Decision Making within a Product Network*

Yusufcan Masatlioglu† Elchin Suleymanov‡

November 20, 2017

Abstract

A product network comprises a vast number of goods which are linked to one another. This paper investigates decision-making in this new environment by using revealed preference techniques. A decision maker will search within the network to uncover available goods. Due to the constraint imposed by the network structure and the starting point, the decision maker might not discover all available goods. We illustrate how one can deduce both the decision maker's preference and her product network from observed behavior. We also consider an extension of the model where the decision maker will terminate the search before exhausting all the options (limited search).

Keywords: Product network, search, consideration set, bounded rationality, reference-dependent choice, revealed preference.

JEL classification: D11, D81.

*We thank Martha Bailey, Tilman Börgers, Kfir Eliaz, Shaowei Ke, Heng Liu, Daisuke Nakajima, and Dmitriy Stolyarov for helpful comments.

†University of Maryland, Department of Economics, 3114 Tydings Hall, College Park, MD 20742. Email: yusufcan@umd.edu

‡University of Michigan, Department of Economics, Lorch Hall, 611 Tappan Ave., Ann Arbor, MI 48109. Email: elchin@umich.edu

1 Introduction

Consider a Netflix consumer who is searching for a movie. First, she looks up a particular movie recommended by a friend. Then, Netflix recommends several other films. These recommendations help customers who face the overwhelming task of finding the best movie among a huge catalogue. The recommendations form a *product network* in which a large number of movies are linked to one another. Each movie is a node in the product network, and a recommendation represents a link between two movies (see Figure 1). While one of the first e-commerce websites to introduce product networks was Netflix, nowadays almost every e-commerce site offers product networks.[1]

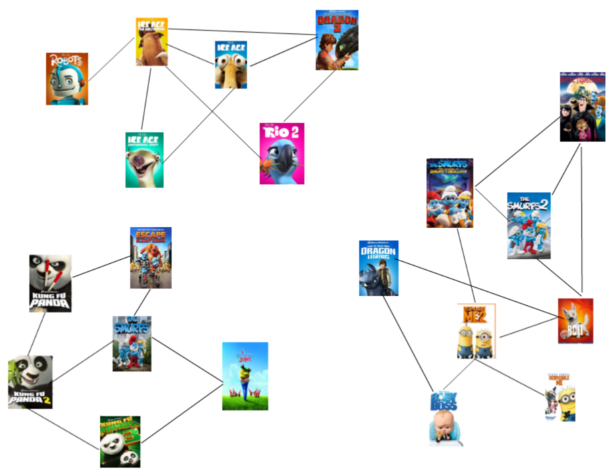

Figure 1: An Example of Product Network

In marketing, researchers recognize the importance of product networks in decision-making (Bucklin and Sismeiro (2003), Hoffman and Novak (1996), Mandel and Johnson (2002)). Moe (2003) claims that "[e]normous potential exists in studying an individual's behavior as they navigate from page to page." This navigation is the key difference between the classical decision-making examined in economics and decision-making within product networks. Shopping within a product network is analogous to walking down the aisles of a supermarket. As in traditional supermarkets, each product has a virtual "shelf position" in the product network, which immensely affects demand for that product (Johnson et al. (2004), Oestreicher-Singer and Sundararajan (2012)). By improving their knowledge of consumer decision-making, firms can improve their marketing strategies. Hence, understanding how a product network shapes consumers' search has become crucial for businesses. Surpris-

[1] As of 2016, Netflix is the world's leading Internet television network with over 93 million members in over 190 countries enjoying over 125 million hours of TV shows and movies per day.

ingly, theoretical work on decision-making within a product network is very limited.

In this paper, we investigate decision-making in this new environment where consumers encounter products in a product network. We consider a customer who picks her most preferred option from the alternatives she can discover (not from the entire feasible set). By studying the observed choice behavior of consumers, our aim is to infer not only their preferences, but also the product network they face. We propose an identification strategy to find out when and how one can deduce both consumer's preferences and the perceived product network she faces from observed choices.[2] Our identification will help the analyst to pinpoint: (i) which products belong to the same subcategory (i.e., connected components)? (ii) which products are more relevant in linking popular products to niche products? (iii) which products are more likely to trigger sales of a particular product? The ability to answer these questions could be of practical importance for firms, regulators, and policy makers.

Throughout this paper, we assume that the product network of a consumer is not observable, as well as the preferences of the consumer. There are three important reasons behind this assumption. First, recommendation systems offer personalized suggestions to match customers with the most appropriate products. Hence, two Netflix customers might face two different product networks because of their unobservable characteristics. Second, there is a distinction between perceived versus exogenous networks. Since the actual search is influenced by different factors such as brand familiarity (Baker et al. (1986)), packaging (Garber (1995)), and color (Aaker (1997)), the perceived network may be different from the product network exogenously provided as in Netflix. The third is that the perceived product network could solely be an outcome of the consumer's memory and the associations she makes between different alternatives. Indeed, cognitive psychologists have illustrated that our memory is organized in an associative network where nodes represent products, and links represent connections between products (Anderson (1983), Anderson and Bower (2014), Collins and Loftus (1975), Hunt and Ellis (1999), Meyer and Schvaneveldt (1976), and Gentner and Stevens (2014)). The fact that the links exist in associative memory does not necessarily mean that the links are activated in reality. Several factors affect the activation of a link such as shape, color, and smell (McCracken and Macklin (1998)). Hence, the presence of a particular product might trigger associated memories and help the consumer recall other associated products.

One of the fundamental assumptions in our model is that the decision maker has well-defined preferences unaffected by the search she undertakes. This assumption has been

[2] The variation in the set of available options is the key for our analysis. Such variations are commonly observed in the real world. For example, Netflix's website removes "Watch Instantly" titles that are not currently available due to different factors, such as delayed license renewals, quality issues, and server technical difficulties. Hence, customers will face different choice sets depending on availability.

used in empirical research in marketing (see Chen and Yao (2016), Honka and Chintagunta (2016), Kim et al. (2010), Koulayev (2010)). Bronnenberg et al. (2016) provide supporting evidence for the assumption by showing that the valuation of an alternative is the same during the search as well as at the time of purchase. Similarly, Reutskaja et al. (2011) provide experimental evidence for the claim that subjects are good at optimizing within the set of options they have explored.

Every search in a network starts from a particular available alternative. We call this alternative the starting point in the consumer's search.[3] In every stage of the search, she will "click on" all the recommendations that appear in her perceived network. Hence, in the baseline model, there is no cognitive limitation other than the one induced by her product network. Our consumer will consider all the goods that are reachable from the starting point. However, as opposed to the standard theory, she might not discover all available alternatives if the product network is sparse.[4]

Our network model is capable of accommodating two seemingly anomalous types of choice behavior. First, given a fixed menu, different starting points can result in different choices (starting point dependence). Moreover, given a fixed starting point, removing irrelevant alternatives from a menu can affect choices (a violation of the weak axiom of revealed preference). While our model helps explain these types of puzzling choice behavior, it also offers distinctive predictions so that we can gain new insights. One of the predictions of the model is that the introduction of a new alternative never diminishes the customer's welfare. Hence, the model generates falsifiable predictions even when the product network are not observable.

Section 2 illustrates how we can infer the preference and the product network of a customer only from observed choices. Identifying preferences is particularly important for welfare analysis. As Bernheim and Rangel (2009) point out, it is typically difficult to identify preference from boundedly rational behavior. Our first result shows that the product network is uniquely identified. However, identifying preferences is tricky. We first provide a list of choice patterns which reveal the consumer's preference. Each of these patterns informs us about the preference of consumer. Hence, even with limited data, it is possible to identify the preferences. This is crucial for empirical studies when the data is limited. We then show that if we have access to complete choice data, all the patterns that reveal preferences can be summarized in a simple manner: x is revealed to be preferred to y if and only if x is chosen when y is the starting point.

[3] Different starting points may result in different consideration sets, and hence different choices (Bronnenberg et al. (2016)). Our initial analysis assumes that we, as the analyst, observe both choices and corresponding starting points. Nevertheless, it is conceivable that starting points will not be observable in some situations. We study the case of unobserved starting points in the Appendix.

[4] The standard model is a special case of our baseline model when the product network is complete (i.e., all the alternatives in the product network are linked).

Our identification strategy relies on the underlying choice procedure where the decision maker maximizes her preference within the alternatives she can reach. It is natural to question the falsifiability of our model. To answer this question, in Section 2, we identify the class of choice behaviors that are consistent with our baseline model. We show that choice data is compatible with the network model if and only if data satisfies three simple behavioral postulates. The first one, Starting Point Contraction, says if the starting point is chosen in a set, then it must be chosen from any subset of it as long as it is still the starting point. The second one, Replacement, states that if an alternative is chosen even though it is not the starting point, then the original starting point and the chosen alternative will induce the same choices in larger sets. The last one, Choice Reversal, dictates that if an alternative causes a choice reversal, then the final choice is not affected if this alternative replaces the starting point. The key feature of our approach is that our assumptions are stated in terms of choice experiments, and therefore a revealed preference type analysis can be used to test our model.

In Section 3, we introduce bounded rationality in our framework. In our baseline model, given the perceived network, the consumer explores all available options which are connected to the starting point. However, many studies show that consumers indeed engage in a very limited search (Goeree (2008), Johnson et al. (2004), Keane et al. (2008), Kim et al. (2010)). We consider a model in which the consumer terminates her search after a number of steps due to limitations such as time pressure or limited cognitive capacity. This model reduces to our baseline model if the number of steps, denoted by K, is larger than the number of products. In the extreme case ($K = 2$), the consumer only considers the starting point and the alternatives which are linked to the starting point. While this model is behaviorally distinct from our baseline model, it coincides with the status quo bias model of Masatlioglu and Ok (2014).

To study K-step network choice, we first provide a characterization for any fixed K. Both Replacement and Choice Reversal hold if the choice set does not have more than K alternatives. In addition, we introduce a new axiom, K-Reduction, which says that if an alternative is chosen in a choice problem, then there must exist a choice problem with at most K alternatives such that the same alternative is chosen. This is a necessary condition in this model since the final choice must be reachable in at most K steps. The baseline model violates this axiom if K is small enough.

Our limited search model highlights an interesting relationship between K-Reduction and K-Replacement. While K-Reduction relates *different* choice problems with the *same* starting point, K-Replacement connects the *same* choice problem with *different* starting points. First, as K decreases, the power of K-Reduction is amplified. While the axiom has no bite when K is big enough, it is very powerful when $K = 2$: removing irrelevant alternatives can never affect choice behavior. Hence, if a consumer follows K-step network choice, given a fixed starting point, fewer violations of "rationality" will be observed as K gets smaller. On the other hand, as K decreases, the power of K-Replacement diminishes. If K is big enough, then replacing the starting point x with another alternative y does not affect the choice as long as x is abandoned for y in a binary comparison. If $K = 2$, then changing the starting point might change the final choice drastically. Hence, a consumer who follows K-step network choice will be more influenced by the starting point as K gets smaller. To summarize, K-step network choice model associates more search (or a higher K) with fewer observations of starting point dependence and more violations of WARP. This discussion highlights the trade-off between the structures imposed by these axioms for different values of K.

Theoretical work on decision making within product networks is limited. The closest paper on the topic that we know of is Masatlioglu and Nakajima (2013). They provide a framework to study behavioral search by utilizing the idea of consideration sets.[5] Their baseline model is quite general and makes very limited predictions. They also consider a particular case which can be represented as a network. However, each model follows completely different choice procedures and they are behaviorally distinct. Their model satisfies *irrelevance of inferior alternatives* and *expansion*, both of which we relax.[6] Hence, on this dimension, their model is more restrictive than ours. On the other hand, their model violates our Replacement axiom. Therefore, these models are distinct.

This paper also contributes to a few branches of decision theory, such as reference dependent choice, limited attention, and search. The other most closely related papers are Masatlioglu and Ok (2005, 2014), Tversky and Kahneman (1991), and Dean et al. (2017) (reference dependent choice); Manzini and Mariotti (2007, 2012), Masatlioglu et al. (2012), and Cherepanov et al. (2013) (limited attention); and Caplin and Dean (2011) (search).

2 Model

The market place consists of a finite set of alternatives, denoted by X. Our consumer's choice problem can be summarized as a pair (S, x) where S is the set of available products at the time of decision and the consumer starts searching from an alternative x in S. Examples of a starting point include (i) the last purchase or status quo, (ii) a product advertised to the consumer, or (iii) a recommendation from someone in the consumer's social network (Masatli-

[5] Caplin and Dean (2011) also study search by employing the revealed preference approach. They assume that an outside observer can view the entire path followed during the search. The main difference is that, in our model, the path is not an input rather an output.

[6] Irrelevance of inferior alternatives states that adding an alternative that is inferior to the starting point cannot affect the choice. Expansion states that if a starting point x dominates both y and z in binary comparisons, then the starting point x will still be chosen when y and z are presented simultaneously.

oglu and Nakajima (2013)). With the explosion of data mining technologies, observability of such data is now plausible.[7] The choice behavior of the consumer is summarized by a choice function c which assigns a single element to each extended choice problem (S, x) where $x \in S$.

We now formally define a product network. A product network $\mathcal{N}$ is a binary relation on X where $(x, y) \in \mathcal{N}$ represents a direct link between x and y. One interpretation of $(x, y) \in \mathcal{N}$ is that y belongs to the set of alternatives recommended to a consumer who considers x. We assume that $\mathcal{N}$ is reflexive and symmetric.[8] Intuitively, this implies that a consumer must always consider the starting point of search. In addition, if a consumer with the starting point x considers y, then she must also consider x when her starting point is y.[9]

We assume that the analyst has no prior knowledge about the product network and tries to pin it down from observed choice behavior. There are two reasons for this assumption. First, it is possible that different consumers face different product networks. This can be due to either personalized recommendations or the use of different stores. For example, while some consumers use Netflix, others use Hulu for movies. Second, even if the exogenous network is the same for everyone (e.g., everyone uses Netflix and gets the same recommendations), it is still possible that different consumers pay attention to different recommended alternatives.

Suppose the consumer faced with the choice problem (S, x) considers all reachable alternatives from x in S in her perceived network. If the perceived network is the same as the exogenous network, this will correspond to a consumer who is not internally constrained (i.e., no search or attention cost), but externally constrained by the network structure. The external constraint imposed by the product network implies that no matter how much search the consumer does, there may be certain alternatives she will never discover. The consumer's consideration set can be written as

$$N_x(S) = \{y \in S|\, \exists\, \{x_1, ..., x_k\} \subseteq S \text{ s.t. } x_1 = x, x_k = y, \text{ and } (x_i, x_{i+1}) \in \mathcal{N} \text{ for } i < k\}. \quad (1)$$

The set $N_x(S)$ denotes all the reachable alternatives from x in S. That is, if $y \in N_x(S)$, then there exists a sequence of linked alternatives connecting x to y (or a path from x to y) in S. We can think of $N_x(S)$ as endogenously determined consideration set of a consumer faced with the choice problem (S, x). We say that c is a network choice if there is a strict preference relation $\succ$ such that for any choice problem (S, x), $c(S, x)$ maximizes $\succ$ in $N_x(S)$.

[7] For unobserved starting points, we refer the interested readers to the Appendix.

[8] A binary relation $\mathcal{N}$ is reflexive if $(x, x) \in \mathcal{N}$ for all $x \in X$. It is symmetric if $(y, x) \in \mathcal{N}$ whenever $(x, y) \in \mathcal{N}$.

[9] Symmetry does not necessarily say that the exogenous network is symmetric. In fact, symmetry may be the result of consumers' symmetric associative memory. That is, if a consumer is recommended y when she considers x, she may recall x when she considers y even in the absence of a direct recommendation. As discussed in the introduction, the availability of y may be necessary for recollection. Recently, Kahana (2002) found evidence suggesting that associative memory is indeed symmetric.

Definition 1. *A choice function c is a **network choice** if there exist a strict preference $\succ$ and a reflexive, symmetric binary relation $\mathcal{N}$ on X and such that*

$$c(S, x) = \text{argmax}(\succ, N_x(S))$$

where $N_x(S)$ is defined as in (1).

Notice that if a product network is complete, that is $(x, y) \in \mathcal{N}$ for all $x, y \in X$, then $N_x(S) = S$ for all S, and hence the choice behavior is compatible with the standard model. Hence, the standard model is a special case of network choice. When the product network is incomplete, seemingly anomalous choice behaviors could be observed. The following example illustrates network choice and the type of choice behavior it allows.

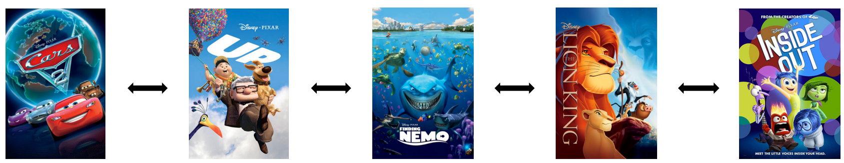

Figure 2: A Linear Network

Example 1. *Altan, who is a six-year-old boy, faces a linear product network which is described in Figure 2.*[10] *Altan's most favorite movie is* Inside Out *and the second best is* Cars.[11] *Altan's starting point is the last movie his older brother has seen. If all the movies are currently available in Netflix, Altan will be aware of all the movies through Netflix's recommendation (because the entire network is connected). Hence, he chooses his most favorite movie* (Inside Out) *independent of the starting point. If* Nemo *becomes unavailable in Netflix, the implied network is no longer connected, and hence the starting point matters. In this case, if the last movie his older brother has seen is* Lion King, *then* Inside Out *is still reachable and will be chosen. However, if the starting point is* Up, *Altan ends up watching* Cars.

In the above example, Altan exhibits two types of choice patterns we mentioned in the introduction. First, changing the starting point affects his final choice (when the choice set is fixed):

$$\text{Inside Out} = c(S, \text{Lion King}) \neq c(S, \text{Up}) = \text{Cars}.$$

where $S = \{\text{Cars, Up, Lion King, Inside Out}\}$. Second, including irrelevant alternatives affects his choice behavior (when the starting point is fixed):

$$\text{Inside Out} = c(S \cup \text{Nemo}, \text{Up}) \neq c(S, \text{Up}) = \text{Cars}$$

[10]We imagine that Netflix recommends at most two available movies in this example. If the movie is one of the extreme points of the network, Netflix provides only one recommendation.

[11]His entire preference is Inside Out $\succ$ Cars $\succ$ Nemo $\succ$ Up $\succ$ Lion King.

These two choice patterns will play crucial role in our analysis later on.

Characterization of Search Behavior

Before we provide behavioral postulates to characterize our choice procedure, we would like to investigate the properties of consideration sets ($N_x(S)$) in this model. These properties will help us to identify and motivate behavioral postulates in the next subsection.

We first define the notion of starting-point-dependent consideration sets. Let $\mathcal{X}(y)$ be the set of all subsets of X including y. We say that $\{\Gamma_x\}_{x\in X}$ is a starting-point-dependent consideration set mapping if for any $x \in X$, the mapping $\Gamma_x : \mathcal{X}(x) \to \mathcal{X}(x)$ satisfies $x \in \Gamma_x(S) \subseteq S$ for all $S \in \mathcal{X}(x)$. Now the question is what type of properties on the consideration sets guarantee that they are induced by a search over a product network.

To provide an answer for this question, we introduce three properties. The first property says that the set of all reachable alternatives does not shrink as the choice problem expands. That is, $\Gamma_x(T) \subseteq \Gamma_x(S)$ whenever $T \subseteq S$ for a fix starting point. The second property says that if y belongs to $\Gamma_x(S)$, the consideration set $\Gamma_y(S)$ must be equal to $\Gamma_x(S)$. In our model, if y is reachable from x in S, then any alternative that is reachable from y must also be reachable from x and vice versa. The last property says that if z belongs to $\Gamma_x(S)$ but not to $\Gamma_x(S \setminus y)$, then y must be considered even in the absence of z, i.e., $y \in \Gamma_x(S \setminus z)$. In addition, the property says that z must be considered in the absence of x when y is the starting point, i.e., $z \in \Gamma_y(S \setminus x)$. In our model, the "if part" of the property implies that y must be in the middle of the path from x to z in the product network. Hence, z must be reachable from y and y must be reachable from x.

We now formally state these properties and provide the result.

A.1 (Upward Monotonicity) $\Gamma_x(T) \subseteq \Gamma_x(S)$ whenever $T \subseteq S$.

A.2 (Symmetry) If $y \in \Gamma_x(S)$, then $\Gamma_y(S) = \Gamma_x(S)$.

A.3 (Path Connectedness) If $z \in \Gamma_x(S)$ and $z \notin \Gamma_x(S\setminus y)$, then $y \in \Gamma_x(S\setminus z)$ and $z \in \Gamma_y(S\setminus x)$ provided that x, y, and z are distinct alternatives.

Lemma 1. $\{\Gamma_x\}_{x\in X}$ *satisfies Upward Monotonicity, Symmetry, and Path Connectedness if and only if there exists a reflexive, symmetric binary relation* $\mathcal{N}$ *such that* $\Gamma_x(S) = N_x(S)$ *for all* $S \in \mathcal{X}$ *and* $x \in S$.[12]

Notice that these properties are defined on consideration sets which are usually unobservable. Nevertheless, Lemma 1 can be used to determine whether a consumer follows network

[12] All proofs are provided in the Appendix.

choice when we have data on the consumer's consideration sets. For example, Reutskaja et al. (2011) find that the average number of items seen by the subjects increases with set size. This suggests an evidence for Upward Monotonicity. In addition, Lemma 1 also allows us to compare our model with existing limited consideration models. We discuss the relationship with other models in Section 3.

Characterization of Choice Behavior

As we discussed before, it is unlikely that an outside observer will have much information on implied consideration sets. We now propose three simple axioms on observed choice behavior. We then discuss how each axiom is related to three properties on consideration sets discussed above.

The first axiom is similar to standard contraction axiom (also known as α-axiom) for a fixed starting point. It says that if the starting point is chosen in some choice problem, then it must also be chosen in any subset of it as long as it stays as the starting point.

Axiom 1. *(Starting Point Contraction) If $c(S, x) = x$ and $T \subseteq S$, then $c(T, x) = x$.*

Axiom 1 is much weaker than α-axiom. This axiom is silent when an alternative different from the starting point is chosen. On the other hand, α-axiom says that removing irrelevant alternatives from a menu cannot affect the choice irrespective of which alternative is chosen. But in our network model, this conclusion is no longer true when the chosen alternative is different from the starting point since an alternative that is considered in a bigger choice set is not necessarily considered in a smaller choice set.

Axiom 1 is directly implied by the first property of consideration sets (A.1). To see this, suppose x is selected when the choice problem is (S, x). This means the starting point is the best among all reachable alternatives. If some of the alternatives are removed, the number of possible paths decreases, and hence the implied consideration set can only get smaller. Since the starting point is always available and it was chosen when the choice set was larger, it must also be chosen when the choice set is smaller.

The second axiom says that if y is chosen in some choice problem (T, x), then the choices corresponding to choice problems (S, x) and (S, y) must be the same for all S containing T. In other words, once the starting point is abandoned for some alternative, replacing the original starting point with the chosen alternative should not affect choices in larger choice sets.

Axiom 2. *(Replacement) If $c(T, x) = y$ and $T \subseteq S$, then $c(S, x) = c(S, y)$.*

To see why Axiom 2 holds in our model, notice that the first and second properties of consideration sets imply a property which we call strong symmetry: if $y \in \Gamma_x(T)$ and $T \subseteq S$, then $\Gamma_x(S) = \Gamma_y(S)$. That is because if $y \in \Gamma_x(T)$ and $T \subseteq S$, then $y \in \Gamma_x(S)$ by A.1, and hence $\Gamma_x(S) = \Gamma_y(S)$ by A.2. Now since y is chosen (and hence considered) when the choice problem is (T, x), by strong symmetry, for any $S \supseteq T$, the consideration sets corresponding to choice problems (S, x) and (S, y) are the same. Therefore, the consumer must choose the same alternative.

From Example 1, we can see that the network model allows for choice reversal patterns which are not allowed by the standard model. However, it does so in a very special way. The last axiom imposes a structure on observed choice reversals. Suppose y is the alternative which causes a choice reversal. Axiom 3 says that if we remove the starting point from the choice set and make y the starting point, then the choice must be the same. Furthermore, if we remove the chosen alternative from the choice set, replacing the original starting point with y leads to the same choice.

Axiom 3. *(Choice Reversal) If $c(S, x) = z \neq c(S \setminus y, x)$, then $c(S \setminus x, y) = z$ and $c(S \setminus z, x) = c(S \setminus z, y)$ provided that x, y, and z are distinct alternatives.*

Axiom 3 is an implication of the properties on consideration sets. To see this suppose z is chosen when the choice problem is (S, x), but it is not chosen when the choice problem is $(S \setminus y, x)$. Since, by A.1, the consideration sets can get only smaller as the choice set becomes smaller, it must be the case that z is not considered when the choice problem is $(S \setminus y, x)$. Then, by A.3, y must be on the path between x and z, that is $y \in \Gamma_x(S \setminus z)$ and $z \in \Gamma_y(S \setminus x)$. Notice that, by A.2, $\Gamma_x(S \setminus z) = \Gamma_y(S \setminus z)$. Therefore, we must have $c(S \setminus z, x) = c(S \setminus z, y)$. By strong symmetry property we defined above, we also have $\Gamma_x(S) = \Gamma_y(S)$. Since $z \in \Gamma_y(S \setminus x) \subseteq \Gamma_y(S)$ and we know that z is the best alternative in $\Gamma_x(S) = \Gamma_y(S)$, z must also be the best alternative in $\Gamma_y(S \setminus x)$. This implies $c(S \setminus x, y) = z$.

The following theorem provides a foundation for the network choice model.

Theorem 1. *A choice function c satisfies Starting Point Contraction, Replacement, and Choice Reversal if and only if it is a network choice.*[13]

Theorem 1 shows that network choice is captured by three simple behavioral postulates. This makes it possible to test our model non-parametrically by using the standard revealed preference technique. We next derive the decision maker's preferences and network from observed choice data.

[13]The fact that the axioms in Theorem 1 and all the subsequent theorems are logically independent is shown in a separate appendix which is available upon request.

Revealed Preference and Network

In this section, we discuss how we can reveal preference and network from choice data given that the consumer follows the network choice. The standard theory suggests that choices directly reveal preferences. That is, x is preferred to y when x is chosen in the presence of y. To justify such an inference, one must implicitly assume that y is considered. In our model, the decision maker is constrained by the network. As a result, the decision maker might not compare all available alternatives before making a choice. Therefore, eliciting preference from choices is no longer trivial. The next example illustrates that there may be multiple preferences representing given choice behavior.

Example 2. *Now consider another decision maker, Mehmet, who is very fond of Lego movies:* Lego Movie, Batman, Ninjago*. While we do not have any knowledge about his underlying choice procedure, we can observe his choice behavior. He always chooses* Ninjago *independent of the starting point as long as it is available. However, in the absence of* Ninjago*, his choice is dictated by the starting point. That is, he watches what he encounters first.*

Since Mehmet's behavior satisfies Starting Point Contraction, Replacement, and Choice Reversal, Theorem 1 guarantees that a network choice representation of his behavior exists. The fact that Mehmet chooses Ninjago from each binary choice problem implies that Ninjago is directly linked to both other movies. On the other hand, since Mehmet chooses different alternatives from {Lego Movie, Batman} depending on the starting point, there is no link between Lego Movie and Batman (notice there is a path between them through Ninjago). Hence, his choices reveal his entire network. This is a general feature of the baseline model. To reveal the underlying network, it suffices to check binary choice data with different starting points. If the decision maker's choices between two options are independent of starting points, then we reveal that these two options are directly linked. Otherwise, there is no direct link (see Figure 3).

Figure 3: Mehmet's Inferred Network (Example 2)

While the entire network is pinned down, revealed preference is not unique in this example. Indeed, Mehmet's preference may be either of the following: Ninjago $\succ_1$ Lego Movie $\succ_1$ Batman and Ninjago $\succ_2$ Batman $\succ_2$ Lego Movie. Since there can be multiple preferences representing choice behavior, we need to define what we mean by revealed preference. Following Masatlioglu et al. (2012), we say that an option is revealed preferred to another option if the first option is ranked higher than the second one in all possible representations.

Definition 2. *Suppose c is a network choice and let $\{(\succ_i, \mathcal{N})|\ i = 1, \ldots, N\}$ be all possible representations of c. Then, we say that x is revealed to be preferred to y if $x \succ_i y$ for all $i = 1, \ldots, N$.*

Once we reveal the network, we can compare two options in terms of preference only if there is a direct link or a path between them. If there is no link or path between these two options, there will be no choice problem in which these alternatives are considered at the same time, and hence we cannot tell which option is more preferred.

One might think that if there is a path between two alternatives, then we can reveal which one is more preferred. However, Example 2 illustrates that the existence of a path is not enough either. In that example, there is a path between Lego Movie and Batman, but we still cannot tell which alternative is more preferred. We can reveal preference between two alternatives if and only if there is a choice set in which (i) a direct link or a path between the two alternatives exists, and (ii) one alternative is chosen over the other.

Consider an observation where the decision maker chooses x when y is the starting point, i.e., $x = c(S, y)$. This observation satisfies both conditions: (i) there is a path between x and y in S, and (ii) x is chosen over y. We can list many other choice patterns that satisfy both (i) and (ii), and hence reveal preference. Consider the following three sets of observations.

1. $x = c(S, z)$ and $y = c(T, z)$ where $T \subset S$,
2. $x = c(S, z)$ and $c(T, y) = c(T, z)$ where $T \subset S$,
3. $x = c(S, z)$ and $y \neq c(T, z) \neq c(T \setminus y, z)$ where $T \subseteq S$.

In all of the observations above, we learn that there is a path between y and z in T. This guarantees that, in any choice set containing T, y must be considered when z is the starting point. Hence, the observation that x is chosen in the choice problem (S, z) implies that x is preferred to y. The list above is far from being exhaustive. For example, suppose $x = c(S, t)$, $c(T, z) = c(T, y)$, and $c(T, z) = c(T', t)$ where $T, T' \subset S$. These observations tell us that there is a path from both t to z and z to y in S. Hence, y is considered when x is chosen.

It may appear that revealed preference can be quite complicated in this setup. However, if we have access to complete choice data and Axiom 1-3 are satisfied, then by applying the axioms we can verify that all of the above observations are possible only if there exists a choice set such that x is chosen when y is the starting point. As an example, Replacement guarantees that $c(S, y) = x$ in the case of Observation 1. The next proposition states that with complete choice data, revealed preference can be summarized by one simple observation: x is preferred to y if and only if x is chosen when y is the starting point.

Proposition 1. *Suppose c is a network choice. Then,*

- *x is revealed preferred to y if and only if $c(S, y) = x$ for some S,*
- *x is revealed to be linked to y if and only if $c(\{x, y\}, x) = c(\{x, y\}, y)$.*

Proposition 1 provides a necessary and sufficient condition for revealed preference in our model. As discussed, the result implicitly assumes that the analyst observes entire choice behavior. With limited dataset, we can use the observations listed above to reveal consumer preference.

We now revisit Example 1 to illustrate how to utilize Proposition 1. We pretend that we do not know either Altan's preferences nor his network structure and try to reveal them from observed choices. Since Altan chooses Inside out when everything is available and the starting point is Cars, this implies that Altan prefers Inside Out over Cars. Similarly, since Cars is chosen when Nemo is the starting point and the menu is $\{\text{Cars}, \text{Nemo}, \text{Up}\}$, Cars is better than Nemo. Choosing Nemo from the choice problem $(\{\text{Nemo}, \text{Up}\}, \text{Up})$ reveals that Nemo is preferred to Up. Finally, $c(\{\text{Nemo}, \text{Up}, \text{Lion King}\}, \text{Lion King}) = \text{Up}$ yields that Lion King is the worst one. Hence, we infer Altan's entire preferences. It is routine to check that all the binary choice problems reveal the product network given in Figure 2.

3 Limited Search

Netflix offers thousands of movies, which make it impossible to do exhaustive search. When people are confronted with overwhelming number of options, they must implement a limited search. In this section, we consider an agent who terminates her search after a certain number of rounds due to limitations such as time pressure or limited cognitive capacity.

We assume that the consumer starts searching from a certain starting point and considers all the alternatives that are linked to the starting point. The decision maker stops search after K steps where $K \geq 2$. For example, if $K = 2$, then the decision maker only considers the starting point and the alternatives which are directly linked to the starting point. If the number of steps is larger than the number of alternatives, then this model reduces to our baseline model. We provide a characterization for K-step network choice where K is fixed and discuss the properties of the consideration sets induced by K-step search.

Consider a consumer faced with the choice problem (S, x). The K-step consideration set is given by

$$N_x^K(S) = \{y \in S | \exists \{x_1, \ldots, x_k\} \subseteq S \text{ such that } x_1 = x, x_k = y, k \leq K, \\ \text{and } (x_i, x_{i+1}) \in \mathcal{N} \text{ for } i < k\} \tag{2}$$

We say that a consumer makes a K-step network choice if the consumer picks the best element in the K-step consideration set.

Definition 3. *A choice function c is a K-step network choice if there exist a strict preference $\succ$ and a reflexive, symmetric binary relation $\mathcal{N}$ on X such that*

$$c(S, x) = \text{argmax}(\succ, N_x^K(S))$$

where $N_x^K(S)$ is as in (2).[14]

The following example illustrates the properties of K-step network choice.

Example 3. *Efe is the little brother of Altan from Example 1. Both brothers share the same preferences and face the same product network. However, Efe always stops the search after two steps ($K = 2$) due to his cognitive limitations. Even though he does a limited search, his choice behavior satisfies WARP for each starting point. Hence, Efe acts "as if" he is a classical utility maximizer for a fixed starting point.*

This example highlights the power of our choice data. The richness of our data prohibits us making wrong claims about the revealed preference. For example, assume that we only observe Efe's choice data when the starting point is Nemo. In this case, Efe's behavior satisfies WARP. If we apply the standard revealed preference, we must conclude that Nemo is strictly preferred to both Inside Out and Cars, which are the best two options for Efe. In other words, we will mistakenly reveal that Inside Out is worse than Nemo even though it is the best alternative for Efe. With richer choice data—having observations with multiple starting points, we can observe that Efe's choices depend on the starting point which is in contrast to the standard model.

Characterization of Search Behavior

Before moving on to characterization, we first show the properties the consideration sets arising from K-step network choice satisfy. First, notice that Upward Monotonicity (A.1) and Path Connectedness (A.3) must still be satisfied in the modified model. However, the

[14] If $K = 2$, then it is without loss of generality to assume that $\mathcal{N}$ is symmetric.

original Symmetry property (A.2) is no longer satsified. That is because in K-step search model, even if y is reachable from x in K steps and z is reachable from y in K steps, this does not necessarily imply that z is reachable from x in K steps. If the choice set has fewer than K alternatives, then Symmetry follows.

For any set S, we use the notation $\mathcal{P}_{\leq K}(S)$ to denote all nonempty subsets of S with at most K elements.

B.2 (K-Symmetry) If $y \in \Gamma_x(S)$ for $S \in \mathcal{P}_{\leq K}(X)$, then $\Gamma_y(S) = \Gamma_x(S)$.

Finally, consideration sets in K-step search model also satisfy K-Reduction which has no analog in the baseline model. This property says that if y is reachable from x in S, then there exists $T \subseteq S$ consisting of at most K alternatives such that y is reachable from x in T. To see why this must be true, consider the set consisting of x, y, and the alternatives connecting x to y. This set must have at most K alternatives.

B.4 (K-Reduction) If $y \in \Gamma_x(S)$, then there exists $T \in \mathcal{P}_{\leq K}(S)$ such that $y \in \Gamma_x(T)$.

Lemma 2 shows that if a collection of consideration set mappings $\{\Gamma_x\}_{x \in X}$ satisfies the four properties described above, then we can treat them as a K-step consideration set on a product network.

Lemma 2. *$\{\Gamma_x\}_{x \in X}$ satisfies Upward Monotonicity, K-Symmetry, Path Connectedness, and K-Reduction if and only if there exists a reflexive, symmetric binary relation $\mathcal{N}$ such that $\Gamma_x(S) = N_x^K(S)$ for all $S \in \mathcal{X}$ and $x \in S$.*

Lemma 2 is useful in comparing our model with existing models of limited consideration. In the recent literature on limited consideration, a decision maker chooses the best alternative from a small subset of the available alternatives. Such models include rational shortlisting (Manzini and Mariotti (2007)), considering alternatives that belong to the best category (Manzini and Mariotti (2012)), considering alternatives that are optimal according to some rationalizing criteria (Cherepanov et al. (2013)), and limited attention (Lleras et al. (2017), Masatlioglu et al. (2012)). While all of these models have an element of limited attention, choices are not affected by a starting point in these models. Even though the domains of these models are different than ours, we contrast our model with these models by fixing the starting point.

These models satisfy one of two following properties. The first condition says that the consideration set is unaffected by removing an alternative which does not attract attention.

$$x \notin \Gamma(S) \text{ implies } \Gamma(S) = \Gamma(S \setminus x).$$

The second captures the idea that attention is relatively scarcer in larger choice sets. That is, if an alternative attracts attention in a larger set, it also attracts attention in subsets of it in which it is included.

$$x \in \Gamma(S) \text{ implies } x \in \Gamma(T) \text{ if } x \in T \subset S.$$

It is routine to show that for a fixed starting point, Upward Monotonicity and Path Connectedness imply the first condition. On the other hand, our model assumes the opposite of the second condition.

Tyson (2013) considers a model in which each menu has an associated preference relation. The smaller the choice sets the more fine-grained the preferences are. The intuition is that in more complex decision problems it is more difficult to compare alternatives. In the first stage, these menu-dependent preferences are used to eliminate some alternatives. In the second stage, ties are broken by a salience relation. From four properties that are satisfied by consideration sets in our model, only Upward Monotonicity holds in his model.

Characterization of Choice Behavior

We now propose four axioms on choices characterizing K-step network choice.

The first axiom says that every choice set has some dominant alternative: for any set S, there exists some alternative x^* that is the best in S. Suppose given a choice problem (T', z), the consumer picks x^*. Then, if we extend the choice set and consider the choice problem (T, z), the consumer must still consider x^*. Since x^* is the best element in S, if the consumer picks an alternative that belongs to S, it must be x^*.

Axiom 4. *(Dominant Alternative) For any S, there exists $x^* \in S$ such that for any $z \in T' \subset T$, if $c(T', z) = x^*$ and $c(T, z) \in S$, then $c(T, z) = x^*$.*

Axiom 4 posits the existence of such an alternative. If we find a choice set such that no alternative satisfying Axiom 4 exists, then we can conclude that the consumer does not follow K-step network choice.

Axiom 5 is a modification of Replacement axiom that we have in the main model. It says that if for some choice set T, the starting point is abandoned for another alternative, then replacing the original starting point with the chosen alternative does not alter the choice for any choice set containing T as long as the choice set does not have more than K alternatives.

Axiom 5. *(K-Replacement) If $c(T, x) = y$ and $T \subseteq S \in \mathcal{P}_{\leq K}(X)$, then $c(S, x) = c(S, y)$.*

To see why it holds in our model, suppose the consumer chooses y when the choice problem is (T, x). Then, y is reachable from x in T in K steps. By Upward Monotonicity and K-Symmetry properties of consideration sets, for any choice set S with at most K alternatives that includes T, the consideration sets corresponding to choice problems (S, x) and (S, y) must be the same. Therefore, the consumer must make the same choice.

Axiom 6 is a modification of Choice Reversal axiom that we had in the main model. In K-step model, Choice Reversal axiom holds if the choice set does not have more than K alternatives.

Axiom 6. *(K-Choice Reversal) If $c(S, x) = z \neq c(S \backslash y, x)$ for $S \in \mathcal{P}_{\leq K}(X)$, then $c(S \backslash x, y) = z$ and $c(S \setminus z, x) = c(S \setminus z, y)$ provided that x, y, and z are distinct alternatives.*

To see why the original version is no longer true, suppose a choice set S has more than K alternatives and $c(S, x) = z \neq c(S \setminus y, x)$. This tells us that z is reachable from x in S, but it is not reachable from x in $S \setminus y$. Then, y must be on the path from x to z. In other words, z is reachable from y in $S \setminus x$ in K steps. However, it is not necessarily the case that $c(S \setminus x, y) = z$ because there may be another alternative $t \in S \setminus x$ that is more preferred to z and reachable from y in $S \setminus x$ in K steps. The alternative t need not be reachable from x in K steps. By a similar argument, $c(S \setminus z, x) = c(S \setminus z, y)$ is not necessarily true either. If a choice set S has fewer than K alternatives, then we are back to the main model, and Choice Reversal axiom holds.

The last axiom is K-Reduction which says that if y is chosen when the choice problem is (S, x), then there exists $T \subseteq S$ with at most K alternatives such that y is chosen when the choice problem is (T, x).

Axiom 7. *(K-Reduction) If $c(S, x) = y$, then there exists $T \in \mathcal{P}_{\leq K}(S)$ such that $c(T, x) = y$.*

To see why it is true note that if $c(S, x) = y$, then y is reachable from x in S in at most K steps. Consider T which consists of x, y, and the alternatives connecting x to y. Since $c(S, x) = y$ implies that y is preferred to all the other alternatives in T, we must have $c(T, x) = y$.

Theorem 2 says that Axiom 4-7 are necessary and sufficient to characterize K-step network choice.

Theorem 2. *A choice function c satisfies Dominant Alternative, K-Replacement, K-Choice Reversal, and K-Reduction if and only if it is a K-step network choice.*

When $K \geq |X|$, this model becomes equivalent to our baseline model. Hence, Theorem 2 provides an alternative characterization of the baseline model. Notice that in this case K-Reduction is redundant. It must be clear that Dominant Alternative implies Starting Point

Contraction. It can also be shown that Starting Point Contraction and Replacement imply Dominant Alternative.

Theorem 2 also covers another important special case ($K = 2$). Notice that when $K = 2$, K-Replacement states that if $c(\{x, y\}, x) = y$, then $c(\{x, y\}, y) = y$. This is implied by Dominant Alternative if we set $y = z$, $T' = \{y\}$, and $T = S = \{x, y\}$. In addition, 2-Choice Reversal is vacuously true, and 2-Reduction can be written as "$c(S, x) = y \Rightarrow c(\{x, y\}, x) = y$." Hence, Dominant Alternative and 2-Reduction characterize 2-step network choice.

The 2-step model has previously been discussed by Masatlioglu and Ok (2014). However, unlike us they assume observability of choices when there is no status quo. In our domain, this is impossible as every search starts from some starting point. The characterization of their model without assuming observability of choices when there is no status quo was an open question and is answered by Theorem 2.

It is illustrative to compare the characterization in Masatlioglu and Ok (2014) and our characterization. Since they assume observability of choices with no starting points, their axiomatic structure looks quite different. The two axioms in their setup that can be translated to our setting are Weak Axiom of Revealed Preference (WARP) and Weak Status Quo Bias (WSQB). WARP states that if $c(S, x) = y$ and $\{x, y\} \subseteq T \subseteq S$, then $c(T, x) = y$. Obviously, WARP implies 2-Reduction but not vice versa. However, WARP and Dominant Alternative are independent axioms. Since WARP relates choices across different sets for a single starting point and Dominant Alternative relates choices across different sets and different starting points, it is easy to construct choice behavior that satisfies WARP but not Dominant Alternative. In the opposite direction, by Theorem 2, 3-step network choice satisfies Dominant Alternative, but it does not necessarily satisfy WARP. Dominant Alternative combined with 2-Reduction implies WARP.

In our setting, their other axiom, Weak Status Quo Bias, states that "$c(\{x, y\}, x) = y \Rightarrow c(\{x, y\}, y) = y$. This is just 2-Replacement axiom and it is implied by Dominant Alternative as discussed above. Masatlioglu and Ok (2014) also discuss a stronger axiom, Strong Status Quo Bias (SSQB). Even though their model does not satisfy SSQB, replacing WSQB with SSQB in their axiomatic structure provides a characterization of Masatlioglu and Ok (2005). In our setting, SSQB states that if $c(S, x) = y$, then $c(S, y) = y$. This is an implication of Replacement satisfied by our baseline model. Recall that unlike Masatlioglu and Ok (2005) our baseline model does not necessarily satisfy WARP.

Our characterization also allows us to see how an increase in K (or more search) is reflected in choice behavior. We can answer this question by utilizing the characterization provided by Theorem 2. First, K-Replacement gets stronger as K increases. An implication of K-Replacement is different levels of Status Quo Bias. When $K = 2$, we get Weak Status

Quo Bias described above. When $K \geq |X|$, we get Strong Status Quo Bias. For intermediate values of K, we get Level-K Status Quo Bias: If $|S| \leq K$ and $c(S,x) = y$, then $c(S,y) = y$. On the other hand, K-Reduction is weaker for larger values of K. K-Reduction and Dominant Alternative imply Level-K WARP: if $c(S,x) = y$, then there exist at most $K-2$ elements in $S \setminus \{x,y\}$ the removal of which can affect choice behavior. Hence, when $K \geq |X|$, the removal of any alternative from S (other than $\{x,y\}$) can lead to y being unchosen. When $K = 2$, the removal of no alternative can change choice behavior. To summarize, a higher K imposes more structure on choice behavior across different starting points but less structure across different sets for a fixed starting point.

K-step model is also related to the work Tversky and Kahneman (1991). They introduce a reference-dependent choice model in which the reference point affects the utility of an individual by overweighting losses relative to the reference point. In contrast, the DM in our model has a fixed preference and the reference point affects the decision through the attention channel. In terms of choice behavior, their model allows cycles in choices across reference points, i.e., $c(\{x,y\},x) = y$, $c(\{y,z\},y) = z$, and $c(\{x,z\},z) = x$. This type of choice behavior is not allowed in our model. For a fixed starting point, while our model allows choice reversals, their model satisfies WARP.

Dean et al. (2017) propose a reference-dependent model with limited attention. For a choice problem (S,x), the consideration set is given by $\Gamma_x(S) = (\Gamma(S) \cup x) \cap Q(x)$ where $x \in \Gamma(S)$ implies $x \in \Gamma(T)$ if $x \in T \subset S$. The consideration sets of this model violate all our properties. Indeed, they assume downward monotonicity which is the exact opposite of upward monotonicity property in our model. As opposed to our model, they allow the decision maker not to choose the reference point in the smaller choice set while choosing it in the larger choice set.

Revealed Preference and Network

In this section, we discuss how one can reveal preference and network given that the consumer makes K-step network choice. First, notice that since $K \geq 2$ network revelation is exactly the same as in the main model for all K. In particular, x is revealed to be linked to y if and only if $c(\{x,y\},x) = c(\{x,y\},y)$, and x is revealed not to be linked to y if and only if $c(\{x,y\},x) \neq c(\{x,y\},y)$.

Remember Efe from Example 3. We again pretend that we only observe Efe's choices and try to reveal his preferences from his choices. We illustrate how we can infer his preference. Remember,

$$\begin{aligned} \text{c}(\{\text{Cars,Nemo,Up}\},\text{Up}) &= \text{Cars} \\ \text{c}(\{\text{Nemo,Up}\},\text{Up}) &= \text{Nemo} \end{aligned}$$

These choices immediately reveal that Nemo and Cars are strictly better than Up. Since both Nemo and Cars are directly linked to Up, Efe also considered Nemo when Efe has chosen Cars in the first choice observation. Therefore, Cars must be better than Nemo.

We now generalize this observation. For any starting point, an alternative that is chosen in a bigger set must be more preferred. That is an implication of Upward Monotonicity property of consideration sets. For any $x \neq y$, we define

$$xPy \text{ if there exists } z \in X \text{ and } T \subset S \subseteq X \text{ such that } c(S,z) = x \text{ and } c(T,z) = y$$

Let P_R denote the transitive closure of P. It is easy to see that if xP_Ry, then x must be revealed preferred to y. Proposition 2 says that if x is revealed to be preferred to y, then we must also have xP_Ry. In other words, Proposition 2 provides a characterization of the revealed preference in K-step search model.

Proposition 2. *(Revealed Preference) Suppose c is a K-step network choice. Then x is revealed to be preferred to y if and only if xP_Ry.*[15]

Proposition 2 highlights an interesting feature of K-step search model. The revealed preference of this model is independent of the value of K. Even if we are unsure about the exact value of K for a decision maker, we can still do revealed preference analysis. This is useful in applications where we have limited knowledge about K.

Since the baseline model is a special case of the K-step model where K is big enough, one might wonder the relationship between Proposition 1 and 2. In general, the revealed preference defined in Proposition 2 is richer than the one defined in Proposition 2 for each K. While the definition of revealed preferences are different in both propositions, they coincide when K is big enough and we have complete choice data.

4 Conclusion

Many real life decision-making problems involve a search over a product network. In this paper, we show how one can reveal preference and network from individual choice data and provide characterizations of the models of decision making within a product network. We explore the cases of "perfectly rational" and "boundedly rational" consumer.

There are several interesting open questions. First, this paper only discusses symmetric links or undirected product networks. An obvious open question is how the implications of such a model change when the links are asymmetric. Second, while we treat the number

[15] The proof of Proposition 2 directly follows the proof of Theorem 2.

of steps as an exogenously given, one can endogenize the number of search steps that the decision maker takes. It is plausible that the number of search steps depends on the complexity of a product network. Third, one can also think about alternative ways of modeling bounded rationality. For example, there may be a temptation ranking which determines what advertised products the decision maker considers.

Another avenue to explore is to study network choice with a random network. The randomness of a network could arise from two factors: (i) the exogenous product network that we take as given may be random (for example, Netflix's recommendation algorithm may produce random links between alternatives), (ii) the decision maker may pay random attention to presented alternatives.

References

Aaker, J. L. (1997). Dimensions of brand personality. *Journal of Marketing Research*, 34(3):347–356.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3):261–295.

Anderson, J. R. and Bower, G. H. (2014). *Human associative memory*. Psychology press.

Baker, W., Hutchinson, J., Moore, D., and Nedungadi, P. (1986). Brand familiarity and advertising: effects on the evoked set and brand preference. *Advances in Consumer Research*, 13:637.

Bernheim, B. D. and Rangel, A. (2009). Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124(1):51–104.

Bronnenberg, B. J., Kim, J. B., and Mela, C. F. (2016). Zooming in on choice: How do consumers search for cameras online? *Marketing Science*, 35(5):693–712.

Bucklin, R. E. and Sismeiro, C. (2003). A model of web site browsing behavior estimated on clickstream data. *Journal of Marketing Research*, 40(3):249–267.

Caplin, A. and Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, 6(1):19–48.

Chen, Y. and Yao, S. (2016). Sequential search with refinement: Model and application with click-stream data. *Management Science*.

Cherepanov, V., Feddersen, T., and Sandroni, A. (2013). Rationalization. *Theoretical Economics*, 8(3):775–800.

Collins, A. M. and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6):407.

Dean, M., Kıbrıs, Ö., and Masatlioglu, Y. (2017). Limited attention and status quo bias. *Journal of Economic Theory*, 169:93–127.

Garber, J. L. L. (1995). The package appearance in choice. *Advances in Consumer Research*, 22:653–660.

Gentner, D. and Stevens, A. L. (2014). *Mental models*. Psychology Press.

Goeree, M. S. (2008). Limited information and advertising in the US personal computer industry. *Econometrica*, 76(5):1017–1074.

Hoffman, D. L. and Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing*, 60(3):50–68.

Honka, E. and Chintagunta, P. (2016). Simultaneous or sequential? search strategies in the US auto insurance industry. *Marketing Science*, 36(1):21–42.

Hunt, R. R. and Ellis, H. C. (1999). *Fundamentals of cognitive psychology*. New York, NY, US: McGraw-Hill.

Johnson, E. J., Moe, W. W., Fader, P. S., Bellman, S., and Lohse, G. L. (2004). On the depth and dynamics of online search behavior. *Management Science*, 50(3):299–308.

Kahana, M. J. (2002). Associative symmetry and memory theory. *Memory & Cognition*, 30(6):823–840.

Keane, M. T., O'Brien, M., and Smyth, B. (2008). Are people biased in their use of search engines? *Communications of the ACM*, 51(2):49–52.

Kim, J. B., Albuquerque, P., and Bronnenberg, B. J. (2010). Online demand under limited consumer search. *Marketing Science*, 29(6):1001–1023.

Koulayev, S. (2010). Estimating demand in online search markets, with application to hotel bookings. *Boston University*.

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

Mandel, N. and Johnson, E. J. (2002). When web pages influence choice: Effects of visual primes on experts and novices. *Journal of Consumer Research*, 29(2):235–245.

Manzini, P. and Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review*, 97(5):1824–1839.

Manzini, P. and Mariotti, M. (2012). Categorize then choose: Boundedly rational choice and welfare. *Journal of the European Economic Association*, 10(5):1141–1165.

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

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

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

Masatlioglu, Y. and Ok, E. A. (2014). A canonical model of choice with initial endowments. *Review of Economic Studies*, 81(2):851–883.

McCracken, J. C. and Macklin, M. C. (1998). The role of brand names and visual cues in enhancing memory for consumer packaged goods. *Marketing Letters*, 9(2):209–226.

Meyer, D. E. and Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science*, 192(4234):27–33.

Moe, W. W. (2003). Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of Consumer Psychology*, 13(1):29–39.

Oestreicher-Singer, G. and Sundararajan, A. (2012). The visible hand? Demand effects of recommendation networks in electronic markets. *Management Science*, 58(11):1963–1981.

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.

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

Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106(4):1039–1061.

Tyson, C. J. (2013). Behavioral implications of shortlisting procedures. *Social Choice and Welfare*, 41(4):941–963.

Appendix A: Proofs

Proof of Lemma 1

Proof. ($\Leftarrow$) A.1: Suppose $y \in \Gamma_x(T)$. Then, there exists $\{x_1, \ldots, x_k\} \subseteq T \subseteq S$ such that $x_1 = x, x_k = y$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $i < k$. By definition, $y \in \Gamma_x(S)$.

A.2: Suppose $y \in \Gamma_x(S)$. This implies that there exists $\{x_1, \ldots x_j\} \subseteq S$ with $x_1 = x$, $x_j = y$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $i < j$. If $z \in \Gamma_y(S)$, then there exists $\{x_j, \ldots, x_k\} \subseteq S$ such that $x_j = y$, $x_k = z$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $j \leq i < k$. Consider $\{x_1, \ldots, x_j, \ldots, x_k\} \subseteq S$. It satisfies the conditions that $x_1 = x$, $x_k = z$, $(x_i, x_{i+1}) \in \mathcal{N}$ for $i < k$. Therefore, $z \in \Gamma_x(S)$. Now suppose $z \in \Gamma_x(S)$. Firstly, let $x'_i = x_{j-i+1}$ for $i \leq j$. Then, $\{x'_1, \ldots, x'_j\} \subseteq S$ satisfies the conditions that $x'_1 = y$, $x'_j = x$, and $(x'_i, x'_{i+1}) \in \mathcal{N}$ for $i < j$. Since $z \in \Gamma_x(S)$, there exists $\{x'_j, \ldots, x'_k\} \subseteq S$ with $x'_j = x$, $x'_k = z$, and $(x'_i, x'_{i+1}) \in \mathcal{N}$ for $j \leq i < k$. Consider $\{x'_1, \ldots, x'_j, \ldots, x'_k\}$. It satisfies the conditions that $x'_1 = y$, $x'_k = z$, $(x'_i, x'_{i+1}) \in \mathcal{N}$ for $i < k$. Therefore, $z \in \Gamma_y(S)$.

A.3: Suppose $z \in \Gamma_x(S)$ and $z \notin \Gamma_x(S \setminus y)$. Since $z \in \Gamma_x(S)$ there exists $\{x_1, \ldots, x_k\} \subseteq S$ with $x_1 = x$, $x_k = z$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $i < k$. Furthermore, since $z \notin \Gamma_x(S \setminus y)$ there exists $j \in \{2, \ldots, k-1\}$ such that $x_j = y$. Consider $\{x_1, \ldots, x_j\} \subseteq S \setminus z$. It satisfies the conditions that $x_1 = x$, $x_j = y$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $i < j$. Therefore, $y \in \Gamma_x(S \setminus z)$. Now consider $\{x_j, \ldots, x_k\} \subseteq S \setminus x$. It satisfies the conditions that $x_j = y$, $x_k = z$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $j \leq i < k$. Therefore, $z \in \Gamma_y(S \setminus x)$.

($\Rightarrow$) Suppose $\{\Gamma_x\}_{x \in X}$ satisfies A.1-A.3. Let $(x, y) \in \mathcal{N}$ iff $y \in \Gamma_x(\{x, y\})$. Note that if $y \in \Gamma_x(\{x, y\})$, by A.2, $\Gamma_y(\{x, y\}) = \Gamma_x(\{x, y\})$. Therefore, $(x, y) \in \mathcal{N}$ iff $(y, x) \in \mathcal{N}$ for all $x, y \in X$. Given $\mathcal{N}$, we define $N_x(S)$ as in the definition. Note that, by the previous part, $N_x(S)$ defined as such satisfies A.1-A.3. Firstly, we show that $N_x(S) \subseteq \Gamma_x(S)$. Suppose $y \in N_x(S)$. Then, there exists $\{x_1, \ldots, x_k\} \subseteq S$ such that $x_1 = x$, $x_k = y$, and $x_{i+1} \in \Gamma_{x_i}(\{x_i, x_{i+1}\})$ for $i < k$. Therefore, by A.1, $x_{i+1} \in \Gamma_{x_i}(S)$ for $i < k$. By A.3, $\Gamma_{x_{k-1}}(S) = \Gamma_{x_{k-2}}(S) = \cdots = \Gamma_{x_1}(S)$. Then, $y = x_k \in \Gamma_{x_{k-1}}(S) = \Gamma_{x_1}(S) = \Gamma_x(S)$.

Now we show that $\Gamma_x(S) \subseteq N_x(S)$. The proof is by induction. Firstly, note that if $y \in \Gamma_x(\{x, y\})$, then $y \in N_x(\{x, y\})$ by definition. Now suppose for all S with $|S| < n$ we have that $y \in \Gamma_x(S) \Rightarrow y \in N_x(S)$. Pick S with $|S| = n$ and suppose $y \in \Gamma_x(S)$. If there exists $z \in S \setminus \{x, y\}$ such that $y \in \Gamma_x(S \setminus z)$, then since $|S \setminus z| < n$ we have that $y \in N_x(S \setminus z)$ and, by A.1, $y \in N_x(S)$. Now suppose for all $z \in S \setminus \{x, y\}$, $y \notin \Gamma_x(S \setminus z)$. Pick one such z. Then, by A.3, $z \in \Gamma_x(S \setminus y)$ and $y \in \Gamma_z(S \setminus x)$. By induction hypothesis, $z \in N_x(S \setminus y)$ and $y \in N_z(S \setminus x)$. By A.1, $z \in N_x(S)$ and $y \in N_z(S)$. By A.2, $N_x(S) = N_z(S)$, and hence $y \in N_x(S)$. $\square$

Proof of Theorem 1

Necessity is obvious from the discussion in the main text. We prove sufficiency.

For any $x \neq y$, define

$$xPy \text{ if and only if } \exists\, S \supseteq \{x, y\} \text{ such that } c(S, z) = x \text{ for all } z \in S$$

Claim 1. *P is transitive.*

Proof. Suppose $xPyPz$. Then there exist $S \supseteq \{x, y\}$ and $T \supseteq \{y, z\}$ such that $c(S, s) = x$ for all $s \in S$ and $c(T,t) = y$ for all $t \in T$. By Axiom 2, $c(S \cup T, s) = c(S \cup T, x)$ for all $s \in S$ and $c(S \cup T, t) = c(S \cup T, y)$ for all $t \in T$. Since $y \in S$, this implies that $c(S \cup T, s) = c(S \cup T, t)$ for all $s \in S$ and $t \in T$. Now if $c(S \cup T, t) = t$ for some $t \in T \setminus y$, then, by Axiom 1, $c(T,t) = t$ which is a contradiction. Similarly, $c(S \cup T, s) = s$ is not possible for $s \in S \setminus x$. Hence, $c(S \cup T, s) = c(S \cup T, t) = x$ for all $s \in S$ and $t \in T$. By definition, xPz and we are done. □

Now let $\succ$ be a completion of P. We define $\mathcal{N}$ as

$$(x,y) \in \mathcal{N} \text{ if and only if } c(\{x,y\},x) = c(\{x,y\},y)$$

First, note that $\mathcal{N}$ is reflexive and symmetric. Define $N_x(S)$ as

$$N_x(S) = \{y \in S | \exists \{x_1, \ldots, x_k\} \subseteq S \text{ such that } x_1 = x, x_k = y, \\ \text{and } (x_i, x_{i+1}) \in \mathcal{N} \text{ for } i < k\}$$

Claim 2. *$c(S,x) \in N_x(S)$.*

Proof. Firstly, let $S = \{x,y\}$. If $c(\{x,y\},x) = x$, then the result is trivial. If $c(\{x,y\},x) = y$, by Axiom 2, $c(\{x,y\},x) = c(\{x,y\},y)$, and hence $y \in N_x(\{x,y\})$. Now suppose the claim is true for all S with $|S| < n$. Let S with $|S| = n$ be given. If there exists $z \in S$ such that $c(S,x) = c(S \setminus z, x)$, then by induction hypothesis, $c(S,x) \in N_x(S \setminus z)$, and by A.1, $c(S,x) \in N_x(S)$. Suppose for all $z \in S \setminus x$, $c(S,x) \neq c(S \setminus z, x)$. Pick $z \neq c(S,x)$.[16] By Axiom 3, $c(S \setminus x, z) = c(S,x)$. By induction hypothesis, $c(S,x) \in N_z(S \setminus x)$, and by A.1, $c(S,x) \in N_z(S)$. Furthermore, by Axiom 3, $c(S \setminus c(S,x), x) = c(S \setminus c(S,x), z)$. By induction hypothesis, $c(S \setminus c(S,x), x) \in N_x(S \setminus c(S,x))$ and $c(S \setminus c(S,x), x) \in N_z(S \setminus c(S,x))$. Then, by A.2, $N_x(S \setminus c(S,x)) = N_{c(S \setminus c(S,x),x)}(S \setminus c(S,x)) = N_z(S \setminus c(S,x))$ which implies $z \in N_x(S \setminus c(S,x))$, and hence by A.1, $z \in N_x(S)$. Finally, by A.2, $N_x(S) = N_z(S)$ which implies $c(S,x) \in N_x(S)$. □

Claim 3. *If $y \in N_x(S)$, then $c(N_x(S), y) = c(S,x)$*

Proof. Suppose $y \in N_x(S)$. Then, there exists $\{x_1, \ldots, x_k\} \subseteq N_x(S) \subseteq S$ with $x_1 = x$, $x_k = y$ such that $c(\{x_i, x_{i+1}\}, x_i) = c(\{x_i, x_{i+1}\}, x_{i+1})$ for $i < k$. By Axiom 2, $c(N_x(S), x_1) = c(N_x(S), x_2) = \cdots = c(N_x(S), x_n)$ which implies $c(N_x(S), x) = c(N_x(S), y)$. By Claim 2, $c(S,x) \in N_x(S)$, and therefore $c(N_x(S), c(S,x)) = c(N_x(S), x)$. Furthermore, by Axiom 2, $c(S,x) = c(S, c(S,x))$, and by Axiom 1, $c(S, c(S,x)) = c(N_x(S), c(S,x))$. Therefore, $c(N_x(S), y) = c(N_x(S), x) = c(N_x(S), c(S,x)) = c(S,x)$. □

Claim 4. *$c(S,x) = \text{argmax}(\succ, N_x(S))$.*

Proof. By Claim 2, $c(S,x) \in N_x(S)$. By Claim 3, $c(N_x(S), y) = c(S,x)$ for all $y \in N_x(S)$. Therefore, by definition of $\succ$, either $c(S,x) = y$ or $c(S,x) \succ y$. This completes the proof. □

Proof of Proposition 1

[16] We can do this since $|S| > 2$.

Proof. Notice that if c is a network choice and $c(S,y) = x$, then $c(N_y(S),z) = x$ for all $z \in N_y(S)$. The rest of the proof directly follows the proof of Theorem 1. $\square$

Proof of Lemma 2

Proof. Necessity can be easily verified as in the proof of Lemma 1. We prove sufficiency.

Let $(x,y) \in \mathcal{N}$ iff $y \in \Gamma_x(\{x,y\})$. By B.2, $\mathcal{N}$ is symmetric. Given $\mathcal{N}$, define $N_x^K(S)$ as before. Note that $N_x^K(S)$ defined as such must satisfy A.1, A.3, B.2, and B.4. We first show that $N_x^K(S) \subseteq \Gamma_x(S)$. Let $y \in N_x^K(S)$. Then, there exists $\{x_1,\dots,x_k\} \subseteq S$ with $x_1 = x$, $x_k = y$, $k \leq K$ such that $x_{i+1} \in \Gamma_{x_i}(\{x_i,x_{i+1}\})$ for $i < k$. Let $T = \{x_1,\dots,x_k\}$. By A.1, $x_{i+1} \in \Gamma_{x_i}(T)$ for $i < k$. Since $|T| \leq K$, by B.2, $\Gamma_{x_1}(T) = \Gamma_{x_2}(T) = \dots = \Gamma_{x_k}(T)$. Therefore, $y \in \Gamma_x(T)$. Finally, $T \subseteq S$ and A.1 imply $y \in \Gamma_x(S)$.

We now show that $\Gamma_x(S) \subseteq N_x^k(S)$. The proof is by induction. Firstly, if $S = \{x,y\}$, the claim is obvious. Suppose the claim is true for all S with $|S| < n$. Let S with $|S| = n$ be given, and suppose $y \in \Gamma_x(S)$. By B.4, there exists $T \subseteq S$ with $|T| \leq K$ such that $y \in \Gamma_x(T)$. If $T \subset S$, then by induction hypothesis, $y \in N_x^K(T)$ and by A.1, $y \in N_x^K(S)$. If $T = S$, then $|S| \leq K$. Since there exists no strict subset T of S with $y \in \Gamma_x(T)$ we have that $y \notin \Gamma_x(S \setminus z)$ for all $z \in S \setminus x$. Pick z distinct from x and y. Now we have that $|S| \leq K$, $y \in \Gamma_x(S)$, and $y \notin \Gamma_x(S \setminus z)$. By A.3, $z \in \Gamma_x(S \setminus y)$ and $y \in \Gamma_z(S \setminus x)$. Then, by induction hypothesis, $z \in N_x^K(S \setminus y)$ and $y \in N_z^K(S \setminus x)$. By A.1, $z \in N_x^K(S)$ and $y \in N_z^K(S)$. Since $|S| \leq K$, by B.2, $N_x^K(S) = N_z^K(S) = N_y^K(S)$ which implies $y \in N_x^K(S)$. $\square$

Proof of Theorem 2

Necessity is obvious from the discussion in the main text. We prove sufficiency.

Let xPy if there exists $z \in T \subset S$ such that $c(S,z) = x$ and $c(T,z) = y$.

Claim 1. *P is acyclic.*

Proof. Suppose $x_1 P x_2 P \cdots P x_n P x_1$. Then, there exists $\{T_i, T_i', z_i\}_{i=1}^n$ with $z_i \in T_i' \subset T_i$ such that $c(T_i, z_i) = x_i$, $c(T_i', z_i) = x_{i+1}$ for $i < n$, and $c(T_n', z_n) = x_1$. Consider the set $S = \{x_1,\dots,x_n\}$. For all $x \in S$, there exists $z \in T' \subset T$ such that $c(T',z) = x$ and $c(T,z) \in S$, but $c(T,z) \neq x$. This contradicts Axiom 4. $\square$

Let $\succ$ be a completion of P. Define $\mathcal{N}$ as

$$(x,y) \in \mathcal{N} \text{ if and only if } c(\{x,y\},x) = c(\{x,y\},y)$$

and $N_x^K(S)$ as

$$N_x^K(S) = \{y \in S \mid \exists \{x_1,\dots,x_k\} \subseteq S \text{ with } x_1 = x, x_k = y, k \leq K \\ \text{and } (x_i,x_{i+1}) \in \mathcal{N} \text{ for } i < k\}$$

Claim 2. $c(S,x) \in N_x^K(S)$.

Proof. First, note that if $c(\{x,y\},x) = x$, then the claim is trivial. If $c(\{x,y\},x) = y$, then by Axiom 5, $c(\{x,y\},x) = c(\{x,y\},y)$, which implies $y \in N_x^K(\{x,y\})$. Now suppose the claim is true for all S with $|S| < n$. Let S with $|S| = n$ be given. By Axiom 7, there exists $T \subseteq S$ with $|T| \leq K$ such that $c(T,x) = y$. If $T \subset S$, then by induction hypothesis, $c(S,x) \in N_x^K(T)$, and by A.1, $c(S,x) \in N_x^K(S)$. Suppose $T = S$ so that $|S| \leq K$. Since there exists no strict subset T of S with $c(T,x) = y$ we must have that for all $z \in S \setminus x$, $c(S,x) = y \neq c(S \setminus z, x)$. Pick z distinct from x and y. By Axiom 6, $c(S \setminus x, y) = z$ and $c(S \setminus y, x) = c(S \setminus y, z)$. Since $c(S \setminus x, y) = z$, by induction hypothesis, $z \in N_y^K(S \setminus x)$, and by A.1, $z \in N_y^K(S)$. Let $t = c(S \setminus y, x) = c(S \setminus y, z)$. By induction hypothesis, $t \in N_x^K(S \setminus y)$ and $t \in N_z^K(S \setminus y)$. By A.1, $t \in N_x^K(S)$ and $t \in N_z^K(S)$. Since $|S| \leq K$, B.2 implies that $N_x^K(S) = N_t^K(S) = N_z^K(S)$. Furthermore, $z \in N_y^K(S)$ implies $N_y^K(S) = N_z^K(S)$. Therefore, $N_x^K(S) = N_y^K(S)$, and hence $y \in N_x^K(S)$. □

Claim 3. *If $y \in N_x^K(S)$, then there exists $T \subseteq S$ such that $c(T,x) = c(T,y)$.*

Proof. Suppose $y \in N_x^K(S)$. Then, there exists $\{x_1, \ldots, x_k\} \subseteq S$ with $x_1 = x$, $x_k = y$, $k \leq K$, and $(x_i, x_{i+1}) \in \mathcal{N}$ for $i < k$. By definition, $(x_i, x_{i+1}) \in \mathcal{N}$ if and only if $c(\{x_i, x_{i+1}\}, x_i) = c(\{x_i, x_{i+1}\}, x_{i+1})$. Let $T = \{x_1, \ldots, x_k\}$. Since $c(\{x_i, x_{i+1}\}, x_i) = c(\{x_i, x_{i+1}\}, x_{i+1})$ we have that either $c(\{x_i, x_{i+1}\}, x_i) = x_{i+1}$ or $c(\{x_i, x_{i+1}\}, x_{i+1}) = x_i$. Then, since $|T| \leq K$, by Axiom 5, we have $c(T, x_1) = c(T, x_2) = \cdots = c(T, x_n)$. □

Claim 4. $c(S,x) = \text{argmax}(\succ, N_x^K(S))$

Proof. By Claim 2, $c(S,x) \in N_x^K(S)$. Pick $y \in N_x^K(S)$. By Claim 3, there exists $T \subseteq S$ such that $c(T,x) = c(T,y)$. By definition of P, we have either $c(T,x) = y$ or $c(T,x)Py$. Furthermore, since $S \supseteq T$, either $c(S,x) = c(T,x)$ or $c(S,x)Pc(T,x)$. Since $\succ$ includes P, we have that either $c(S,x) = y$ or $c(S,x) \succ y$. □

Appendix B: Unobserved Starting Points

In the main text, we assume that we can observe the starting point of the consumer. Here, we investigate network choice with standard choice data. We first show that if we impose no structure on starting points, then any choice behavior can be justified. Suppose we observe choice function c where $c(S)$ is the element chosen by the consumer when the choice set is S. If our model is correct, then we must have $c(S) = c(S,x)$ where x is a starting point in S. If any alternative in S can be a starting point (i.e., there is no condition on how starting points in different sets are related), then we can let $c(S) = c(S, c(S))$ for all S. That is, the consumer always chooses the starting point. But then any choice behavior is allowed under this model. Hence, the model does not make any prediction.

In what follows, we impose a structure on starting points that helps us infer preferences and network with standard choice data. Following Salant and Rubinstein (2008) and Masatlioglu and Nakajima (2013), we assume that we observe induced choice correspondence where each possible choice corresponds to a different starting point. The induced choice correspondence reflects the data available to an outside observer who knows that the choices of the decision maker are affected by the starting point, but lacks information about the actual starting point. Salant and Rubinstein explore a model in which the decision maker is allowed to make different choices under different frames. Given a choice correspondence C, the model is given by $C(S) = \{x \in S | \, x = c(S, f) \text{ for some } f \in F\}$ where F is the set of frames and $c(S, f)$ is frame dependent choice function. Masatlioglu and Nakajima use a similar idea with starting points.

Suppose the decision maker actually follows a network model denoted by c, but we do not observe her starting point. Let C stand for an induced choice correspondence where for every alternative x in $C(S)$, there exists a starting point y such that $x = c(S, y)$. In other words, x maximizes preference among all reachable alternatives from y in S.

Definition 4. *A choice correspondence C is an induced network choice if there exists a strict preference $\succ$ and a reflexive, symmetric binary relation $\mathcal{N}$ on X such that*

$$C(S) = \{x \in S | x = \text{argmax}(\succ, N_y(S)) \textit{ for some } y \in S\}$$

where $N_y(S)$ is defined as before.

Suppose we observe that the decision maker chooses different alternatives when faced with the same choice set. In standard theory, this would happen only if the decision maker is indifferent between chosen alternatives. However, in our model the decision maker with strict preference over all alternatives may still choose different alternatives when faced with the same choice set if the choice set is *not* connected (there are alternatives in the choice set such that no path between them exists).

Characterization

Before moving on to characterization, notice that using the symmetry property of consideration sets we can write the induced network choice as

$$C(S) = \{x \in S | \, x = \text{argmax}(\succ, N_x(S))\}$$

The alternative representation says that given a choice set S, an alternative x is chosen if and only if it is the best alternative among all the alternatives reachable from x in S. If y is reachable from x in S or vice versa, then the consideration sets corresponding to choice problems (S, x) and (S, y) are the same by A.2. Therefore, the original and the alternative representations are exactly the same.

We propose four simple axioms which characterize induced network choice. Axiom 8 is the standard contraction axiom. It says that if x is chosen when the choice set S, then x must also be chosen in any subset of S containing x.

Axiom 8. *(Contraction) If $x \in C(S)$, then $x \in C(T)$ for all $x \in T \subseteq S$.*

Note that Axiom 8 is a direct implication of the monotonicity property of consideration sets. Since consideration sets can only shrink as the choice set gets smaller, an alternative that is chosen in a bigger choice set must also be chosen in a smaller choice set as long as it is available.

Contraction axiom tells us what we should expect if x is chosen in some choice set S. Axiom 9 tells us what we should expect if x is not chosen. In particular, it posits the existence of an alternative y that dominates x. That is, if x is not chosen when the choice set is S, then there must exist an alternative y and a subset T of S containing x such that y is uniquely chosen.

Axiom 9. *(Dominating Alternative) If $x \notin C(S)$, then there exist $y \in C(S)$ and $T \subseteq S$ containing x such that $C(T) = y$.*

To see why Axiom 9 holds, suppose x is not chosen when the choice set is S. Then, x is not the best element in $N_x(S)$. Suppose the best element in $N_x(S)$ is y, and let $T = N_x(S) \subseteq S$. Then, since T is a connected set, meaning that exists a path between any two alternatives, y must be uniquely chosen when the choice set is T.

Axiom 10 is similar to standard expansion property. It says that if x is uniquely chosen when the choice set is T, y is uniquely chosen when the choice set is S, and T and S have a nonempty intersection, then either x or y must be uniquely chosen when the choice set is $T \cup S$.

Axiom 10. *(Expansion) Suppose $C(T) = x$ and $C(S) = y$. If $T \cap S \neq \emptyset$, then $C(T \cup S) = x$ or y.*

To see why it holds, suppose x is chosen when the choice set is T and y is chosen when the choice set is S. In our model, this can only happen if T and S are connected sets. If T and S have a nonempty intersection, then $T \cup S$ must also be a connected set. Therefore, a unique element must be chosen when the choice set is $T \cup S$. Given that x is the best alternative in T, and y is the best alternative in S, the only possible choices are x and y.

The next property follows from an observation that given a network we can divide any connected set into two connected sets with a nonempty intersection.

Axiom 11. *(Separability) Suppose $|S| \geq 3$. If $C(S) = x$, then there exist non-singleton $T_1, T_2 \subset S$ with $T_1 \cap T_2 \neq \emptyset$ and $T_1 \cup T_2 = S$ such that $C(T_1) = x$ and $C(T_2) = y$ for some $y \in S$.*

Suppose x is uniquely chosen when the choice set is S. Then, S must be a connected set. Given the network structure we can separate S into two connected sets, say T_1 and T_2, with a nonempty intersection. If x is in T_1, then x must be uniquely chosen when the choice set is T_1, and the best element in T_2 must be uniquely chosen when choice set is T_2.

Theorem 3 shows that Axiom 8-11 are necessary and sufficient to characterize the induced network choice.

Theorem 3. *A choice correspondence C satisfies Contraction, Dominating Alternative, Expansion, and Separability if and only if it is an induced network choice.*

Proof of Theorem 3

Necessity is obvious from the previous discussion. We prove sufficiency. Let xPy if there exists $S \supseteq \{x, y\}$ such that $C(S) = x$

Claim 1. *P is acyclic.*

Proof. Suppose $x_1 P x_2 P \cdots P x_n P x_1$. Then, there exist $S_1, \ldots S_n$ with $S_i \supseteq \{x_i, x_{i+1}\}$ for $i < n$ and $S_n \supseteq \{x_1, x_n\}$ such that $C(S_i) = x_i$. Consider the set $T = S_1 \cup S_2 \cup \cdots \cup S_n$. Note that, by Axiom 8, we cannot have $x_i \in C(T)$ since $x_i \in S_{i-1} \subseteq T$ for $i > 1$ and $x_1 \in S_n \subseteq T$, but $C(S_i) = x_i$. Furthermore, we cannot have $y \in C(T)$ for any $y \notin \{x_1, \ldots, x_n\}$ since $y \in S_i$ for some i, but $y \notin C(S_i)$. Hence, we cannot assign any alternative to $C(T)$. Therefore, P is acyclic. □

Let $\succ$ be a completion of P. Define $\mathcal{N}$ as

$$(x, y) \in \mathcal{N} \text{ if and only if } C(\{x, y\}) \text{ is a singleton}$$

and let $N_x(S)$ be given by

$$N_x(S) = \{y \in S | \exists \{x_1, \ldots, x_k\} \subseteq S \text{ with } x_1 = x, x_k = y, \text{ and } (x_i, x_{i+1}) \in \mathcal{N} \text{ for } i < k\}$$

Claim 2. *Let $\{x_1, \ldots, x_k\}$ be such that $C(\{x_i, x_{i+1}\})$ is a singleton for $i < k$. Then, $C(\{x_1, \ldots, x_k\})$ is a singleton.*

Proof. Suppose $C(\{x_1, x_2\})$ and $C(\{x_2, x_3\})$ are singletons. Since $\{x_1, x_2\} \cap \{x_2, x_3\} \neq \emptyset$, by Axiom 10, $C(\{x_1, x_2, x_3\})$ is a singleton. Now suppose $C(\{x_1, \ldots, x_j\})$ and $C(\{x_j, x_{j+1}\})$ are singletons. Since $\{x_1, \ldots, x_j\} \cap \{x_j, x_{j+1}\} \neq \emptyset$, $C(\{x_1, \ldots, x_j, x_{j+1}\})$ is a singleton. Iterating this procedure we get that $C(\{x_1, \ldots, x_k\})$ is a singleton. □

Claim 3. *If $C(S) = y$, then $y \in N_x(S)$ for all $x \in S$.*

Proof. Notice that the claim is trivial if $S = \{x, y\}$. Suppose the claim is true for S with $|S| < n$. Let S with $|S| = n$ be given and suppose $C(S) = y$. By Axiom 11, there exist non-singleton $T_1, T_2 \subset S$ with $T_1 \cap T_2 \neq \emptyset$ and $T_1 \cup T_2 = S$ such that $C(T_1) = y$ and $C(T_2) = z$. Pick $t \in T_1 \cap T_2$. By induction hypothesis, $y \in N_t(T_1)$ and $z \in N_t(T_2)$. Since $T_1, T_2 \subset S$, by A.1, $y, z \in N_t(S)$. Now pick $x \in S$. Either $x \in T_1$ or $x \in T_2$. If $x \in T_1$, then by induction hypothesis, $y \in N_x(T_1)$, and by A.1, $y \in N_x(S)$. If $x \in T_2$, then by induction hypothesis, $z \in N_x(T_2)$, and by A.1, $z \in N_x(S)$. But then, $z \in N_t(S)$ and $z \in N_x(S)$. By A.2, $N_x(S) = N_z(S) = N_t(S)$. Since $y \in N_t(S)$ we should have $y \in N_x(S)$. □

Claim 4. $C(S) = \{x \in S | x = \text{argmax}(\succ, N_x(S))\}$

Proof. Firsty, suppose $x \in C(S)$. We show that $x = \text{argmax}(\succ, N_x(S))$. Pick $z \in N_x(S)$. By definition, there exists $\{x_1, \ldots, x_k\} \subseteq S$ with $x_1 = x$, $x_k = z$ such that $C(\{x_i, x_{i+1}\})$

is a singleton for $i < k$. By Claim 2, $C(\{x_1, \ldots, x_k\})$ is a singleton. Since $x \in C(S)$ and $x \in \{x_1, \ldots, x_k\} \subseteq S$, by Axiom 8, $C(\{x_1, \ldots, x_k\}) = x$. Therefore, xPz, and hence $x \succ z$.

Now suppose $x = \text{argmax}(\succ, N_x(S))$. We show that $x \in C(S)$. Suppose $x \notin C(S)$. Then, by Axiom 9, there exist $y \in C(S)$ and $T \subseteq S$ containing x such that $C(T) = y$. By definition of P, we have yPx and hence $y \succ x$. By Claim 3, $y \in N_x(T)$, and by A.1, $y \in N_x(S)$. This contradicts the fact that $x = \text{argmax}(\succ, N_x(S))$. $\square$