$ .*<sup>19</sup>

This provides a very simple test of Markovian CIS using only limited data. In addition, Proposition 4 makes it possible to identify behavior that is consistent with the general model but not with the Markovian model, even absent data on starting points. Therefore, standard choice data alone are enough to make the distinction between the general and Markovian CIS models.

Finally, we discuss how and to what extent one can infer the preference of the DM following the Markovian CIS model from this limited data. First, let us make it clear what a revealed preference means in this context. Since multiple Markovian CIS models can generate the same choice correspondence  $C$ , naturally we define  $x$  to be revealed preferred to  $y$  if and only if the preference of every Markovian CIS model generating  $C$  ranks  $x$  above  $y$ . The proof of Proposition 4 shows that any completion of the acyclical binary relation rationalizing  $C$  represents one of Markovian CIS models generating  $C$  (by choosing  $\Omega^*$  appropriately). Therefore, the transitive closure of this relation is the full characterization of the revealed preference in this context.

## 5 Related Literature

CIS is related to three branches of decision theory: search, limited consideration, and reference dependence. We discuss the relationships of both the general CIS model and the Markovian CIS model to these branches.

### Search

Caplin and Dean [2011] also study search by employing the revealed preference approach. Their model also assumes decision makers have a stable preference. They utilize a different type of auxiliary data called “choice process data.” These data include what the decision

---

<sup>18</sup>We say that a choice correspondence  $C$  is generated by an acyclical binary relation  $P$  if  $C(S) = \{x \in S \mid \text{there exists no } y \text{ such that } yPx\}$ .

<sup>19</sup>When the starting point is not observable, the Markovian CIS model is observationally equivalent to Köszegi and Rabin [2006]’s personal equilibrium, which is also generated by an acyclic preference as shown by Gul and Pesendorfer [2006].

maker would choose at any given point in time if she were suddenly forced to quit searching. The entire path followed during the course of search is the input of their model, rather than its output. The novelty of our approach is that the path of interim choices (the choice process data) can be uniquely identified by imposing certain structures — such as that considered in our Markovian CIS model — on consideration sets.<sup>20</sup>

### Limited Consideration

In the recent literature on limited consideration, a DM chooses the best alternative from a small subset of the available alternatives. In contrast with the CIS model, most of them focus on some other aspects of decision-making than searching. For instance, in many models of limited consideration, the DM, prior to choosing the most preferred element, applies some criteria other than her preference to the *entire* feasible set to eliminate some alternatives. Such models include the rational shortlisting (Manzini and Mariotti [2007]), considering only alternatives that belong to the best category (Manzini and Mariotti [2012]) and considering only alternatives that are optimal according to some rationalizing criteria (Cherepanov et al. [2008]). These procedures cannot be implemented without knowing the entire feasible set, so these models are inconsistent with an environment where the DM must find her alternatives through search.

Salant and Rubinstein [2008] study a decision making process in which the DM considers only the top  $n$  elements according to some ranking and chooses her most preferred element from that restricted consideration set. The model is consistent with a search environment if the ranking is, for instance, the amount of advertising for an element, because limited consideration arises even when the DM is unaware of the entire feasible set. However, this procedure assumes that she is following a fixed stopping rule and she does not adjust her search strategy during the search process.

Masatlioglu et al. [2012] provide a model of limited attention where the consideration set depends only on the budget set, denoted by  $\Gamma(S)$ . Unlike the CIS model, their model is agnostic about how such a consideration set is formed. Instead, they impose a condition satisfied in many environments and procedures. The condition says that the consideration set is unaffected by removing an alternative which does not attract attention:  $x \notin \Gamma(S)$  implies  $\Gamma(S) = \Gamma(S \setminus x)$ . The final consideration set generated by the Markovian CIS model

---

<sup>20</sup>Papi [2010], Horan [2010], and Raymond [2010] also analyze choice as a process of search. All of these papers assume that the order of search is the part of the decision problem and observable to the analyst. In contrast, in our approach the order of search (except the starting point) is in the mind of the decision maker and unobservable to the analyst.

(fixing the starting point) satisfies this property.<sup>21</sup> Thus, one can view Masatlioglu et al. [2012] as a reduced-form model where one of the possible underlying structures is CIS. Nevertheless, the strong structure of the Markovian CIS model makes it possible to identify more preferences than the limited attention model in the sense of set inclusion. In Section 3, we illustrate three types of choice patterns that reveals the DM’s preference, whereas only choice reversal does so in the limited attention model.

## Reference Dependence

The CIS model is also related to the literature on reference-dependent preferences. In particular, the Markovian CIS model captures dynamic adjustments of the reference point in search, where the reference point at any given stage is the best element in the current consideration set. Each reference point restricts the set of options the DM is willing to currently consider. The DM will continue to adjust her reference point until the current reference point is the best within its consideration set. Previous models of reference dependent choices such as Tversky and Kahneman [1991], Masatlioglu and Ok [2005], and Salant and Rubinstein [2008] are *static*—the reference point is exogenously given and fixed.

Tversky and Kahneman [1991] introduce the first reference-dependent model where losses have a greater impact than corresponding gains on utility (loss aversion). Their loss aversion model allows strict cycles: a DM strictly prefers  $y$  to  $x$  when endowed with  $x$ , strictly prefers  $z$  to  $y$  when endowed with  $y$ , and strictly prefers  $x$  to  $z$  when endowed with  $z$ , i.e.,  $x \succ_z z \succ_y y \succ_x x$ . In the CIS model, the reference point is abandoned only when there is a welfare improvement, so their model is not a special case of the CIS model. Unlike the DM in this reference-dependent model, the DM in the CIS model can violate IIA even for a fixed starting point. Therefore, these two models are independent.<sup>22</sup>

Masatlioglu and Ok [2005] propose a reference-dependent choice model where a DM follows a simple two-stage procedure: elimination and optimization. In the elimination stage, the DM discards all alternatives other than those dominating the status quo (say  $r$ ) in every (endogenously driven) aspect. Formally, the set of surviving alternatives in  $S$  is  $\{x \in S \mid u_i(x) \geq u_i(r) \text{ for all } i\} \cap S$ , where  $u_i$  represents the ranking in aspect  $i$ . In the optimization stage, the DM simply chooses the best alternative from the set of surviving alternatives. For comparison purposes, we interpret the set of the alternatives that can survive over the status quo  $r$  to be the consideration set of  $r$ , denoted by  $\Omega^*(r)$ . Thus, the DM’s choice from  $S$  is the best alternative in  $\Omega^*(r) \cap S$ . The first difference between models

---

<sup>21</sup>We should note that the final consideration set generated by the general CIS model might not satisfy this property, hence their model is more restrictive than our general CIS model.

<sup>22</sup>Since the model of Köszegi and Rabin [2006] is closely related to Tversky and Kahneman [1991], these differences also apply to their formulation of reference dependence.

is that in the Masatlioglu and Ok [2005]’s model, decisions are made just after the first round (static), whereas the final consideration set dynamically involves in the Markovian CIS model. In addition, because of the multi-utility representation, Masatlioglu and Ok [2005]’s consideration set enjoys a nested structure:  $x \in \Omega^*(y)$  implies  $\Omega^*(x) \subset \Omega^*(y)$ , which is not imposed in the Markovian CIS model.

However, the difference between the two models disappears when we impose the nested structure also on the consideration set of the Markovian CIS model. This is because once the DM finds the best element in the consideration set of the starting point, she will not find any more new alternative in the next consideration set given the nested structure so she stops searching immediately, which is behaviorally equivalent to Masatlioglu and Ok [2005]’s model. Thus, the difference is solely generated by Masatlioglu and Ok [2005]’s particular structure of the consideration set driven by the multi-utility representation.

Masatlioglu and Ok [2010] provide a more general version of the status quo bias model by eliminating the nested structure on  $\Omega^*$ . Given the behavioral equivalence between Masatlioglu and Ok [2005] and the Markovian CIS model with the nested structure, one may suspect that Masatlioglu and Ok [2010]’s model is equivalent to the Markovian CIS model without any particular structure on  $\Omega^*$ . However, the static nature of the status quo model now makes these two models distinguishable. Let  $x$  be the best alternative within  $\Omega^*(r)$ . Unlike Masatlioglu and Ok [2005]’s model, it is possible that the consideration set of  $x$  contains something better than  $x$ . Nevertheless, Masatlioglu and Ok [2010]’s DM simply picks  $x$  while the Markovian CIS’s DM will continue searching. Indeed, the model in Masatlioglu and Ok [2010] does not allow choice reversals for a fixed status quo but the Markovian CIS does. Conversely, the Markovian CIS model satisfies the following condition: if  $c(S, x) = y$  then  $c(S, y) = y$ , which is not implied by their model. In sum, the status quo bias models and the Markovian CIS model are behaviorally equivalent only when the consideration set or the set of surviving alternatives has the nested structure but that is not the case in general.

Salant and Rubinstein [2008] study choices in the presence of a framing effect using the  $(A, f)$  model, which is the most general reference-dependent choice model. A frame  $f$  is irrelevant in the rational assessment of alternatives but nonetheless affects choice. The CIS model can be thought of as a special case of the  $(A, f)$  model by letting the frame be the starting point.

## 6 Concluding Remarks

We recognize that people must engage in a dynamic process to learn about feasible alternatives, especially in complicated decision problems such as searching for a house. Thus, we incorporate the idea of search and consideration sets into decision theory to get new insights into how people make decisions in such circumstances.

We illustrate how to infer a DM's preference and search process from the choice data in the Markovian model. The choice data partially distinguish the DM's preference and consideration sets. This is useful for firms. For instance, if a product is unpopular, it is important for a firm to know why. Is it because consumers do not like the product or because they cannot find the product in the course of search? These identification results can help firms identify the reason and react accordingly.

We mention two interesting future directions to explore. The first one is to study the implications of the CIS model using extended market share data (such as 40% of consumers starting from  $x$  end up choosing  $y$ ) rather than extended individual choice data. With such limited data, we cannot observe how each individual reacts when the feasible set or the starting point changes; we can only observe reactions at the aggregated level. Then a reasonable question is how to identify the *distributions* of preferences and consideration sets if each individual follows a Markovian CIS model.

The second one is to study the effect of framing on consideration sets. While most of our work makes the useful simplification that consideration sets are only affected by the history and the feasible set, one could consider cases where consideration sets are affected by more general framing effects (such as presentation and advertising). In these cases, the consideration set can be written as  $\Omega(A, S, f)$  or  $\Omega^*(x, f)$ , where  $f$  is a frame.

## References

- R. Agaev and F. Aleskerov. Interval choice: classic and general cases. *Mathematical Social Sciences*, 26(3):249–272, 1993.
- A. Ambrus and K. Rozen. Revealed conflicting preferences: Rationalizing choice with multi-self models. Cowles Foundation Discussion Paper No. 1670, 2008.
- B. D. Bernheim and A. Rangel. Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124(1):51–104, 2009.
- M. Brown, C. Flinn, and A. Schotter. Real-time search in the laboratory and the market. *American Economic Review*, 101(2):948–974, 2011.
- A. Caplin and M. Dean. Search, choice and revealed preference. *Theoretical Economics*, 6: 19–48, 2011.
- V. Cherepanov, T. Feddersen, and A. Sandroni. Rationalization. mimeo, Northwestern University, 2008.
- I. Chioveanu and J. Zhou. Price competition with consumer confusion. Working papers, UCL, 2009.
- K. Eliaz and R. Spiegel. Consideration sets and competitive marketing. *The Review of Economic Studies*, 78(1):235–262, 2011.
- X. Gabaix, D. Laibson, G. Moloche, and S. Weinberg. Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96 (4):1043–1068, 2006.
- G. Gigerenzer and R. Selten. *Rethinking Rationality*, chapter 1, pages 1–12. MIT Press, 2001.
- M. Goeree. Limited information and advertising in the us personal computer industry. *Econometrica*, 76:1017–1074, 2007.
- J. Green and D. Hojman. Choice, rationality and welfare measurement. mimeo, Harvard Institute of Economic Research Discussion Paper No. 2144., 2008.
- F. Gul and W. Pesendorfer. The case for mindless economics. mimeo, 2006.
- J. R. Hauser. Consideration-set heuristics. 2010.
- J. R. Hauser and B. Wernerfelt. An evaluation cost model of consideration sets. *The Journal of Consumer Research*, 16(4):393–408, 1990.
- K. Hendricks, A. Sorensen, and T. Wiseman. Observational learning and demand for search goods. *American Economic Journal: Microeconomics*, 4(1):1–31, 2012.
- J. D. Hey. Search for rules for search. *Journal of Economic Behavior and Organization*, 3 (1):65 – 81, 1982.

- J. D. Hey. Still searching. *Journal of Economic Behavior and Organization*, 8(1):137 – 144, 1987.
- S. Horan. Sequential search and choice from lists. 2010.
- P. Jeziorski and I. Segal. What makes them click: Empirical analysis of consumer demand for search advertising. Economics working paper archive, The Johns Hopkins University, Department of Economics, 2010.
- C. Kogut. Consumer search behavior and sunk costs. *Journal of Economic Behavior & Organization*, 14(3):381–392, 1990.
- B. Köszegi and M. Rabin. A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4):1133–1165, 2006.
- P. Manzini and M. Mariotti. Sequentially rationalizable choice. *American Economic Review*, 97(5):1824–1839, December 2007.
- P. Manzini and M. Mariotti. Consumer choice and revealed bounded rationality. *Economic Theory*, 41(3):379–392, 2009.
- P. Manzini and M. Mariotti. Categorize then choose: Boundedly rational choice and welfare. *Forthcoming in Journal of the European Economic Association*, 2012.
- P. Manzini, M. Mariotti, and C. Tyson. Manipulation of choice behavior. SSRN Working Paper, 2011.
- Y. Masatlioglu and E. A. Ok. Rational choice with status quo bias. *Journal of Economic Theory*, 121:1–29, 2005.
- Y. Masatlioglu and E. A. Ok. A canonical model of choice with initial endowments. mimeo, 2010.
- Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay. Revealed attention. *American Economic Review*, 102(5):2183–2205, 2012.
- J. McCall. Economics of information and job search. *The Quarterly Journal of Economics*, 84(1):113–126, 1970.
- S. Moorthy, B. T. Ratchford, and D. Talukdar. Consumer information search revisited: Theory and empirical analysis. *The Journal of Consumer Research*, 23(4):263–277, 1997.
- D. Mortensen. Job search, the duration of unemployment, and the Phillips curve. *American Economic Review*, 60(5):847–862, 1970.
- P. Nedungadi. Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *The Journal of Consumer Research*, 17(3):263–276, 1990.
- M. Papi. The satisficing choice procedure. 2010.
- M. Piccione and R. Spiegler. Price competition under limited comparability. *The quarterly journal of economics*, 127(1):97–135, 2012.

- C. Raymond. Behavioral foundations for search. mimeo, University of Michigan, 2010.
- J. H. Roberts and M. J. Lattin. Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 28(4):429–440, 1991.
- Y. Salant and A. Rubinstein. Choice with frames. *Review of Economic Studies*, 75:1287–1296, 2008.
- A. Sen. Internal consistency of choice. *Econometrica*, 61(3):495–521, 1993.
- R. Spiegler and K. Eliaz. On the strategic use of attention grabbers. *Theoretical Economics*, 6(1), 2011.
- G. J. Stigler. The economics of information. *The Journal of Political Economy*, 69(3): 213–225, 1961.
- A. Tversky and D. Kahneman. Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061, 1991.
- R. Zwick, A. Rapoport, A. K. C. Lo, and A. V. Muthukrishnan. Consumer sequential search: Not enough or too much? *Marketing Science*, 22(4):503–519, 2003.

## 7 Appendix

### 7.1 Independence of Axioms 1-4

The following examples illustrate that, in our characterization for the Markovian model, no axiom can be disposed without affecting the characterization. We drop each axiom at a time in their numerical order.

- If the consideration set mapping,  $\Omega^*$ , is allowed to depend on the starting point, one can create examples where A1 is violated and the rest of the axioms are satisfied.
- If the pair  $(\succ, \Omega^*)$  is allowed to depend on the composition of budget set, one can create examples where A2 is violated and the rest of the axioms are satisfied.
- Consider a decision maker who stays with her status quo except once ( $c(X, x) = y \neq x$ ). This example violates A3 and satisfies the rest of the axioms are satisfied.
- If one modifies the model and considers the one-step-limited-search model (search terminates after the first step, i.e.,  $c(S, x) = \arg \max(\succ, \Omega^*(x) \cap S)$ ), this model satisfies A1-3, but violates A4.

### 7.2 Proof of Theorem 2

The proof of the “if” part is left to readers. Now suppose  $c$  satisfies A1-A4. For any distinct  $x$  and  $y$ , define  $x \triangleright_c y$  if  $x = c(S, z)$  and  $y = c(\{y, z\}, z)$  for some  $z \neq x$  and  $y \in S$ .

$$x \triangleright_c y \text{ if } x = c(S, z) \text{ and } y = c(\{y, z\}, z) \text{ for some } z \neq x \text{ and } y \in S.$$

Notice that this definition does not exclude the possibility that  $y$  is equal to  $z$ .

We now show that A1 guarantees  $\triangleright_c$  being acyclical. Assume  $c$  satisfies A1 and  $\triangleright_c$  is cyclic. If elements of  $S$  form a cycle in terms of  $\triangleright_c$ , this means that for every alternative  $x \in S$ , there exist  $z_x$  and  $y \in S \setminus \{x\}$  such that  $c(\{x, z_x\}, z_x) = x$  and  $y = c(T, z_x)$  for some  $T, z$ , which violates A1. Hence,  $\triangleright_c$  cannot have a cycle.

Since  $\triangleright_c$  is acyclical, so we can find a strict ranking  $\succ$  which includes  $\triangleright_c$ . Define  $\Omega^*(x) \equiv \{y \mid c(\{x, y\}, x) = y\}$  and note that  $x \in \Omega^*(x)$  for all  $x$  in  $X$ . Then  $(\succ, \Omega^*)$  generates a Markovian CIS. Let us denote it by  $c^*$ , and we will show that  $c^* = c$ . First, we need to prove several claims.

**Claim 1.** If  $c(S, x) = x$ , then  $c(T, x) = x$  whenever  $T \subset S$ .

Proof of Claim 1: Assume  $c(S, x) = x \neq c(T, x)$  and  $T \subset S$ . By A3, there exists  $z \in T \setminus x$  such that  $c(\{x, z\}, x) = z$ . By applying A4,  $c(S, x) = x$  and  $c(\{x, z\}, x) = z$  imply  $c(S \cup z, x) = c(S \cup z, z)$ . Since  $z$  also belongs to  $S(\supset T)$ ,  $c(S, x) = c(S, z)$ . Hence  $c(S, z) = x$ . A1 is violated because of  $c(S, z) = x$  and  $c(\{x, z\}, x) = z$ .  $\square$

**Claim 2.**  $\Omega^*(x) \cap S = \{x\}$  if and only if  $c(S, x) = x$ .

Proof of Claim 2:  $S \cap \Omega^*(x) = \{x\}$  implies that  $x = c(\{x, y\}, x)$  for all  $y \in S$ . Then, by A3, we have  $c(S, x) = x$ . On the other hand, if  $c(S, x) = x$  then, by Claim 1,  $x = c(\{x, y\}, x)$  for all  $y \in S$ . Hence  $S \cap \Omega^*(x) = \{x\}$ .  $\square$

**Claim 3.** Let  $x \neq y$ . Then  $c(S, x) = c(\{x, y\}, x) = y$  and  $\Omega^*(x) \cap S = \{x, y\}$  implies  $c(T, x) = y$  whenever  $\{x, y\} \subset T \subset S$ .

Proof of Claim 3: By A3, it is  $c(T' \setminus y, x) = x$  because  $c(\{x, z\}, x) = x$  for any  $z \in T' \setminus y$  by construction of  $\Omega^*$  for any  $T' \subset S$ . Together with  $c(\{x, y\}, x) = y$ , we can particularly get  $c(S, y) = c(S, x) = y$  and  $c(T, y) = c(T, x)$  by A4. The former and Claim 1 imply  $c(T, y) = y$  so it must be  $c(T, x) = y$ .  $\square$

**Claim 4.** Let  $x, y, z \in S$  be all distinct alternatives. If  $c(S, x) = c(\{x, y\}, x) = y$  and  $z \in \Omega^*(x)$ , then  $c(T, x) = c(T \setminus z, x)$  for any  $T$  including  $x$  and  $y$ .

Proof of Claim 4: By construction of  $\Omega^*$ , we have  $z = c(\{x, z\}, x)$  so we can get the desired result by proving  $c(\{x, y, z\}, x) = y$  so that A2 is directly applicable. First note that  $c(\{x, y, z\}, x)$  cannot be  $x$  because by Claim 2 (notice that  $y, z \in \Omega^*(x)$ ). To see why it cannot be  $z$ , if so, it must be  $c(\{x, y, z\}, x) = c(\{x, z\}, x) = z$  by A2 and  $z \in \Omega^*(x)$ . Thus, it must be  $z \triangleright_c y$ . On the other hand,  $c(\{x, z\}, x) = z$  and  $c(S, x) = y$  requires  $y \triangleright_c z$ , which is a contradiction.  $\square$

**Claim 5.** Let  $x \neq y$ . Then  $c(S, x) = c(\{x, y\}, x) = y$  implies  $c(T, x) = y$  whenever  $\{x, y\} \subset T \subset S$ .

Proof of Claim 5: Let  $Q = (\Omega^*(x) \setminus \{x, y\}) \cap S$ . By applying Claim 4 repeatedly,  $c(T \setminus Q, x) = c(T, x)$  and  $c(S \setminus Q, x) = c(S, x) = y$ . Since  $T \setminus Q \subset S \setminus Q$  and  $(S \setminus Q) \cap \Omega^*(x) = \{x, y\}$ , Claim 3 requires  $c(T \setminus Q, x) = y$ . Hence,  $c(T, x) = y$ .  $\square$

**Claim 6.** If  $x \succ y$  then  $c(S, x) = c(S \cup y, x)$ .

Proof of Claim 6: Let

$$\begin{aligned} \mathcal{D} &= \{(x, y, S) \mid x \succ y \text{ and } c(S, x) \neq c(S \cup y, x)\}. \\ \mathcal{D}_{\min} &= \{(x, y, S) \in \mathcal{D} \mid \text{there is no } (x', y', S') \in \mathcal{D} \text{ and } S' \subsetneq S\} \end{aligned}$$

We shall show that  $\mathcal{D}_{\min}$  is empty. Then,  $\mathcal{D}$  must be empty as well, which is the statement of Claim 6.

**Observation 1.** If  $(x, y, S) \in \mathcal{D}$ , then (i)  $c(T, x) \neq y$  for any  $T$ , (ii)  $c(\{x, y\}, x) = x$  (iii)  $x, y, c(S, x), c(S \cup y, x)$  are all distinct and (iv)  $c(\{x, c(S \cup y, x)\}, x) = x$ .

Proof of Observation 1:

- (i) If not, we have  $y \triangleright_c x$ , which contradicts to  $x \succ y$ .
- (ii) This is just a special case of (i) when  $T = \{x, y\}$ .
- (iii) If either  $c(S, x)$  or  $c(S \cup y, x)$  is equal to  $y$ , it contradicts  $x \succ y$ . If  $x = c(S, x)$ , by (ii) and A3, we have  $x = c(S, x) = c(S \cup y, x)$ , a contradiction. If  $x = c(S \cup y, x)$ , by Claim 1, we have  $x = c(S, x) = c(S \cup y, x)$ , a contradiction.
- (iv) If  $c(\{x, c(S \cup y, x)\}, x) = c(S \cup y, x)$ , by Claim 5, we have  $c(S, x) = c(S \cup y, x)$ , a contradiction.  $\triangle$

**Observation 2.** If  $(x, y, S) \in \mathcal{D}_{\min}$ , there exists unique  $t \in S \setminus x$  such that  $c(\{x, t\}, x) = t$ .

Proof of Observation 2:

Existence: Suppose there is no such  $t$ . Then,  $c(S, x) = x$  by A3, which is a contradiction (Observation 1 (iii)).

Uniqueness: Suppose that there is another element  $t' \in S \setminus x$  such that  $c(\{x, t'\}, x) = t'$ . Then,  $c(\{x, t, t', x\}, x) \neq x$  by Claim 2. Without loss of generality, let  $c(\{x, t, t'\}, x) = c(\{x, t\}, x) = t$ . Then by A2,  $c(T, x) = c(T \cup t', x)$  whenever  $t \in T$ . Particularly, we have  $c(S \setminus t', x) = c(S, x)$  and  $c(S \cup y \setminus t', x) = c(S \cup y, x)$ . Since  $c(S, x) \neq c(S \cup y, x)$ , we have  $c(S \setminus t', x) \neq c(S \cup y \setminus t', x)$ . Thus,  $(x, y, S \setminus t') \in \mathcal{D}$ , which contradicts  $(x, y, S) \in \mathcal{D}_{\min}$ .  $\triangle$

**Observation 3.** If  $(x, y, S) \in \mathcal{D}_{\min}$  and let  $z = c(S, x)$  and  $z' = c(S \cup y, x)$ . Then there exists  $w \in S \setminus z$  such that  $(w, y, S) \in \mathcal{D}_{\min}$ ,  $c(\{w, z\}, w) = z$ ,  $z = c(S, w)$  and  $z' = c(S \cup y, w)$ .

Proof of Observation 3:

(The easy case) If  $z = c(\{x, z\}, x)$ , then let  $w = x$ .

(The difficult case) Suppose  $x = c(\{x, z\}, x)$ . By Observation 2,  $S \setminus x$  has only one element  $t$  such that  $t = c(\{x, t\}, x)$ . Thus, A3 implies  $c(S \setminus t, x) = x$ , so  $c(S, x) = c(S, t)$  by A4. Similarly, we get  $c(S \cup y, x) = c(S \cup y, t)$  because  $c(\{x, y\}, x) = x$  by Observation 1. Therefore, we have  $z = c(S, t) \neq c(S \cup y, t) = z'$ . Furthermore, it must  $t \succ y$  because  $t \triangleright_c x \succ y$ . Hence,  $(t, y, S) \in \mathcal{D}_{\min}$ . Applying Observation 1 (iii), it must be  $t \neq z$  so  $t = c(\{x, t\}, x)$  and  $z = c(S, x)$  implies  $z \triangleright_c t$ . Thus, it must be  $z \succ t \succ x$ .

If  $z = c(\{t, z\}, t)$ , then let  $w = t$ . If  $t = c(\{t, z\}, t)$ , repeating this argument (the difficult case), we can find another  $t' \in S \setminus x$  such that  $(t', y, S) \in \mathcal{D}_{\min}$  and  $z \succ t' \succ t$ . If  $t' = c(\{t', z\}, t')$ , repeat this process. Each time, we find a better alternative than the previous one so this case cannot be repeated forever. At some point, we must go to the easy case.  $\triangle$

Given these observations, we shall complete the proof of the claim. Suppose  $(x, y, S) \in \mathcal{D}_{\min}$  and  $z = c(S, x)$ ,  $z' = c(S \cup y, x)$ . Let  $w$  be the element satisfying the conditions in Observation 3. By Observation 1 and Observation 2,  $c(\{w, z\}, w) = z (\neq w)$  and  $c(\{w, w'\}, w) = w$  for all  $w' \in S \setminus z$ . Together with  $c(\{w, y\}, w) = w$ , A3 implies  $c(T, w) = w$  for any  $T \subset S \cup y$  so it must be  $c(T, w) = c(T, z)$  by A4. Particularly, consider  $T = S, S \cup y$ , and  $\{w, y, z\}$ , we have  $c(S, z) = c(S, w) = z$ ,  $c(S \cup y, z) = c(S \cup y, w) = z'$ , and  $c(\{w, y, z\}, z) = c(\{w, y, z\}, w)$ . By Claim 2,  $c(\{w, y, z\}, w)$  is not  $w$  because  $c(\{w, z\}, w) = z$  and not  $y$  because  $w \succ y$  so it is  $z$ . Applying A3 to the first and the third, we get  $c(S \cup y, z) = z$ , which is a contradiction.  $\square$

We next define the concept of replacement of a dominated starting point. That is, if  $y$  is a replacement of  $x$  at  $S$ , then (i) the DM will reach the final choice in both decision problems  $(T, x)$  and  $(T, y)$  for all subsets of  $S$  including both  $x$  and  $y$ , and (ii) any alternative in  $\Omega^*$  cannot be chosen when the starting point is  $y$ . Formally,

**Definition 1.** Suppose  $c(S, x) \neq x$ . Then  $y \in S \setminus x$  is called a replacement of  $x$  at  $S$  if and only if (i)  $c(T, x) = c(T, y)$  for any  $T$  such that  $\{x, y\} \subset T \subset S$  (ii) for any  $z \in S$  if  $c(\{x, z\}, x) = z$ , then  $c(\{y, z\}, y) = y$ .

**Claim 7.** Any replacement of  $x$  at  $S$  belongs to  $\Omega^*(x)$ .

Proof of Claim 7: Let  $y$  be one of replacements of  $x$  at  $S$ . (i) implies that  $c(\{x, y\}, x) = c(\{x, y\}, y)$ . Substituting  $z = x$  into (ii) says if  $c(\{x, x\}, x) = x$  (always true) then  $c(\{x, y\}, y) = y$ . Hence, it must be  $c(\{x, y\}, x) = y$ , i.e.,  $y \in \Omega^*(x)$ .  $\square$

**Claim 8.** Suppose  $c(S, x) \neq x$ . Then  $c(S \cap \Omega^*(x), x)$  is the unique replacement for  $x$  in  $S$ .

Proof of Claim 8: By Claim 2, we have  $S \cap \Omega^*(x) \neq x$ . Let  $y = c(S \cap \Omega^*(x), x)$ . Then  $y$  must be the  $\succ$ -best element in  $S \cap \Omega^*(x)$ . This is because  $y = c(\{x, y, z\}, x)$  for all  $z \in S \cap \Omega^*(x)$  by Claim 5. Hence, we have  $y \triangleright_c z$  for all  $z \in \Omega^*(x) \setminus y$ . Since  $\succ$  is a completion of  $\triangleright_c$ ,  $y$  must be the  $\succ$ -best element in  $S \cap \Omega^*(x)$ .

Next, we shall prove that  $y$  satisfies (i) and (ii). Take any  $T$  such that  $\{x, y\} \subset T \subset S$ . Let

$$T' = (T \setminus \Omega^*(x)) \cup \{x, y\}$$

By the definition of  $\Omega^*$ ,  $T'$  contains only one element ( $y$ ) such that the DM is willing to move from  $x$  in a pairwise comparison ( $c(T' \setminus y, x) = x$ ). Thus, A4 implies

$$c(T', x) = c(T', y)$$

Since we have  $y \triangleright_c z$  for all  $z \in \Omega^*(x) \setminus y$ , by applying Claim 6 recursively, we have  $c(T', y) = c(T, y)$ . Similarly, since  $c(\{x, y, z\}, x) = c(\{x, y\}, x)$  for all  $z \in \Omega^*(x) \setminus y$ , by applying A2 recursively, we have  $c(T', x) = c(T, x)$ . Thus, we have  $c(T, x) = c(T, y)$  so  $y$  satisfies (i). To see (ii), notice that the presumption implies  $z \in \Omega^*(x)$  so if it was  $z = c(\{y, z\}, y)$  then we would have  $z \triangleright_c y$ , but we have already shown that  $y \triangleright_c z$  for all  $z \in \Omega^*(x) \cap x \setminus y$ , which contradicts A1.

Now, we shall show the uniqueness. Suppose  $z \neq y = c(S \cap \Omega^*(x), x)$  is the replacement. Then it must be  $z \in \Omega^*(x)$  by Claim 7. We have already proven that  $c(\{x, y, z\}, x) = y$ . Since  $z$  satisfies (i), we must have  $c(\{x, y, z\}, z) = y$ . A3 and  $z \in \Omega^*(x)$  imply  $c(\{y, z\}, z) = y$ . This contradicts to the fact that  $z \in \Omega^*(x)$  and (ii) imply  $c(\{y, z\}, z) = z$ .  $\square$

**Claim 9.**  $c(S \cap \Omega^*(x), x) = c^*(S \cap \Omega^*(x), x)$  for any  $(S, x)$ .

Proof of Claim 9: If  $S \cap \Omega^*(x)$  includes only  $x$ , the statement is trivial so assume that it includes some elements other than  $x$ , hence  $c(S \cap \Omega^*(x), x) \neq x$  by Claim 2. Let  $y$  be the  $\succ$ -best element in  $S \cap \Omega^*(x)$ . By definition of  $c^*$ ,  $y = c^*(S \cap \Omega^*(x), x)$ . Suppose  $c(S \cap \Omega^*(x), x) = z \neq y$ . Since  $c(S \cap \Omega^*(x), x) \neq x$ , Claim 8 implies  $z$  is a replacement for  $x$  in both  $S$  and  $S \cap \Omega^*(x)$  (since  $(S \cap \Omega^*(x)) \cap \Omega^*(x) = S \cap \Omega^*(x)$ ). By the definition of a replacement,  $c(\{x, y, z\}, x) = c(\{x, y, z\}, z)$ .

We also know that  $c(\{y, z\}, z) = z$  for all  $y \in S \cap \Omega^*(x)$  since  $z$  is a replacement. By A3, we have  $c(S \cap \Omega^*(x), z) = z$ . By Claim 5,  $c(\{x, y, z\}, z) = z$ , so  $c(\{x, y, z\}, x) = z$ . On the other hand,  $y = c(\{x, y\}, x)$  since  $y \in \Omega^*(x)$ . These two imply  $z \triangleright_c y$ , i.e.  $z \succ y$ . This contradicts that  $y$  is the  $\succ$ -best element within  $S \cap \Omega^*(x)$ .  $\square$

**Claim 10.**  $c(S, x) = c^*(S, x)$  for any  $(S, x)$ .

Proof of Claim 10: For any  $S$ , make its partition recursively as follows:

$$I_0^S = \{x \in S \mid S \cap \Omega^*(x) = \{x\}\}$$

and

$$I_k^S = \{x \in S \setminus (I_0^S \cup \dots \cup I_{k-1}^S) \mid \Omega^*(x) \cap S \subset \{x\} \cup I_0^S \cup \dots \cup I_{k-1}^S\}$$

We will prove the assertion of Claim 10 by induction. Take any  $x \in I_0^S$ . Then we have  $x = c^*(S, x)$ . By Claim 2,  $c(S, x) = x$  as well.

Now, assume  $c = c^*$  for any  $x \in I_{k'}^S$  with  $k' \leq k-1$ . We shall show that  $c(S, x) = c^*(S, x)$  for any  $x \in I_k^S$ .

Since  $k > 0$ ,  $S \cap \Omega^*(x)$  includes some element other than  $x$ . Therefore, by Claim 2, it cannot be  $x = c(S, x)$ . Let  $y$  be the  $\succ$ -best element in  $S \cap \Omega^*(x)$ . Then by definition of  $c^*$ , we have  $y = c^*(S \cap \Omega^*(x), x)$  and  $c^*(S, x) = c^*(S, y)$ . Then

$$\begin{aligned} c(S, x) &= c(S, c(S \cap \Omega^*(x), x)) \text{ by Claim 8} \\ &= c(S, c^*(S \cap \Omega^*(x), x)) \text{ by Claim 9} \\ &= c(S, y) \end{aligned}$$

Finally, since  $x \in I_k^S$  and  $y \in S \cap \Omega^*(x)$ , it must be  $y \in I_l^S$  for some  $l < k$ . Therefore, by the inductive hypothesis we have  $c(S, y) = c^*(S, y)$ . Therefore,  $c(S, x) = c^*(S, x)$ .  $\square$

By Claim 10, we have reached the required result.  $\blacksquare$

### 7.3 Proof of Proposition 2

The only-if part has been shown in the main body so we shall give the proof of the if-part. Take any completion of  $\gg'_c$ , denoted by  $\succ$ . We shall show that  $c = c_{(\succ, \Omega^*)}$ .

Case 1:  $c(S, x) = x$ . Take any element  $y \in \Omega^*(x) \cap S$  ( $y \neq x$ ) if any. By the definition of  $\gg'_c$ ,  $x \gg'_c y$  so it must be  $x \succ y$ . Thus,  $c_{(\succ, \Omega^*)}(S, x) = x$ .

Case 2:  $c(S, x) = y \neq x$ . Then, we shall show that  $c(S, x) = c(S, x')$  where  $x'$  is the  $\succ$ -best element in  $\Omega^*(x) \cap S$ . By recursively applying this result, we can conclude  $c(S, x) = c_{(\succ, \Omega^*)}(S, x)$ .

Claim 7 and 8 in the proof of Theorem 2 shows that if  $S' = \{a \in S \mid c(\{a, x\}, x) = a\}$  and  $z = c(S', x)$  then (a)  $z \neq x$ , (b)  $c(T, x) = c(T, z)$  for all  $T \subset S$ , and (c)  $c(\{a, z\}, z) = z$  for any  $a \in S'$ .

We shall show that  $z$  is the  $\succ$ -best element in  $\Omega^*(x) \cap S$ . It is easy to see  $z \in \Omega^*(x)$  because  $c(\{x, z\}, x) = z$  (by definition of  $S'$ ) and the contrapositive of (ii)). Thus, we only need to show  $z \succ a$  for all  $a \in \Omega^*(x) \cap S \setminus \{z\}$ .

If  $a \in S'$ , then it must be  $a \in \Omega^*(x)$  as well so it must be  $z \gg'_c a$  so  $z \succ a$ . If  $a \notin S'$ , then it must be  $x \gg'_c a$  because  $c(\{a, x\}, x) = x$  (Let  $z = x$  into the definition of  $\gg'_c$ ). Since  $z \gg'_c x$  (because  $x \in S'$ ), we obtain  $z \gg'_c x \gg'_c a$ . Thus, it must be  $z \succ a$ .  $\blacksquare$

### 7.4 Proof of Proposition 3

Suppose choice correspondence  $C$  is induced by a CIS represented by  $(\succ, \Omega)$ . Then the  $\succ$ -best element in  $S$  will be a bliss point of  $S$ . This is because for any  $T \subset S$ ,  $\Omega(\cdot, T)$  never includes an element which is  $\succ$ -better than  $x$  so  $x = c(T, x)$  for all  $T \subset S$  (as long as  $x \in T$ ).

Thus, we have  $x \in C(T)$  for any  $T \subset S$  including  $x$ .

Conversely, let  $C$  satisfy the BP axiom. To construct  $\succ$  from the choice correspondence, we utilize the notion of bliss points. The BP axiom guarantees the existence of a bliss point in any choice problem, particularly in  $X$ . Since bliss points of  $X$  are never eliminated, it is natural to put them on the top of the ranking. In the next step, remove all the bliss points of  $X$  from  $X$ . Now consider bliss points of this strictly smaller set and put them into the second layer, and so on. Formally, we recursively define  $\succ$  starting from the grand set,  $X$  and construct a partition of  $X$ . Let  $X = X_0$  and

$$I_0 = \{x \in X_0 \mid x \in C(T) \text{ for all } T \text{ s.t. } x \in T \subset X_0\}.$$

Note that  $I_0$  is the set of bliss points of  $X_0$ . BP implies that  $I_0$  is non-empty. Define  $X_1 = X_0 \setminus I_0$ . If it is non-empty, define

$$I_1 = \{x \in X_1 \mid x \in C(T) \text{ for all } T \text{ s.t. } x \in T \subset X_1\}.$$

BP also implies that  $I_1$  is non-empty. Then define recursively,

$$I_k = \{x \in X_k \mid x \in C(T) \text{ for all } T \text{ s.t. } x \in T \subset X_k\}$$

where  $X_k = X_{k-1} \setminus I_{k-1}$  until  $\bigcup I_k = X$ . Note that  $\{I_n\}$  is a partition of  $X$ , and  $X_n$ 's are nested. Given the partition  $\{I_n\}$ , define  $x \succ y$  if  $x \in I_k$ ,  $y \in I_l$  and  $k < l$ . Since  $X$  is finite, this recursive process will end in finite time. Since it is incomplete as of now, take any completion of it. By construction,  $\succ$  is a preference. Now define

$$\Omega(A, S) = \begin{cases} A \cup \{y_S\} & \text{if } A = \{x\} \text{ and } x \notin C(S) \\ A & \text{otherwise} \end{cases}$$

where  $y_S$  is the  $\succ$ -best element in  $S$  with  $y_S \succ x$ . Note that  $A \subset \Omega(A, S) \subset S$  for all  $A \subset S$ .

Let  $c$  be a CIS represented by  $(\succ, \Omega)$ . We need to show that  $c$  induces  $C$ . Take any  $x \in S$ . If  $x \in C(S)$ , then  $x = c(S, x)$  because  $\Omega(\{x\}, S) = \{x\}$  so  $x \in C_c(S)$ . On the other hand, suppose  $x \in C_c(S)$ . Since  $C_c(S) = C(S) \cup \{y_S\}$ , we can conclude  $x \in C(S)$  by showing  $y_S \in C(S)$ . Suppose  $y_S \in I_k$ . Since  $y_S$  is  $\succ$ -best element in  $S$ ,  $S \subset X_k$  by definition of  $\succ$ . Since  $y_S$  is a bliss point of  $X_k$ , it must be  $y_S \in C(S)$ . Therefore, we conclude  $C(S) = C_c(S)$ . ■

## 7.5 Proof of Proposition 4

We first prove a claim.

**Claim 11.** *Let  $c$  be a Markovian CIS. Then  $c(S, y) = x$  implies  $c(S, x) = x$ .*

Proof of Claim 11: Assume  $c(S, y) = x$  and  $c$  is represented by  $(\succ, \Omega)$  where  $\Omega$  defined by (1). As we show before that under the Markovian structure, the search stops whenever the best item so far is also best within its own consideration set. Hence  $x$  is the best alternative in  $\Omega^*(x) \cap S$ . Then when we consider the choice problem  $(S, x)$ , we have

$$\Omega(\{x\}, S) = (\Omega^*(x) \cap S) \cup \{x\} = \Omega^*(x) \cap S = \Omega(\Omega^*(x) \cap S, S).$$

Therefore,  $x$  is the best alternative in the final consideration set for the choice problem  $(S, x)$ , i.e.,  $c(S, x) = x$ . □

Now suppose choice correspondence  $C_c$  is induced by a Markovian CIS,  $c$ , represented by  $(\succ, \Omega^*)$ . Then define  $xPy$  if  $x = c(\{x, y\}, x)$  for any distinct  $x$  and  $y$ . By A1,  $P$  is

acyclical. Let  $C$  be represented by  $P$ . We need to show that  $c$  induces  $C$ . Take any  $x \in S$ . If  $x \in C(S)$ , then there exists no  $y \in S$  such that  $yPx$ . By definition of  $P$ ,  $x = c(\{x, y\}, x)$  for all  $y \in S$ . By A3, we have  $c(S, x) = x$ , so  $x \in C_c(S)$ . On the other hand, suppose  $x \in C_c(S)$ . That is, there exists  $z \in S$  such that  $c(S, z) = x$ . By Claim 11, we have  $c(S, x) = x$ . By Claim 1, we have  $c(\{x, y\}, x) = x$  for all  $y \in S$ . Hence there is no  $y \in S$  with  $yPx$ ,  $x \in C(S)$ . Therefore, we conclude  $C(S) = C_c(S)$ .

Conversely, let  $C$  be represented by an acyclical binary relation  $P$ . Take any preference  $\succ$  including  $P$  and define  $\Omega^*(x) = \{y \in X \mid yPx\}$ . Let  $c$  be a Markovian CIS represented by  $(\succ, \Omega^*)$ . Again we need to show that  $c$  induces  $C$ . Take any  $x \in S$ . If  $x \in C(S)$ , then  $\Omega^*(x) \cap S = \emptyset$  by the definition of  $\Omega^*$ . Hence  $c(S, x) = x$ , so  $x \in C_c(S)$ . If  $x \in C_c(S)$  then there exists  $z \in S$  such that  $c(S, z) = x$ . By Claim 11, we have  $c(S, x) = x$ . Therefore, there exists no  $y \in \Omega^*(x) \cap S$  such that  $y \succ x$ . By  $P \subset \succ$  and the definition of  $\Omega^*$ , that is equivalent to say that there exists no  $y \in S$  such that  $yPx$ , so  $x \in C(S)$ . Therefore, we conclude  $C(S) = C_c(S)$ . ■