Decision Making within a Product Network*

Yusufcan Masatlioglu† Elchin Suleymanov‡

December 7, 2016

Abstract

A product network consists of a vast number of goods which are linked to one another. This paper investigates decision-making in this new environment by utilizing revealed preference techniques. The decision maker starts her search from a particular starting point. Due to the network structure, the decision maker will not have access to all available alternatives. We illustrate how one can deduce both the decision maker's preference and the network she faces from the observed behavior. We provide two characterizations of the model with observed and unobserved starting points. We also consider extensions of the model with limited search and random network.

Keywords: Product network, search, consideration set, bounded rationality, random choice, Luce rule, revealed preference, revealed network.

JEL classification: D11, D81.

---

*We thank Martha Bailey, Tilman Börgers, Kfir Eliaz, Shaowei Ke, Heng Liu, 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 that she has heard. Then, Netflix recommends several other films. These recommendations create 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 the edge between two movies. While one of the first e-commerce websites to introduce product networks was Netflix, nowadays almost every e-commerce site (Amazon, Barnes & Noble, YouTube, Yelp, iTunes, etc.) offers their product networks. Retail websites use phrases such as "customers who viewed this item also viewed," "you might also like," "similar items," etc., to recommend new products.

In marketing, researchers recognize the importance of product network in decision-making (Hoffman and Novak [1996], Bucklin and Sismeiro [2003], 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. Consumers discover alternatives through search in a product network. Shopping within a product network is analogous to walking the aisles of a supermarket. Each product has a virtual "shelf position" in the product network. As in the traditional supermarkets, the virtual shelf position immensely affects demand for each product (Johnson et al. [2004], Oestreicher-Singer and Sundararajan [2012]). By improving their knowledge of the consumer's decision-making, firms can improve their marketing strategies. Hence understanding how a product network shapes consumers' search has become more crucial for businesses. Surprisingly, theoretical work on the decision making within a product network is very limited.

Throughout the paper, we assume that the product network of a consumer is not observable, as well as the preferences of the consumer. There are two important reasons behind this assumption. Firstly, we want to make a distinction between perceived vs. exogenous networks. Since the actual search can be influenced by different factors such as brand familiarity (Baker et al. [1986]), packaging (Garber [1995]), color (Aaker [1997]), the perceived network may be different than the product network exogenously provided as in Amazon or Netflix. The second is that the perceived product network could solely be an outcome of what our brains remind us of other relevant 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 and Bower [1973], Collins and Loftus [1975], Meyer and Schvaneveldt [1976], Anderson [1983], Ellis and Hunt [1993], and Gentner and Stevens [1983]). The fact that the links exist in memory does not necessarily mean that the links are activated. 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 remembrance and helps us to make associations.

In this paper, we investigate the decision-making in this new environment where consumers encounter products in a product network utilizing revealed preference techniques. We consider a customer who picks her most preferred option from the alternatives she can reach, not from the entire feasible set. As we do not observe either consumer preferences or the product network, a choice can be attributed to either preference or limited search due to the product network. Our premise is that we can infer not only the preferences of consumers but also the network by studying their behaviors. 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. Our identification will help to pinpoint: (i) which products belong to the same sub-category (connected components)? (ii) which products are most relevant regarding linking popular products to niche products? and (iii) which products are most likely to trigger sales of a particular product? The ability to answer these questions could be of vital practical importance for firms, regulators, and policy makers.

One of the fundamental assumption in our model that the decision maker has well-defined preferences, which are not affected by the search she undertakes. This assumption has been utilized in empirical research in marketing (see, Chen and Yao [2015], Honka and Chintagunta [2015], Kim et al. [2010], Koulayev [2010]). Recently, Bronnenberg et al. [2016] provide a supporting evidence by showing that the valuation of an alternative is the same during search and at the time of choice for online search. Similarly, Reutskaja et al. [2011] also provide experimental evidence for the fact that subjects are good at optimizing within options they have explored.

Every search in a network starts from a particular available alternative. We call this alternative a starting point in the consumer's search. Different starting points may result in various 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. The starting point of a search is the alternative that the consumer initially pays attention. There are many potential reasons for a particular starting point. For example, a starting point could be (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 (Masatlioglu and Nakajima [2013]). With the explosion of data mining technologies, observability of such data is plausible nowadays.[1]

In our baseline model, we assume that the perceived product network induces the only limitation the consumer faces. In every stage of the search, she is willing to "click on" all the alternatives that are presented in her perceived network. Hence, there is no other 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 due to the sparsity of the network. The standard model will be a special case of this model where the product network is complete (i.e., all the alternatives in the product network are linked).

Section 2 illustrates how we can infer the preference and the product network of a decision maker 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 result shows that while the product network is uniquely identified, we can partially pin down preferences in our baseline model. We show x is revealed preferred to y if x is chosen from some set including y independent of the starting point. We also show x is revealed to be linked to y if the starting point does not influence the choice between x and y in a binary comparison. These results implicitly assume that the analyst observes the entire choice behavior. We also provide a list of choice patterns which reveal the consumer's preference and the perceived network when we have partial data. This is crucial for empirical studies when the data is limited.

Our identification strategy relies on the underlying choice procedure, where she maximizes her preference within the alternatives she can reach. It is natural to question the falsifiability of our model. To answer this question, in 2, we identify the class of choice behaviors that are consistent with our baseline model. We show that choice data are compatible with our model if and only if data satisfy three simple behavioral postulates. The first one, Starting Point Contraction, says if the starting point is

[1] Nevertheless, it is conceivable that starting points will not be observable in some situations. We also consider cases we only observe the classical choice data in Section 4.

chosen in a set, then it must be chosen from any subset of it as long as it is still starting point. The second one, Replacement, states 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 bigger 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. For example, if our Netflix consumer's perceived network is the same as the Netflix's original massive network, she will end up spending a lot of time to uncover all available options. Except for the product network, online shopping could be seen frictionless due to the amount of extensive information about products. However, many studies show that consumers indeed engage in very limited search (Johnson et al. [2004], Keane et al. [2008], Goeree [2008], Kim et al. [2010]). One might imagine that due to other limitations (such as time pressure, limited cognitive capacity) consumer terminates her search after certain steps. For example, the customer only considers the alternatives directly linked to the starting point. Of course, this model reduces to our baseline model if the number of steps is larger than the number of options. We provide a characterization for K-step network choice for any fixed K and discuss the properties the consideration sets resulting from K-step search satisfy.

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. Nevertheless, we can imagine situations where we do not observe the starting point, but only the standard choice data. For example, if the starting point is what the decision maker expects to buy in the market, it can be hard to elicit such information. Given possible limitations in the data, we would still like to know whether we can identify decision makers who follow our procedure. Section 4 studies network choice with only classical choice data. We also illustrate how to extend our results to revealed preference and network in the baseline model to the case with unobserved starting points.

Section 5 investigates decision-making with a random network. The randomness of a network can arise from two factors: (i) the exogenous product network that we take as given may be random (for example, Amazon's recommendation algorithm may produce random links between alternatives), (ii) the decision maker may randomly pay attention to presented options. A general random network choice could be very complex and intractable. We consider a particular case which we believe is more realistic and tractable than others. Here, given the realization of a random network, the decision maker only considers alternatives which are linked to the starting point. Since the links between options are random the decision maker's consideration set will be random. As we do not observe the realized consideration set, the choices will appear as random. However, this is different than the random utility models where randomness comes from utility rather than product network. We provide a characterization result and illustrate how to infer preference under random network.

Theoretical work on decision making within product networks is limited. The closest paper we know of is Masatlioglu and Nakajima [2013]. They provide a framework to study behavioral search by utilizing the idea of consideration sets. Their baseline model is quite general, which 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. We show that these models are behaviorally distinct (see Section 6).

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

2 Model

Let X be a finite set of alternatives. An extended choice problem is (S, x) where S is a nonempty subset of X and x is a starting point in S. Intuitively, if the choice problem is (S, x), then the set of available alternatives is S and the consumer starts searching from an alternative x in S. A choice function c assigns a single element to each extended choice problem (S, x).

Let $\mathcal{N}$ stand for the product network. If $(x, y) \in \mathcal{N}$, we say that there is a link from x to y. Intuitively, if there is a link from x to y, then the agent who is considering alternative x must also consider alternative y. We let $(x, x) \in \mathcal{N}$ for all $x \in X$. Given $\mathcal{N}$ we can define a link function $\gamma : X \times X \rightarrow \{0, 1\}$ where $\gamma(x, y) = 1$ if and only if $(x, y) \in \mathcal{N}$.[2] We assume that γ is symmetric, that is $\gamma(x, y) = \gamma(y, x)$ for all $x, y \in X$. In other words, the links between alternatives are undirected. If there is a link from x to y, then there must also exist a link from y to x.

We do not specify how the product network is formed. In other words, the analyst has no prior knowledge about the product network and tries to pin it down from the observed choice behavior. This assumption is based on the fact that each agent potentially faces a different product network. For example, while some of them utilize Amazon, others might use Barnes & Noble for buying books. Furthermore, it is well known that many websites use algorithms that produce different recommendations based on consumer's characteristics.[3] Even if the exogenous network is the same for everyone (e.g., everyone uses Amazon and gets the same recommendations) it is still possible that each agent perceives the network differently. Due to all of these observations, we assume that the product network is unobservable and try to reveal it from the choices.

Suppose the agent 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 the agent 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 agent does there may be certain alternatives that she will never discover. In this case, given γ, the agent's consideration set will be given by

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

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. In this case, we also say that there is a path from x to y. We can think of $N_x(S)$ as endogenously determined consideration set of an agent faced with the choice problem

[2]This notation is useful when we consider probabilistic networks.

[3]See Linden et al. [2003] for descriptions of some common recommendation algorithms used in the industry.

(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), $c(S, x)$ is $\succ$-best element in $N_x(S)$.

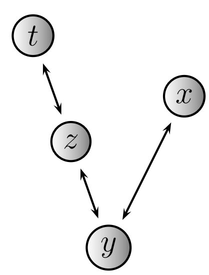
Definition 1. *A choice function c is a **network choice** if there exists a strict preference relation $\succ$ on X and a symmetric link function γ such that*

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

where $N_x(S)$ is defined as above.

Notice that if we have a complete product network, that is $\gamma(x, y) = 1$ for all $x, y \in X$, then $N_x(S) = S$ for all S and we are back to the standard model. In other words, the standard model is a special case network choice. When we have an incomplete product network, we can see more interesting choice behaviors. The following example illustrates network choice and the type of choice behavior it allows.

Example 1. *Efe faces a particular product network. The network is described in the following figure. His preference is $t \succ x \succ z \succ y$. Notice that the network is illustrated in such a way that the height of an alternative reflects his preference.*



Then, we can derive Efe's choices.[4]

	$xyzt$	xyz	xyt	xzt	yzt	xy	xz	xt	yz	yt	zt
x	t	x	x	x		x	x	x			
y	t	x	x		t	x			z	y	
z	t	x		t	t		z		z		t
t	t		t	t	t			t		t	t

[4]For brevity, we represent each decision problem with the available alternatives. For example, $xyzt$ represents $\{x, y, z, t\}$.

In this example, for some choice problems, Efe makes the same choices no matter what the starting point is (e.g., the column of $\{x, y, z, t\}$). This means that these choice problems are connected (i.e., there is a path between any two alternatives in the set), hence starting points have no influence on the choice as in the standard theory. On the other hand, for example, when the choice set is $\{x, z, t\}$ (look at its column), Efe chooses x when the starting point is x and he chooses t otherwise (*starting point effect*).

Another interesting pattern is choice reversal for a fixed starting point (now look at across a row). Consider the choices when the starting point is x. Efe chooses t when the choice set is $\{x, y, z, t\}$, but he picks x when the choice problem is $\{x, y, t\}$ or $\{x, z, t\}$ (*network effect*). That is because when we remove y or z from the choice problem even though t is still among the available alternatives, it is no longer reachable from x.

2.1 Characterization

Before moving on to characterization, we want to investigate the properties of consideration sets ($N_x(S)$) in this model. Let $\mathcal{P}_{\geq 1}(X)$ denote all nonempty subsets of X. We say that $\Gamma_x : \mathcal{P}_{\geq 1}(X) \to \mathcal{P}_{\geq 1}(X)$ is a consideration set mapping if for any $x \in S \subseteq X$, $x \in \Gamma_x(S) \subseteq S$. Let $\{\Gamma_x\}_{x \in X}$ be a collection of consideration set mappings. Now the question is that what type of properties on $\{\Gamma_x\}_{x \in X}$ guarantees that they are induced by a search over a product network.

Lemma 1 provides an answer for that question. The first property says that if an alternative is considered when the choice problem is (T, x), it must also be considered when the choice problem (S, x) contains T. Intuitively, all the alternatives that are reachable in T in her perceived network must be reachable when the choice problem is bigger. Of course, it is possible that the decision maker considers more since there are more alternatives which are potentially reachable from the starting point.

The second property says that if y is considered when the choice problem is (S, x), then the consideration sets corresponding to choice problems (S, x) and (S, y) must be the same. To see why it is true, note that 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 is considered when the choice problem is (S, x), but it is not considered when we remove y from the choice problem, then y must be reachable from x without z. It also imposes that z must be reachable from y in the absence of x. The "if part" of the statement implies that y must be on the path from x to z in the product network. Hence, both x and z must be reachable from y.

We now formally state these properties and provide the result.

A.1 (Upward Monotonicity) If $y \in \Gamma_x(T)$, then $y \in \Gamma_x(S)$ for all $S \supseteq T$.

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

A.3 (Path Connectedness) If $z \in \Gamma_x(S)$, $z \notin \Gamma_x(S \setminus y)$, then $y \in \Gamma_x(S \setminus z)$ and $z \in \Gamma_y(S \setminus x)$.[5]

Lemma 1. *$\{\Gamma_x\}_{x \in X}$ satisfies Upward Monotonicity, Symmetry, and Path Connectedness if and only if there exists a symmetric link function γ such that $\Gamma_x(S) = N_x(S)$ for all $x \in S \subseteq X$.*[6]

Notice that these properties are defined on consideration sets, which are usually unobservable. If we have some information about consideration sets, Lemma 1 will be handy. Then by verifying these properties one can claim that the agent faces network choice. As an example, Reutskaja et al. [2011] find that the number of alternatives considered increase as the choice set gets bigger. This suggests evidence for Upward Monotonicity. In addition to that, Lemma 1 also allow us to compare our model with existing limiting consideration models such as Manzini and Mariotti [2007, 2012], Masatlioglu et al. [2012], Cherepanov et al. [2013] (limited attention), Rubinstein and Salant [2006] (choice from lists), Manzini and Mariotti [2014] (random attention), Caplin and Dean [2011] (search).

As we discussed before, it is unlikely that we will have such information on implied consideration sets. We now propose three simple axioms on observed choice behavior (not on consideration sets). We then discuss how each axiom is related to three properties on consideration sets we discussed above.

The first axiom is similar to standard contraction axiom for the starting point. It says that if the starting point is chosen in some set S, then it must also be chosen in any subset T of S as long as it is the starting point. In the standard model, this would be true of any alternative in T that is chosen when the choice problem is (S, x). But

[5] x, y, and z are distinct alternatives.

[6] All proofs are provided in the Appendix.

in the network model, it is no longer true for all alternatives since an alternative that is considered in a bigger choice set is not necessarily considered in a smaller choice set unless it is the starting point.

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

Note that Axiom 1 is directly implied by the first property of consideration sets. 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 paths decreases, and hence the implied consideration set can only get smaller by A.1. Since the starting point is always available and it was chosen when the choice set was bigger, 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 make a difference in the bigger 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, first 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 by A.1, $y \in \Gamma_x(S)$ and by A.2, $\Gamma_x(S) = \Gamma_y(S)$. Now suppose y is chosen when the choice problem is (T, x) which implies that y is considered when the choice problem is (T, x). By strong symmetry property, for any $S \supseteq T$, the consideration sets corresponding to choice problems (S, x) and (S, y) are the same. Therefore, the agent 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* $c(S \setminus z, x) = c(S \setminus z, y)$.

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 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 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 network choice.

Theorem 1. *A choice function c satisfies Starting Point Contraction, Replacement, and Choice Reversal if and only if it is network choice.*[7]

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 the observed choice data.

2.2 Revealed Preference and Revealed Network

In this section, we discuss how we can reveal preference and network from the choice data given that the consumer follows the network choice. The standard theory suggests 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 the consumer's preference is no longer trivial. First, we illustrate that there may be multiple preferences representing given choice behavior.

[7]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.

Example 2. *Altan, our decision maker, always chooses x independent of the starting point as long as it is available. However, in the absence of x, his choice is dictated by the starting point. His behavior is summarized below.*

	xyz	xy	xz	yz
x	x	x	x	
y	x	x		y
z	x		x	z

First, we need to make sure that Altan's behavior can be written as a network choice before we discuss revealed preference. To check that we utilize Theorem 1. It is routine to verify that Altan's behavior satisfies Axiom 1-3. Therefore by Theorem 1 we are guaranteed that a network choice representation exists. After establishing this fact, we first discuss the revealed preference and the revealed network structure.

Firstly, the fact that the decision maker chooses x from the choice problem $(\{x, y\}, y)$ implies that there is a link from y to x. Similarly, we also know that there is a link between x and z. Since the decision maker chooses different alternatives from the choice problems $(\{y, z\}, y)$ and $(\{y, z\}, z)$, there is no link between y and z (notice there is a path between them through x). Hence, the network is fully revealed. It turns out that is a general feature of the model. To reveal network, it suffices to look at binary choice data with different starting points. If the decision maker's choices from two choice problems $(\{x, y\}, x)$ and $(\{x, y\}, y)$ are the same, then we reveal that x is linked to y. Otherwise, we reveal that x is not linked to y.

While the entire network is revealed uniquely, the preference is not unique in this example. Indeed, there are two possible preference relations that can represent the choice behavior: $x \succ_1 y \succ_1 z$ and $x \succ_2 z \succ_2 y$. Given that there can be multiple preferences representing the choice, we need to define what we mean by revealed preference. Following Masatlioglu et al. [2012] we say that x is revealed preferred to y if x is ranked higher than y in all possible representations.

Definition 2. *Suppose c is a network choice and let $\{(\succ_i, \gamma) |\, i = 1, \ldots, N\}$ 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 $i = 1, \ldots, N$.*

Given the revealed network, we can compare two alternatives x and y only if there is a link or path from x to y. If there is no link or path between these two alternatives, there will be no choice problem in which these alternatives are considered at the same time implying that we cannot tell which alternative 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 y and z, but we can still not tell which alternative is more preferred. We can only reveal preference between two alternatives only if (i) there is a choice set in which a link or path between them exists and (ii) one alternative is chosen over the other.

Consider an observation that the decision maker chooses x when y is the starting point. This observation satisfies both conditions above: (i) there is a path between x and y, and (ii) x is chosen over y. The following proposition summarizes the results.

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

- *x is revealed preferred to y if and only if there exists S such that $c(S, y) = x$,*
- *x is revealed to be linked to y if and only if $c(\{x, y\}, x) = c(\{x, y\}, y)$.*[8]

Proposition 1 provides a necessary and sufficient condition for revealed preference in our model. This result implicitly assumes that the analyst observes the entire choice behavior. We will illustrate that Proposition 1 may not be useful for revealing preference with limited data. Let assume that we only observe $x = c(S \cup y, z)$ and $y = c(\{y, z\}, z)$ nothing else. By Proposition 1, we know that both x and y are revealed preferred to z. However, Proposition 1 is silent about the relative ranking of x and y with this limited data. In our model, one can prove that x is preferred to y by seeking more choice data and/or applying the axioms to fill in the missing choice data. Indeed, we must have a set $T \subset S$ such that $c(T, y) = x$, hence x is preferred to y. Therefore, if we had the entire choice data, Proposition 1 would have informed us about the relative ranking of x and y.

We would like to list more choice patterns which reveal the decision maker's preference. These patterns will inform us about revealed preference when we have partial data, which is more useful for empirical studies.

1. $x = (S, y)$,

[8] All results in this section can be easily verified by the proof of Theorem 1.

2. $x = c(S \cup y, z) \neq c(S, z)$,

3. $x = c(S, z)$ and $y = c(T, z)$. for $T \subset S$.

The first one directly follows from Proposition 1. The second one 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 on the path from z to x. That is the only way it can affect the decision maker's choice. We can conclude that x is preferred to y because we know that y was considered and x was chosen.

In the third choice pattern, $y = c(T, z)$ reveals that y is reachable from z in T. Therefore, y is reachable from z in S since T is a subset of S. In the decision problem (S, z), y is considered and x is chosen. Therefore, x is revealed preferred to y.

Any preference that can represent c must be consistent with the above revelations. Indeed, we might get more revelation due to the transitivity. For example, suppose we learn x is revealed preferred z from the first observation and z is revealed preferred to y from the third choice pattern. Then we must have x is revealed to preferred y. Formally, given a network choice 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 network choice. The next proposition states these facts formally.

Proposition 2. *(Revealed Preference) Suppose c is a network choice. Then, x is revealed to be preferred to y iff $x \gg_c^{TC} y$.*

Proposition 1 shows us how to reveal network uniquely if we have access to binary choice data. If we do not have such data, we may still be interested in whether there exists a path between two alternatives in a given set since it gives us information about consideration sets. Let $x, y \in S$ be given and suppose one of the following is observed.

1. $c(S, x) = y$ or $c(S, y) = x$,

2. $c(S, x) = z \neq c(S \setminus y, x)$ or $c(S, y) = z \neq c(S \setminus x, y)$,

3. $c(S, x) = c(S, y)$.

Then in each case, our model implies that there is a path between x and y in any set containing S. The first case is obvious. For the second one, an irrelevant alternative can cause choice reversal only if it is on the path between the starting point and the chosen alternative. The last one follows from (1) and the fact that $c(S, x) = c(S, y) = z$ for some $z \in S$.

As an example, suppose $c(\{x, y, z, t\}, x) = t \neq c(\{x, z, t\}, x)$ and $c(\{x, y, z\}, y) = z$. Proposition 1 is not useful with these observations. However, the choice reversal pattern tells us that there is a path between x and y in $\{x, y, z\}$. Since z is chosen when the choice problem is $(\{x, y, z\}, y)$ we also learn that there is a path between y and z in $\{x, y, z\}$. Then it must also be true that there is a path between x and z in $\{x, y, z\}$ through y.

Formally, for any $x, y \in T$, define

$$\widehat{\gamma}_T(x, y) = 1 \text{ if there exists } S \subseteq T \text{ such that one of the above is observed}$$

and let $\gamma_T(x, y) = 1$ if there exists $\{x_1, \ldots, x_n\} \subseteq T$ with $x_1 = x$ and $x_n = y$ such that $\widehat{\gamma}_T(x_i, x_{i+1}) = 1$ for $i < n$. The following proposition says that if $\gamma_T(x, y) = 1$, then we are guaranteed that there exists a path between x and y in T.

Proposition 3. *Suppose c is a network choice. If $\gamma_T(x, y) = 1$, then there exists a path between x and y in T.*

3 Limited Search: K-step

In this section, we introduce bounded rationality in our framework. Recall that in the original model the agent considers all the alternatives that appear in her perceived network. In reality, if the product networks are not sparse the search process can take a long time. For example, if our Netflix consumer's perceived network is the same as the Netflix's original huge network, she will end up spending a lot of time to uncover all available options. One might imagine that due to other limitations (such as time pressure, limited cognitive capacity) consumer terminates her search after certain steps. Here we consider an agent who has some search cost or experiences fatigue after a certain number steps in a search process.

We assume that the agent 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 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 for any fixed K and discuss the properties the consideration sets resulting from K-step search satisfy.

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

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

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

Definition 3. *A choice function c is a K-step network choice if there exists a preference $\succ$ over X and a symmetric link function γ such that*

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

where $N_x^K(S)$ is defined as above.

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

Example 3. *Remember Efe from Example 1. Suppose Efe always stops the search after two steps due to his cognitive limitations. His choice data we observe will be as follows.*

	$xyzt$	xyz	xyt	xzt	yzt	xy	xz	xt	yz	yt	zt
x	x	x	x	x		x	x	x			
y	x	x	x		z	x			z	y	
z	t	z		t	t		z		z		t
t	t		t	t	t			t		t	t

One interesting feature of Efe's choice data is that for a fixed starting point, Efe acts "as if" he is classical utility maximizer. For example, if we only have choice data with the starting point y and we do not observe the starting point we will verify that Efe's behavior satisfies WARP. However, our revealed preference relation will be $x \succ z \succ y \succ t$. In other words, we will mistakenly reveal that t is the worst alternative even though it is the best alternative. 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

Before moving on to characterization, we first show the properties the consideration sets arising from K-step network choice satisfy. Lemma 2 shows that if a collection of consideration set mappings $\{\Gamma_x\}_{x\in X}$ satisfies the four properties described below, then we can treat them as a K-step consideration set on a product network.

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. We now state these properties.

B.1 (Upward Monotonicity) If $y \in \Gamma_x(T)$, then $y \in \Gamma_x(S)$ for all $S \supseteq T$.

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

B.3 (Path Connectedness) If $z \in \Gamma_x(S)$, $z \notin \Gamma_x(S \setminus y)$, then $y \in \Gamma_x(S \setminus z)$ and $z \in \Gamma_y(S \setminus x)$.

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

The first property is upward monotonicity which says that as choice sets get bigger, the consideration sets do not get smaller. The idea for this property is the same as in the original model. When there are more alternatives available, there are also more alternatives that can potentially be reached in K steps.

The second property is a modified version of symmetry that we have in the main model. Recall that in the main model, we showed that if y is reachable from x in S, then the consideration sets corresponding to choice problems (S, x) and (S, y) must be the same. This is no longer true in K-step search model since an alternative that is reachable from y in K steps is not necessarily reachable from x in K steps. However, if the choice set has at most K alternatives, then symmetry follows.

The third property is path connectedness that we have in the main model. It says that if z is reachable from x in S, but not in $S \setminus y$, then y must be on the path from x to z. The fact that the search is K-step does not affect this property.

The last property is contraction which says that if y is reachable from x in S, then there must be a subset T of 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 which connect x to y. This set must have at most K alternatives.

Lemma 2. *$\{\Gamma_x\}_{x\in X}$ satisfies Upward Monotonicity, Symmetry, Path Connectedness, and Contraction if and only if there exists a symmetric link function γ such that $\Gamma_x(S) = N_x^K(S)$ for all S.*

Even though Lemma 2 is useful in understanding the properties of consideration sets, since we usually do not have data on consideration sets it does not help us in verifying whether the agent makes a K-step network choice. 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 agent picks x^*. Then, if we extend the choice set and make the choice problem (T, z), the agent must still consider x^*. Since x^* is the best element in S, if the agent 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 only talks about the existence of such an alternative. If we find some choice set such that no alternative satisfying Axiom 4 exists, then we can conclude that the agent 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. *(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 agent chooses y when the choice problem is (T, x). Then y is reachable from x in T in K steps. By Upward Monotonicity and Symmetry properties of consideration sets, for any choice set S, that contains T and has at most K alternatives, the consideration sets corresponding to choice problems (S, x) and (S, y) must be the same. Therefore, the agent 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 get choice reversal axiom holds if the choice set does not have more than K alternatives.

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

To see why the original version is no longer true, suppose the 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 in K steps, but not 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, there may be some alternative $t \in S \setminus x$ that is more preferred to z and is reachable from y in $S \setminus x$ in K steps, but not from x in S. Therefore, it is not necessarily the case that $c(S \setminus x, y) = z$. By a similar argument, $c(S \setminus z, x) = c(S \setminus z, y)$ is not necessarily true either. If the choice set S has at most K alternatives, then we are back to the main model, and the choice reversal axiom holds.

The last axiom is contraction which says that if y is chosen when the choice problem is (S, x), then there must exist a subset T of S with at most alternatives K such that y is chosen when the choice problem is (T, x).

Axiom 7. *(Contraction) 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. $c(S, x) = y$ implies that y is preferred to all the other alternatives in T. Therefore, 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, Replacement, Choice Reversal, and Contraction if and only if it is a K-step network choice.*

Revealed Preference and Network

In this section, we discuss how one can reveal preference and network given that the agent 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 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's choices from Example 3. We can partially reveal his preference as follows. Firstly, $c(\{x,y\},y) = x$ and $c(\{y,z\},y) = z$ implies that $x \succ y$ and $z \succ y$. That is because in both cases the starting point y can be abandoned only if the chosen alternative is better than the starting point. Furthermore, $c(\{z,t\},z) = t$ implies $t \succ z$. The choice data with starting points x and t are not useful in revealing his preference since in both cases the starting points are always chosen. However, we can still reveal preference between x and z. Notice that $c(\{y,z\},y) = z$ implies that z is linked to y. Therefore, it must also be considered when the choice problem is $(\{x,y,z\},y)$. Since $c(\{x,y,z\},y) = x$ we conclude that $x \succ z$. There are two possible preferences that can explain the choice data: $x \succ_1 t \succ_1 z \succ_1 y$ and $t \succ_2 x \succ_2 z \succ_2 y$. The preference between x and t is not identified.

In general, 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 4 says that if x is revealed to be preferred to y, then we must also have xP_Ry. To see why this is true, suppose x is revealed preferred to y. Then, there must be a choice problem such that x is chosen while y is being considered. In other words, there exist $z \in S$ and $T \subseteq S$ such that $c(S,z) = x$ and $c(T,z) = c(T,y)$. The second condition guarantees that y is considered in the choice problem (S,z). Now if $c(T,z) = y$, then we have xPy. If $c(T,z) = t = c(T,y)$, then we have xPt and tPy and hence xP_Ry.

Proposition 4. *Suppose c is a K-step network choice. Then*

- *x is revealed to be preferred to y if and only if xP_Ry,*

- *x is revealed to be linked to y if and only if* $c(\{x, y\}, x) = c(\{x, y\}, y)$.[9]

4 Unobserved Starting Points

In the previous sections, we assume that we can observe the starting point of the agent. 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 agent 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 agent always chooses the starting point. But then any choice behavior is possible under this model. Therefore, 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. Firstly, 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.

In the second model, we assume that the agent's starting point is random. Hence, the induced choice is probabilistic even though the network is deterministic. Again, an outside observer, who views the probabilistic choices of the decision maker, knows that the starting points are stochastic and but does not have information on the actual starting point. Under this assumption, our model becomes is a generalization of Luce model (Luce [1959]) and the standard model. If all the alternatives in the product network are isolated (i.e., no links between any two alternatives exist), then the choice probabilities correspond to Luce probabilities. If the product network is complete

[9]The proof of Proposition 4 directly follows the proof of Theorem 2.

(i.e., any two alternatives in the product network are linked), then the randomness in starting point has no effect and the decision maker always chooses the best available alternative, which reduces the model to the standard model.

4.1 Choice Correspondence

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. Formally,

Definition 4. *A choice correspondence C is an induced network choice if there exists a preference relation $\succ$ over X and a symmetric link function γ 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).

The following example illustrates the properties of induced network choice.

Example 4. *Remember Efe from Example 1. Let's assume that we do not observe the starting point. Instead we observe his induced network choice. That is,*

S	$xyzt$	xyz	xyt	xzt	yzt	xy	xz	xt	yz	yt	zt
$C(S)$	t	x	$\{x,t\}$	$\{x,t\}$	t	x	$\{x,z\}$	$\{x,t\}$	z	$\{y,t\}$	t

For example, $C(\{x, y, z, t\}) = t$ implies that no matter what the starting point is in $\{x, y, z, t\}$ Efe always picks the alternative t. Notice that for any choice set, the choices we observe is just the vertical summation of the choices in Example 1. In other words, in both choice data we observe all possible choices, but here we do not observe the relation between starting points and choices.

One interesting thing about this choice behavior is that even though x is not dominated in any of the binary comparisons, x is not chosen when the choice set is $\{x, y, z, t\}$. In addition, even though t is uniquely chosen when the choice set is $\{x, y, z, t\}$ removal of y or z changes the choice behavior. These are the types of behavior that are not allowed in the standard model. Notice that for any choice set a unique alternative will be chosen if and only if the choice set is connected. In this example, $\{x, y, z, t\}$, $\{x, y, z\}$, and $\{y, z, t\}$ are connected sets with more than 2 alternatives.

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. But 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 $x \in T \subseteq S$ 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) If $C(T) = x$ and $C(S) = y$ for $T \cap S \neq \emptyset$, 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 two 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 says 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.*

Revealed Preference and Network

Given our results for revealed preference and network in the main model, it is easy to extend it to the case with unobserved starting points. In the model with observed starting points, we found that x is revealed to be preferred to y if and only if there exists S containing x and y such that $c(S, y) = x$. If there exists such S, then let T be a set containing x, y, and the alternatives on the path from x to y in S. Since T is a connected subset of S it has to be the case that $c(T, z) = x$ for all $z \in S$. Given this observation, we reveal that x is preferred to y if and only if there exists some choice set containing x and y such that x is chosen no matter what the starting point is. In fact, not observing starting points does not lead to any losses in preference revelations.

Similarly, in the model with observed starting point we reveal that x is linked y if and only if $c(\{x, y\}, x) = c(\{x, y\}, y)$. With unobserved starting point this corresponds to the case when $C(\{x, y\})$ is a singleton. If $c(\{x, y\}, x) \neq c(\{x, y\}, y)$ or $C(\{x, y\})$ is not a singleton, we reveal that x is not linked to y. The product network is fully revealed in the model with unobserved starting point as is the case in the model with observed starting point.

The following corollary summarizes the results.

Corollary 1. *Suppose C is a network choice. Then,*

- *x is revealed to be preferred to y if and only if there exists $S \supseteq \{x, y\}$ such that $C(S) = x$,*
- *x is revealed to be directly linked to y if and only if $C(S)$ is a singleton.*[10]

[10] The proof of Corollary 1 directly follows the proof of Theorem 3.

4.2 Probabilistic Starting Point

We now assume that the starting point is determined probabilistically. That is, each alternative could be the starting point with some positive probability. The induced choice, then, is probabilistic even though the network is deterministic. There are many reasons why starting point can be random. For example, Amazon advertises different alternatives on different days on its front page. We can get different recommendations from different people we encounter. It is also possible that when we explore a certain category the first alternative that we are reminded of depends on the environment we are in.

To model this we assume there is a probability distribution β over X. $\beta(x)$ is the probability of x being the starting point when the choice set is X. We assume that every element in X has a positive probability of being a starting point: $\beta(x) > 0$ for all $x \in X$. Then given a choice set S, the probability that x is a starting point in S is given by $\frac{\beta(x)}{\sum_{y\in S}\beta(y)} = \frac{\beta(x)}{\beta(S)}$.

Let $p(x, S)$ denote the probability of x being selected when the choice set is S. We say that p is a network choice if $p(x, S)$ is the sum of probabilities of different starting points for which x is the final choice.

Definition 5. *A probabilistic choice p is a network choice if there exists a preference relation $\succ$ over X, a symmetric link function γ, and a probability distribution β over X such that*

$$p(x, S) = \sum_{\substack{y\in S \\ x=\text{argmax}(\succ, N_y(S))}} \frac{\beta(y)}{\beta(S)}$$

where $\beta(S) = \sum_{y\in S}\beta(y)$ and $N_y(S)$ is defined as before.

There are two special cases of this model.

1. If every alternative is in isolation in the network, (that is $N_x(S) = x$), then the model reduces to Luce rule.

$$p(x, S) = \frac{\beta(x)}{\sum_{y\in S}\beta(y)}$$

2. If the network is complete, (that is, $N_x(S) = S$ for all x and S,) then the model

reduces to standard model. Hence there is no randomness in the choice.

$$p(x,S)=\begin{cases}1 & \text{if } x=\text{argmax}(\succ,S)\\ 0 & \text{if } x\neq\text{argmax}(\succ,S)\end{cases}$$

Recall that if $x \in N_y(S)$, then by A.2, $N_x(S) = N_y(S)$. This implies an alternative represenation for network choice given by

$$p(x,S)=\begin{cases}\frac{\beta(N_x(S))}{\beta(S)} & \text{if } x=\text{argmax}(\succ,N_x(S))\\ 0 & \text{if } x\neq\text{argmax}(\succ,N_x(S))\end{cases}$$

In other words, if x is the best element in it's neighborhood, then the probability of it being chosen is given by the measure of its neighborhood divided by the measure of the choice set. If x is not the best element in it's neighborhod, then it is never chosen.

Example 5. *Let's revisit Efe's choices. We assume that each alternative could be his starting point with equal chances ($\beta(x) = \beta(y) = \beta(z) = \beta(t) = 1/4$). Then his choice data we observe will be as follows.*

$p(\cdot,S)$	$xyzt$	xyz	xyt	xzt	yzt	xy	xz	xt	yz	yt	zt
x	*0*	*1*	*2/3*	*1/3*	-	*1*	*1/2*	*1/2*	-	-	-
y	*0*	*0*	*0*	-	*0*	*0*	-	-	*0*	*1/2*	-
z	*0*	*0*	-	*0*	*0*	-	*1/2*	-	*1*	-	*0*
t	*1*	-	*1/3*	*2/3*	*1*	-	-	*1/2*	-	*1/2*	*1*

Notice that if the choice set is connected, then the best alternative is chosen with probability 1. For example, t is chosen with probability 1 when the choice set is $\{x, y, z, t\}$. If the choice set is not connected, we can have multiple alternatives being chosen with positive probability. For example, when the choice set is $\{x, y, t\}$, x is chosen with 2/3 probability and t is chosen with 1/3 probability. The fact that x is chosen more often does not reflect that it is more preferred than t. In fact, here t is ranked higher than x in the preference ordering. The reason why x is chosen more often is that it is the best alternative in the neighborhood of x and y in $\{x, y, t\}$, whereas t is the best alternative only in the neighborhood of t in $\{x, y, t\}$. In addition, notice that the probability ratio of two alternatives depends on the choice set.

Characterization

First, notice that probabilistic choice is just an extension of the induced choice correspondence model with the addition that we know exact choice probabilities. While, in the induced model, we only know the support of the distribution, here the entire distribution is given to us. Hence the four axioms characterizing the model with choice correspondence must still hold. We state those axioms in the new framework below.

Axiom 12. *(Contraction) If $p(x, S) > 0$, then $p(x, T) > 0$ for $x \in T \subseteq S$.*

Axiom 13. *(Dominating Alternative) If $p(x, S) = 0$, then there exist $y \in S$ with $p(y, S) > 0$ and $T \subseteq S$ containing x such that $p(y, T) = 1$*

Axiom 14. *(Expansion) If $p(x, S) = 1$ and $p(y, T) = 1$ where $S \cap T \neq \emptyset$, then either $p(x, S \cup T) = 1$ or $p(y, S \cup T) = 1$.*

Axiom 15. *(Separability) Suppose $|S| \geq 3$ and $p(x, S) = 1$. Then, there exist non-singleton $T_1, T_2 \subset S$ with $T_1 \cup T_2 = S$ and $T_1 \cap T_2 \neq \emptyset$ such that $p(x, T_1) = 1$ and $p(y, T_2) = 1$ for some $y \in S$.*

In addition to these four axioms, we need axioms on choice probabilities. Recall that in Luce model we have menu independence property which says that choice probability ratio of two alternatives should not depend on the choice set. Since our model is an extension of Luce model this property is no longer satisfied in general. However, under certain conditions given by Axiom 16 we can have menu independence. In particular, Axiom 16 says that the probability ratio of two alternatives x and y in two sets S and $S \cup T$ will be the same if T is a connected set such that no alternative in T is linked to alternatives in connected sets containing x or y.

Axiom 16. *(Weak Menu Independence) Suppose $p(x, S \cup T) > 0$ and $p(y, S \cup T) > 0$ where $\{x, y\} \subseteq S$. If there exists $z \in T$ with $p(z, S \cup T) > 0$ and $p(z, T) = 1$, then*

$$\frac{p(x, S \cup T)}{p(y, S \cup T)} = \frac{p(x, S)}{p(y, S)}$$

To see why this axiom holds, suppose x and y are chosen with positive probability when the choice set is S. This can only happen if S has a connected subset S_1 in which x is the best alternative and a connected subset S_2 in which y is the best alternative. Suppose we extend the choice set by $S \cup T$, and x and y are still chosen with positive probability. If T has no connections with S_1 and S_2, then we expect that the probability ratio of x and y remains the same. That is because the measure of the neighborhood of x and y stays the same in the bigger choice set. The fact that there is an alternative $z \in T$ with $p(z, S \cup T) > 0$ and $p(z, T) = 1$ guarantees that T has no connections with S_1 and S_2.

Axiom 17 is additivity property which allows us to write the ratio of choice probabilities of x and y in some choice set as a sum of choice probability ratios. It says that if S is a connected set in which x is the best alternative and T is a connected set in which y is the best alternative, then the probability ratio of x and y in $S \cup T$ is the sum of probability ratio of z and y in $T \cup z$ for $z \in S$.

Axiom 17. *(Additivity) Suppose $p(x, S) = 1$ and $p(y, T) = 1$. If $p(x, S \cup T) \in (0, 1)$, then*

$$\frac{p(x, S \cup T)}{p(y, S \cup T)} = \sum_{z \in S} \frac{p(z, T \cup z)}{p(y, T \cup z)}$$

To see why it holds, suppose we have two connected sets S and T. Let x be the best alternative in S and y be the best alternative in T and suppose x and y are chosen with positive probability in $S \cup T$. This can only happen if there are no links between the alternatives in S and the alternatives in T. Then, the probability ratio of x and y is just equal to the ratio of the measure of the neighborhood of x (the measure of S) and the measure of the neighborhood of y (the measure of T). Finally, note that the measure of S is just the sum of the measure of each alternative in S.

Theorem 4 says that Axioms 12-17 are necessary and sufficient to characterize probabilistic network choice.

Theorem 4. *Probabilistic choice p satisfies Contraction, Dominating Alternative, Expansion, Separability, Weak Menu Independence, and Additivity if and only if it is a network choice.*

Revealed Preference and Network

The intution behind preference and network revelation is exactly similar to the intuiton in the basic model. We showed that x is revealed to be linked to y if and only if $c(\{x, y\}, x) = c(\{x, y\}, y)$. In probabilstic choice this corresponds to $p(x, \{x, y\}) = 0$

or 1. Similarly, x is revealed not to be linked to y if and only if $p(x, \{x, y\}) \in (0, 1)$. As before the product network is fully revealed.

In the main model, we showed that x is revealed to be preferred to y if and only if there exists S containing x and y such that x is chosen no matter what the starting point is. In probabilistic choice this corresponds to having a set S containing x and y such that $p(x, S) = 1$. The revealed preference is exactly the same as the revealed preference in the previous models.

The results are summarized in Corollary 2.

Corollary 2. *Suppose p is probabilistic network choice. Then,*

- *x is revealed to be preferred to y if and only if there is $S \supseteq \{x, y\}$ such that $p(x, S) = 1$,*
- *x is revealed to be directly linked to y if and only if $p(x, \{x, y\}) = 0$ or 1.*[11]

5 Random Network

So far we only discuss the deterministic networks. In this section, we investigate network choice with a random network. The randomness of a network can arise from two factors: (i) the exogenous product network that we take as given may be random (for example, Amazon's recommendation algorithm may produce random links between alternatives), (ii) the decision maker may pay random attention to presented alternatives. A general network choice with a random network can be very complex and intractable. In this section, we consider a particular case which we believe is more realistic and tractable than others. Here, given the realization of a random network, the decision maker only considers alternatives which are linked to the starting point. Since the links between alternatives are random, the decision maker's consideration set will be random. As we do not observe the realized consideration set, the choices the agent makes will appear as random. Again, this is different than the random utility models where randomness comes from utility rather than product network.

We assume that the decision maker stops searching after 2 steps, i.e., the decision maker only considers the alternatives that are linked to the starting point. To define

[11] The proof of Corollary 2 directly follows the proof of Theorem 4.

random networks, let $\gamma(x, y)$ denote the probability that there is a link between x and y. We assume that $\gamma(x, x) = 1$ and $\gamma(x, y) \in (0, 1)$ for $y \neq x$. We also assume that links are independent. Let $\mathcal{A}_x(D, S)$ denote the probability that $x \in D \subseteq S$ is the consideration set when the choice problem is (S, x). Since the search is 2-step the consideration set probabilities are given by

$$\mathcal{A}_x(D, S) = \prod_{y \in D} \gamma(x, y) \prod_{z \in S \setminus D} (1 - \gamma(x, z))$$

In other words, the decision maker's consideration set is D if all the alternatives in D are linked to x and none of the alternatives in $S \setminus D$ are linked to x.

Let $p_x(y, S)$ denote the probability of y being selected when the choice problem is (S, x). We denote a probabilistic network choice by P where $P(y, (S, x)) = p_x(y, S)$. We say that P is a probabilistic 2-step network choice if, for all (S, x), $p_x(y, S)$ is the event that a consideration set in which y is the best element is realized.

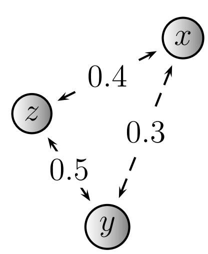
Definition 6. *P is a probabilistic 2-step network choice if there exist a preference relation $\succ$ over X and a symmetric link function γ satisfying $\gamma(x, x) = 1$ and $\gamma(x, y) \in (0, 1)$ for $y \neq x$ such that*

$$p_x(y, S) = \sum_{y \text{ is } \succ\text{-best in } D} \mathcal{A}_x(D, S)$$

where $\mathcal{A}_x(D, S)$ is defined as above.

The following example illustrates probabilistic 2-step network choice.

Example 6. *Suppose Mehmet's preference among three alternatives is $x \succ z \succ y$, and his probabilistic product network is as below.*



We now calculate $p_x(\cdot, \{x, y, z\})$. Given that x is the best alternative and it is the starting point, x will be chosen with probability 1. Hence we have $p_x(x, \{x, y, z\}) = 1$

and $p_x(y, \{x, y, z\}) = p_x(z, \{x, y, z\}) = 0$. It may seem that the magnitude of choice probabilities reflects preference. This is not true. For example, when the choice set is $\{x, y, z\}$ and the starting point is y, x is chosen with 0.3 probability, but y and z are chosen with 0.35 probability even though they are inferior to x. This is due to randomness in the network.

Characterization

There are three axioms which characterize probabilistic 2-step network choice. Axiom 18 is the standard Starting Point Contraction axiom in the probabilistic domain. It says that as the choice set gets bigger, the probability that the starting point is chosen can only get smaller.

Axiom 18. *(Starting Point Contraction) For any S and $x \in T \subseteq S$, $p_x(x, T) \geq p_x(x, S) > 0$.*

Axiom 18 holds in our model because as the choice set gets bigger, there are potentially more preferred alternatives which are realized to be linked to the starting point.

The second axiom says that for any choice set, there must exist a unique dominant alternative, i.e., the alternative which is never abandoned for other alternatives if it is the starting point.

Axiom 19. *(Dominant Alternative) For any S, there exists unique $x \in S$ such that $p_x(x, S) = 1$.*

To see why Axiom 19 is true in our model, given any choice set S pick the most preferred alternative in S. This alternative must be chosen with probability one if it is the starting point. For uniqueness, note that any alternative that is not ranked the highest in the preference ordering has a positive probability of being linked to the most preferred alternative which implies that it cannot be chosen with probability one when it is the starting point.

The last axiom tells us what happens if the dominant alternative is removed from the choice set. Suppose we are given a choice problem (S, x) in which z is the dominant alternative. The axiom tells that the probability that y is chosen in the choice problem (S, x) is equal to the probability that y is chosen in the choice problem $(S \setminus z, x)$ times the probability that x is chosen in the choice problem $(\{x, z\}, x)$.

Axiom 20. *(Removal of Dominant Alternative) Suppose* $p_z(z,S)=1$*. Then,*

$$p_x(y,S)=p_x(y,S\setminus z)p_x(x,\{x,z\})$$

Intuitively, for y to be chosen in the choice problem (S,x) it must dominate all the other alternatives in the choice problem $(S\setminus z,x)$ *and* x must dominate z in the choice problem $(\{x,z\},x)$. Since the links are independent this gives us multiplicative formula.

Theorem 5 gives a characterization of probabilistic 2-step network choice.

Theorem 5. *P satisfies Starting Point Contraction, Dominant Alternative, and RDA if and only if it is a probabilistic 2-step network choice.*

Revealed Preference and Network

Since the probability that any two distinct alternatives are linked is strictly positive, it suffices to check binary choice sets to reveal preference and network. We reveal that x is preferred to y if and only if x is chosen with positive probability in the choice problem $(\{x,y\},y)$, or alternatively, if and only if x is chosen with probability one in the choice problem $(\{x,y\},x)$. Since for any choice set, there exists a unique dominant alternative the two statements are equivalent. Note that this also implies complete and transitive preference relation.

To reveal network we need to check the probability that x is chosen in the choice problem $(\{x,y\},y)$ and the probability that y is chosen in the choice problem $(\{x,y\},x)$. By Dominant Alternative axiom, one and only one of these two probabilities will be equal to zero. The maximum of these probabilities is equal to the probability that x is linked to y.

The following corollary summarizes the results.

Corollary 3. *Suppose P is a probabilisitic 2-step network choice. Then,*

- *x is revealed to be preferred to y if and only if* $p_y(x,\{x,y\})>0$,
- $\gamma(x,y)=\max\{p_x(y,\{x,y\}),p_y(x,\{x,y\})\}$.[12]

[12] The proof of Corollary 3 directly follows the proof of Theorem 5.

6 Related Literature

Behavioral Search

The closest paper we know of is Masatlioglu and Nakajima [2013]. They provide a framework to study behavioral search by utilizing the idea of consideration sets, which evolve dynamically. Their baseline model is quite different than ours. First, their motivation is entirely different, which requires a much more general model. Second, their model is dynamic because interim decisions about what items to pay attention to affect the evolution of the consideration set.

They also study a special case, which is called Markovian consideration set. This special case can be represented as a network. However, their dynamic feature makes this model distinct from ours. Unlike our model where the consumer "clicks on" all linked alternatives, in their model the agent only "clicks on" the best linked alternative. Hence, the consideration sets are preferences dependent. Therefore, each model follows a different procedure to reach a final decision.

Given a fixed preference relation, the consideration sets of the Markovian model in Masatlioglu and Nakajima [2013] violate Upward Monotonicity and Symmetry properties, which is implied by our model (both the baseline model and K-step version). While this established the difference in terms of consideration sets, the models could still generate similar choice behavior. We show that they are also distinct in terms of observed choice behaviors. We now provide an example of choice behavior that can be explained by the Network model but not the Markovian model and vice versa.

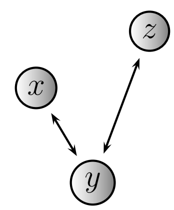
Example 7. *Choice behavior that can be explained by the Network model but not the Markovian model.*

	xyz	xy	xz	yz
x	z	x	x	
y	z	x		z
z	z		z	z

This can be explained by the Network model as follows.

But this choice behavior cannot be explained by the Markovian model since in the Markovian model $c(\{x,z\},x) = x$ *and* $c(\{x,y\},x) = x$ *implies* $c(\{x,y,z\},x) = x$.

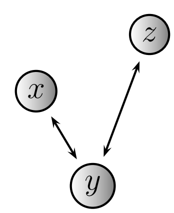
Example 8. *Choice behavior that can be explained by the Markovian model but not*



the Network model.

	xyz	xy	xz	yz
x	x	x	x	
y	z	x		z
z	z		z	z

This choice data can be explained by the Markovian Model as follows.



But this choice behavior cannot be explained by the Network model since Replacement Axiom is violated as $c(\{x, y\}, y) = x$*, but* $c(\{x, y, z\}, x) \neq c(\{x, y, z\}, y)$*.*

In general, it is difficult to say which model approximates consumer behavior better because they utilize different choice procedures. We believe depending on the task one model might fit the consumer behavior more accurately than the other. We hope that future empirical and/or experimental work will shed light on this issue.

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.

Reference Dependent Choice

The network choice is also related to the literature on reference-dependent preferences, especially if we interpret the starting point as the reference point. Even under this interpretation, our model differs from previous models of reference-dependent choices such as Tversky and Kahneman [1991], Masatlioglu and Ok [2005, 2013], Dean et al. [2014]. Contrary to ours, in TK, the reference point affects the utility of individual. In terms of choice behavior, 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 network choice, the starting point is abandoned only when there is a welfare improvement, so their model is not a special case of the network choice. For a fixed starting point, while our model allows choice reversals, their model satisfies WARP. Therefore, these two models are independent.

Masatlioglu and Ok [2005, 2013] propose a reference-dependent choice model consisting of a simple two-stage procedure: elimination and optimization. In the elimination stage, the decision maker discards all alternatives which are "unambiguously" better than the reference point. In the optimization stage, she simply chooses the best alternative from the set of surviving alternatives. The consideration sets of both MO 2005 and MO 2014 violate Symmetry and Path Connectedness properties. On the other hand, both MO 2005 and MO 2014 impose that $\Gamma_x(T) = T \cap \Gamma_x(S)$ for $T \subset S$, which is violated in our model. In terms of choice, the models MO 2005 and 2014 do not allow choice reversals for a fixed status quo but the network choice does.

Dean et al. [2014] propose a reference-dependent model in which the attention is limited. $\Gamma_x(S) = (A(S) \cup x) \cap Q(x)$ where $x \in A(S)$ implies $x \in A(T)$ 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. Instead, they require that if the reference point is chosen in the smaller choice set, then it must also be chosen in the larger choice set.

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 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.

[2013]), limited attention (Lleras et al. [2010], Masatlioglu et al. [2012]. While these models have the element of limited attention, choices are not affected by a starting point. Even though the domains of these models are different than ours, we contrast our model with these models by fixing the starting point. The rest of the section, we focus on choice behavior for a fixed 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 more scarce 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, Symmetry, and Path Connectedness imply the first condition. On the other hand, our model assumes the opposite of the second condition.

Random Attention

Manzini and Mariotti [2014] stochastic choice model is closely related 2-step probabilistic network model. The main difference is that MM fix outside option as a "starting point" while we consider different alternatives in the choice set as different potential starting points. For a fixed starting point, it can be shown that i-Asymmetry and i-Independence in MM imply Axiom 20 in our model.

Brady and Rehbeck [2016] consider a modification of MM by relaxing the assumption that the probability that a set $D \subseteq S$ is a consideration set is independent of the choice set. In particular, they assume there exists a probability measure π on X such that

$$\mathcal{A}(D, S) = \frac{\pi(D)}{\sum_{D' \subseteq S} \pi(D')}$$

Obviously, for a fixed starting point, the 2-step network model is a subset of BR.

7 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 case of "perfectly rational" and "boundedly rational" agent, observed and unobserved starting points, deterministic and random network.

There are several interesting open questions. Firstly, in this paper, we only discuss symmetric links or undirected product network. An obvious open question is how the implications of such a model change when the links are asymmetric. While we treat the number 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 product network. 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. Lastly, one can also consider a K-step random network model for any K.

Appendix

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 $\gamma(x_i, x_{i+1}) = 1$ 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 $\gamma(x_i, x_{i+1}) = 1$ 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 $\gamma(x_i, x_{i+1}) = 1$ for $j \leq i < k$. Consider $\{x_1, \ldots, x_j, \ldots, x_k\} \subseteq S$. It satisfies the conditions that $x_1 = x$, $x_k = z$, $\gamma(x_i, x_{i+1}) = 1$ 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 $\gamma(x'_i, x'_{i+1}) = 1$ 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 $\gamma(x'_i, x'_{i+1}) = 1$ for $j \leq i < k$. Consider $\{x'_1, \ldots, x'_j, \ldots, x'_k\}$. It satisfies the conditions that $x'_1 = y$, $x'_k = z$, $\gamma(x'_i, x'_{i+1}) = 1$ 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 $\gamma(x_i, x_{i+1}) = 1$ 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 $\gamma(x_i, x_{i+1}) = 1$ 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 $\gamma(x_i, x_{i+1}) = 1$ 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 $\gamma(x, y) = 1$ if $y \in \Gamma_x(\{x, y\})$. Note that if $y \in \Gamma_x(\{x, y\})$, then by A.2, $\Gamma_y(\{x, y\}) = \Gamma_x(\{x, y\})$. Therefore, $\gamma(x, y) = \gamma(y, x)$ for all $x, y \in X$. Given γ, we define $N_x(S)$ as in the definition. Note that by $N_x(S)$ 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 $i < k$, and 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 therefore $y \in N_x(S)$. $\square$

Proof of Theorem 1

The proof of the "if part" is left to the reader. We prove the "only if part".

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 acyclical.*

Proof. Suppose $x_1 P x_2 P \cdots P x_n P x_1$. Then, there exists $S_1, \ldots, S_n$ with $S_i \supseteq \{x_i, x_{i+1}\}$ for $i < n$ and $S_n \supseteq \{x_1, x_n\}$ such that $c(S_i, z_i) = x_i$ for all $z_i \in S_i$. Consider the set $T = S_1 \cup S_2 \cup \cdots \cup S_n$. By Axiom 2, $c(T, x_1) = c(T, x_2) = \cdots = c(T, x_n)$ and $c(T, z_i) = c(T, x_i)$ for all i. Therefore, $c(T, x) = c(T, y)$ for all $x, y \in T$. Now by Axiom 1, we cannot have $x_i = c(T, x_i)$ since for $i > 1$, $x_i \in S_{i-1} \subseteq T$, but $x_i \neq c(S_{i-1}, x_i)$, and $x_1 \in S_n \subseteq T$, but $x_1 \neq c(S_n, x_1)$. Furthermore, for any $z \neq \{x_1, \ldots, x_n\}$, we cannot have $z = c(T, z)$ since $z \in S_i \subseteq T$ for some i, but $c(S_i, z) \neq z$. Hence we cannot assign any alternative to $c(T, x)$ without violating axioms. Therefore, P is acyclical. □

Now let $\succ$ be a transitive completion of P. We define γ as follows:

$$\gamma(x, y) = 1 \text{ if and only if } c(\{x, y\}, x) = c(\{x, y\}, y)$$

First, note that γ is symmetric. Define $N_x(S)$ as

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

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 $c(\{x, y\}, x) = y$, then 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 $z \neq c(S, x)$.[13] 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 by A.1, $z \in N_x(S)$. Finally, by A.2, $N_x(S) = N_z(S)$ which implies $c(S, x) \in N_x(S)$. □

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

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 $\{x_1, \dots, x_k\} \subseteq N_x(S) \subseteq S$ 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 Lemma 2

Proof. ($\Leftarrow$) B.1: Suppose $y \in \Gamma_x(T)$. Then, there exists $\{x_1, \dots, x_k\} \subseteq T \subseteq S$ such that $x_1 = x, x_k = y$, $k \leq K$ and $\gamma(x_i, x_{i+1}) = 1$ for $i < k$. Then, by definition, $y \in \Gamma_x(S)$.

B.2: Suppose $y \in \Gamma_x(S)$. This implies that there exists $\{x_1, \dots x_j\} \subseteq S$ with $x_1 = x$, $x_j = y$, $j \leq K$, and $\gamma(x_i, x_{i+1}) = 1$ for $i < k$. If $z \in \Gamma_y(S)$, then there exists $\{x_j, \dots, x_k\} \subseteq S$ such that $x_j = y$, $x_k = z$, $k - j + 1 \leq K$, and $\gamma(x_i, x_{i+1}) = 1$ for $j \leq i < k$. Consider $\{x_1, \dots, x_j, \dots, x_k\} \subseteq S$. Since $|S| \leq K$ it satisfies the conditions that $x_1 = x$, $x_k = z$, $k \leq K$, $\gamma(x_i, x_{i+1}) = 1$ 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, $\{x'_1, \dots, x'_j\} \subseteq S$ satisfies the conditions that $x'_1 = y$, $x'_j = x$, $j \leq K$, and $\gamma(x'_i, x'_{i+1}) = 1$ for $i < j$. Since $z \in \Gamma_x(S)$ there exists $\{x'_j, \dots, x'_k\} \subseteq S$ with $x'_j = x$, $x'_k = z$, $k - j + 1 \leq K$, and $\gamma(x'_i, x'_{i+1}) = 1$ for $j \leq i < k$. Consider $\{x'_1, \dots, x'_j, \dots, x'_k\} \subseteq S$. Since $|S| \leq K$ it satisfies the conditions that $x'_1 = y$, $x'_k = z$, $k \leq K$, and $\gamma(x'_i, x'_{i+1}) = 1$ for $i < k$. Therefore, $z \in \Gamma_y(S)$.

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

B.4: Suppose $y \in \Gamma_x(S)$. Then, there exists $\{x_1, \dots, x_k\} \subseteq S$ such that $x_1 = x, x_k = y$, $k \leq K$ and $\gamma(x_i, x_{i+1}) = 1$ for $i < k$. Let $T = \{x_1, \dots, x_k\}$. Then, $|T| \leq K$ and $y \in \Gamma_x(T)$.

(⇒) Let $\gamma(x,y)=1$ if $y\in\Gamma_x(\{x,y\})$. Notice that γ is symmetric since by B.2, $y\in\Gamma_x(\{x,y\})$ implies $\Gamma_y(\{x,y\})=\Gamma_x(\{x,y\})\ni x$. Define $N_x^K(S)$ as before. Note that $N_x^K(S)$ defined as such satisfies B.1-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 B.1, $x_{i+1}\in\Gamma_{x_i}(T)$ for $i<k$. Since $|T|\leq K$, by B.2, $\Gamma_{x_1}(T)=\Gamma_{x_2}(T)=\cdots=\Gamma_{x_n}(T)$. Therefore, $y\in\Gamma_x(T)$. Since $T\subseteq S$ by B.1, $y\in\Gamma_x(S)$.

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 B.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 B.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 B.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 $y\in N_x^K(S)$. □

Proof of Theorem 2

We prove the "only if" part.

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 $x_1Px_2P\cdots Px_nPx_1$. 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. □

Let $\succ$ be a transitive completion of P. Define γ as

$$\gamma(x,y)=1 \text{ if and only if } c(\{x,y\},x)=c(\{x,y\},y)$$

and $N_x^K(S)$ as

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

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

Proof. First note that if $c(\{x,y\},x) = x$, then $c(\{x,y\},x) \in N_x^K(\{x,y\})$. If $c(\{x,y\},x) = y$, then by Axiom 4 (or Axiom 5), $c(\{x,y\},x) = c(\{x,y\},y)$ which implies $\gamma(x,y) = 1$ and hence $y \in N_x^K(\{x,y\})$. 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 B.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 B.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 B.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 $\gamma(x_i,x_{i+1}) = 1$ for $i < k$. By definition, $\gamma(x_i,x_{i+1})$ 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 that either $c(T,x) = y$ or $c(T,x)Py$. Furthermore, since $S \supseteq T$ we have that 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 $c(S,x) \succ y$. ☐

Proof of Theorem 3

We prove the "only if" part.

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 exists $S_1, \ldots S_n$ 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 the transitive completion of P. Define γ as

$$\gamma(x, y) = 1 \text{ if and only if } C(\{x, y\}) \text{ is a singleton}$$

and let $N_x(S)$ be given by

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

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. $\square$

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

Proof. Notice that the claim is trivial for $S = \{x, y\}$. 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 $y \in N_x(S)$. $\square$

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 exists $y \in C(S)$ 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 $x = \text{argmax}(\succ, N_x(S))$. $\square$

Proof of Theorem 4

We prove the "only if" part.

For any $x \neq y$, define xPy if there exists $S \supseteq \{x, y\}$ such that $p(x, S) = 1$, and let $\succ$ be a transitive completion of P. The following claim shows that P is acyclic, and hence such $\succ$ exists.

Claim 1. *P is acyclic.*

Proof. Suppose $x_1 P x_2 P \cdots P x_n P x_1$. Then, there exists $S_1, \ldots, S_n$ with $S_i \supseteq \{x_i, x_{i+!}\}$ for $i < n$ and $S_1 \supseteq \{x_1, x_n\}$ such that $p(x_i, S_i) = 1$. Consider the set $T = S_1 \cup S_2 \cup \cdots \cup S_n$. Pick an element $x \in T$. Since $x \in T$, we must have that $x \in S_i$ for some i. If $x \neq x_i$, then $p(x, S_i) = 0$, and by Axiom 12, $p(x, T) = 0$. Suppose $x = x_i$. If $i > 1$, then $x \in S_{i-1}$ and $p(x, S_{i-1}) = 0$. If $i = 1$, then $x \in S_n$ and $p(x, S_n) = 0$. In either case, Axiom 12 implies $p(x, T) = 0$. We showed that $p(x, T) = 0$ for all $x \in T$ which is a contradiction. Hence P has no cycle. $\square$

Let $\gamma(x, y) = 1$ if $p(x, \{x, y\}) = 1$ or 0. Given γ, define $N_x(S)$ as

$$N_x(S) = \{y \in S | \exists \{x_1, x_2, \ldots, x_k\} \text{ such that} x_1 = x, x_k = y, \\ \text{and } \gamma(x_i, x_{i+1}) = 1 \text{ for } i < k\}$$

Given any binary choice set $\{x, y\}$ with $p(x, \{x, y\}) \in (0, 1)$, let

$$\frac{\beta(x)}{\beta(y)} = \frac{p(x, \{x, y\})}{p(y, \{x, y\})}$$

In addition, let

$$\sum_{x \in X} \beta(x) = 1$$

We need to show that there exists a solution to the system of equations with $\beta(x) > 0$ for all $x \in X$.

Claim 2. *Suppose $p(x, \{x, y\}), p(x, \{x, z\}), p(y, \{y, z\}) \in (0, 1)$. Then x, y, and z are chosen with positive probability when the choice set is $\{x, y, z\}$.*

Proof. Suppose not. Without loss of generality let $p(x, \{x, y, z\}) = 0$. Then by Axiom 13, either $p(y, \{x, y, z\}) = 1$ or $p(z, \{x, y, z\}) = 1$. Let $p(y, \{x, y, z\}) = 1$. But since $p(x, \{x, y\})$, $p(x, \{x, z\})$, and $p(y, \{y, z\}) \in (0, 1)$ this violates Axiom 15. $\square$

Claim 3. *Suppose $p(x, \{x, y\}), p(x, \{x, z\}), p(y, \{y, z\}) \in (0, 1)$. Then*

$$\frac{p(y, \{y, z\})}{p(z, \{y, z\})} = \frac{p(y, \{x, y\})}{p(x, \{x, y\})} \frac{p(x, \{x, z\})}{p(z, \{x, z\})}$$

Proof. By Claim 2, x, y, and z are chosen with positive probability when the choice set is $\{x,y,z\}$. Then by Axiom 16,

$$\frac{p(y,\{x,y,z\})}{p(z,\{x,y,z\})}=\frac{p(y,\{y,z\})}{p(z,\{y,z\})},\ \frac{p(y,\{x,y,z\})}{p(x,\{x,y,z\})}=\frac{p(y,\{x,y\})}{p(x,\{x,y\})},\ \frac{p(x,\{x,y,z\})}{p(z,\{x,y,z\})}=\frac{p(x,\{x,z\})}{p(z,\{x,z\})}$$

Since

$$\frac{p(y,\{x,y,z\})}{p(z,\{x,y,z\})}=\frac{p(y,\{x,y,z\})}{p(x,\{x,y,z\})}\frac{p(x,\{x,y,z\})}{p(z,\{x,y,z\})}$$

the previous equalities guarantee that

$$\frac{p(y,\{y,z\})}{p(z,\{y,z\})}=\frac{p(y,\{x,y\})}{p(x,\{x,y\})}\frac{p(x,\{x,z\})}{p(z,\{x,z\})}$$

□

Claim 3 guarantees that some equations in the system of equations will be redundant. In particular, if there are 3 alternatives x, y, and z we will have at most 4 equations to solve for $\beta(x)$, $\beta(y)$, and $\beta(z)$ and one of 4 equations will be implied by the others. If there are N alternatives we will have at most N relevant equations.

To see the existence of a solution to the system of equations with $\beta(x)>0$ for all $x\in X$, first let $\tilde{\beta}(x)=0$ for all $x\in X$ such that there exists no $y\neq x$ with $p(x,\{x,y\})\in(0,1)$. For all the other alternatives let

$$\frac{\tilde{\beta}(x)}{\tilde{\beta}(y)}=\frac{p(x,\{x,y\})}{p(y,\{x,y\})}$$

whenever $p(x,\{x,y\})\in(0,1)$. (If there are no such alternatives, then set $\beta(x)=1/N$ for all $x\in X$.) Suppose there are $M>0$ such alternatives. Then, using Claim 3, we will have at most M equations in M unknowns which we can solve for $\tilde{\beta}(x)>0$ uniquely. There are $N-M$ alternatives for which we set $\tilde{\beta}(x)=0$. Now for all x with $\tilde{\beta}(x)>0$, let $\beta(x)=\tilde{\beta}(x)/(N-M+1)>0$, and for all x with $\tilde{\beta}(x)=0$, let $\beta(x)=(1-\sum_{\tilde{\beta}(y)>0}\beta(y))/(N-M)>0$. It is easy to see that β satisfies all the equations and $\beta(x)>0$ for all $x\in X$.

Claim 4. *Suppose $p(x,S)=1$. Then, $x\in N_z(S)$ for all $z\in S$.*

Proof. If $S=\{x,y\}$, the claim is trivial. Suppose the claim is true for all S with $|S|<n$. Pick S with $|S|=n$. By Axiom 15, there exist non-singleton $T_1,T_2\subset S$ with $T_1\cup T_2=S$ and $T_1\cap T_2\neq\emptyset$ such that $p(x,T_1)=1$ and $p(y,T_2)=1$ for some $y\in S$. By induction $x\in N_{z_1}(T_1)$ for all $z_1\in T_1$ and $y\in N_{z_2}(T_2)$ for all $z_2\in T_2$. Since $T_1\cap T_2\neq\emptyset$ there exists $z^*\in T_1\cap T_2$ such that $x\in N_{z^*}(T_1)$ and $y\in N_{z^*}(T_2)$. Now pick

$z \in S$. Either $z \in T_1$ or $z \in T_2$. If $z \in T_1$, then by induction $x \in N_z(T_1)$ and by A.1, $x \in N_z(S)$. Suppose $z \in T_2$. Then, by induction, $y \in N_z(T_2)$ and by A.1, $y \in N_z(S)$. Since $y \in N_{z^*}(S)$ by A.2, $N_z(S) = N_y(S) = N_{z^*}(S)$. Therefore, $x \in N_z(S)$. ☐

Claim 5. *Suppose $p(x, S) = 1$ for some $x \in S$. Then,*

(i) $y \neq \mathrm{argmax}(\succ, N_y(S))$ *for all* $y \neq x$,

(ii) $p(x, S) = \frac{\beta(N_x(S))}{\beta(S)}$.

Proof. (i) By Claim 4, $x \in N_y(S)$ for all $y \in S$. By definition of P, $p(x, S) = 1$ implies xPy for all $y \neq x$. Since $\succ$ includes P we have $x \succ y$ for all $y \neq x$. Therefore, $y \neq \mathrm{argmax}(\succ, N_y(S))$.

(ii) By Claim 4, $x \in N_y(S)$ for all $y \in S$. By A.2, $y \in N_x(S)$ for all $y \in S$. Therefore, $N_x(S) = S$. Hence $p(x, S) = 1 = \frac{\beta(N_x(S))}{\beta(S)}$.

☐

Claim 5 proves the result for all sets S such that there exists $x \in S$ with $p(x, S) = 1$. Now we prove the result for all possible S.

Claim 6. *Suppose $p(x, S) \in (0, 1)$. Then $p(x, N_x(S)) = 1$.*

Proof. Pick $y \in N_x(S)$, $y \neq x$. By definition, there exists $\{x_1, x_2, \ldots, x_k\} \subseteq S$ with $x_1 = x, x_k = y$, and $p(x_i, \{x_i, x_{i+1}\}) = 0$ or 1 for $i < k$. Consider $\{x_1, x_2\}$ and $\{x_2, x_3\}$. We have that $\{x_1, x_2\} \cap \{x_2, x_3\} \neq \emptyset$, $p(x_i, \{x_1, x_2\}) = 1$, and $p(x_{i'}, \{x_2, x_3\}) = 1$ for some $i \in \{1, 2\}$ and $i' \in \{2, 3\}$. By Axiom 14, either $p(x_i, \{x_1, x_2, x_3\}) = 1$ or $p(x_{i'}, \{x_1, x_2, x_3\}) = 1$. Now suppose $p(x_i, \{x_1, \ldots, x_j\}) = 1$ and $p(x_{i'}, \{x_j, x_{j+1}\}) = 1$ for some $i \in \{1, \ldots, j\}$ and $i' \in \{j, j+1\}$. By Axiom 14, either $p(x_i, \{x_1, \ldots, x_j, x_{j+1}\}) = 1$ or $p(x_{i'}, \{x_1, \ldots, x_j, x_{j+1}\}) = 1$. By induction, we must have that $p(x_i, \{x_1, \ldots, x_k\}) = 1$ for some $i \in \{1, \ldots, k\}$. Now notice that $x \in \{x_1, \ldots, x_k\} \subseteq S$ and $p(x, S) > 0$. By Axiom 12, we must have $p(x, \{x_1, \ldots, x_k\}) > 0$. Therefore, $p(x, \{x_1, \ldots, x_k\}) = 1$. Since $\{x_1, \ldots, x_k\} \subseteq N_x(S)$, by Axiom 12, $p(y, N_x(S)) = 0$. Since the choice of $y \in N_x(S)$ was arbitrary, we must have that $p(y, N_x(S)) = 0$ for all $y \in N_x(S)$, $y \neq x$. We conclude that $p(x, N_x(S)) = 1$. ☐

Claim 7. *Suppose $p(x, S) > 0$ and $p(y, S) > 0$ for $x \neq y$. Then, $N_x(S) \cap N_y(S) = \emptyset$.*

Proof. Suppose $z \in N_x(S) \cap N_y(S)$. Then by A.2, $N_x(S) = N_z(S) = N_y(S)$. By Claim 6, $p(x, N_x(S)) = 1$ and $p(y, N_y(S)) = 1$. Since $N_x(S) = N_y(S)$ and $x \neq y$ we have a contradiction. Therefore, $N_x(S) \cap N_y(S) = \emptyset$. ☐

Claim 8. *Let $\{x_1, \ldots, x_n\} \subseteq S$ be such that $p(x_i, S) > 0$ and $\sum_i p(x_i, S) = 1$. Then, $N_{x_1}(S) \cup N_{x_2}(S) \cup \cdots \cup N_{x_n}(S) = S$.*

Proof. It is obvious that $N_{x_1}(S) \cup N_{x_2}(S) \cup \cdots \cup N_{x_n}(S) \subseteq S$. We show that $S \subseteq N_{x_1}(S) \cup N_{x_2}(S) \cup \cdots \cup N_{x_n}(S)$. Pick $y \in S$. If $y = x_i$ for some $i \in \{1, \ldots n\}$, then $y \in N_{x_i}(S)$. Suppose $y \notin \{x_1, \ldots, x_n\}$. By hypothesis, $p(y, S) = 0$. By Axiom 13, there exists $T \subseteq S$ containing y such that $p(x_i, T) = 1$ for some $x_i \in \{x_1, \ldots, x_n\}$. By Claim 4, $x_i \in N_y(T)$ and by A.1, $x_i \in N_y(S)$. By A.2, $y \in N_{x_i}(S)$. □

Claim 9. *Suppose $p(x, S) > 0$ and $p(y, S) > 0$ for $x \neq y$. Then, for any $z \in N_y(S)$, we have that $p(x, N_x(S) \cup z) > 0$ and $p(z, N_x(S) \cup z) > 0$.*

Proof. Since $N_x(S) \cup z \subseteq S$ by Axiom 12, $p(x, N_x(S) \cup z) > 0$. Suppose $p(x, N_x(S) \cup z) = 1$. Then, by Claim 4, $x \in N_z(N_x(S) \cup z)$ and by A.2, $z \in N_x(N_x(S) \cup z)$. Since $N_x(S) \cup z \subseteq S$, by A.1, $z \in N_x(S)$. But then, $N_x(S) \cap N_y(S) \neq \emptyset$. By Claim 7, this is a contradiction. Therefore, $p(x, N_x(S) \cup z) \in (0, 1)$. Furthermore, by Claim 6, $p(x, N_x(S)) = 1$ and hence by Axiom 12, $p(t, N_x(S) \cup z) = 0$ for all $t \in N_x(S)$. Therefore, $p(z, N_x(S) \cup z) \in (0, 1)$. □

Claim 10. *Suppose $p(x, S) > 0$ and $p(y, S) > 0$ for $x \neq y$. Then, for any $z \in N_y(S)$,*

$$\frac{p(x, N_x(S) \cup z)}{p(z, N_x(S) \cup z)} = \frac{\beta(N_x(S))}{\beta(z)}$$

Proof. By Claim 6, $p(x, N_x(S)) = 1$. Pick an element $z \in N_y(S)$. By previous claim, $p(x, N_x(S) \cup z) \in (0, 1)$. By Axiom 17,

$$\frac{p(x, N_x(S) \cup z)}{p(z, N_x(S) \cup z)} = \sum_{t \in N_x(S)} \frac{p(t, \{z, t\})}{p(z, \{z, t\})}$$

Note that we must have $p(z, \{z, t\}) \in (0, 1)$ since otherwise $z \in N_x(S)$ which contradicts Claim 7. By definition, $\frac{p(t,\{z,t\})}{p(z,\{z,t\})} = \frac{\beta(t)}{\beta(z)}$. Then,

$$\frac{p(x, N_x(S) \cup z)}{p(z, N_x(S) \cup z)} = \sum_{t \in N_x(S)} \frac{p(t, \{z, t\})}{p(z, \{z, t\})} = \sum_{t \in N_x(S)} \frac{\beta(t)}{\beta(z)} = \frac{\beta(N_x(S))}{\beta(z)}$$

□

Claim 11. *Suppose $p(x, S) > 0$ and $p(y, S) > 0$ for $x \neq y$. Then,*

$$\frac{p(x, N_x(S) \cup N_y(S))}{p(y, N_x(S) \cup N_y(S))} = \frac{\beta(N_x(S))}{\beta(N_y(S))}$$

Proof. By Claim 6, $p(x, N_x(S)) = 1$ and $p(y, N_y(S)) = 1$. By Claim 7, $N_x(S) \cap N_y(S) = \emptyset$. Furthermore, by Axiom 12, $p(x, N_x(S) \cup N_y(S)) > 0$ and $p(y, N_x(S) \cup N_y(S)) > 0$.

Therefore, by Axiom 17,

$$\frac{p(x, N_x(S) \cup N_y(S))}{p(y, N_x(S) \cup N_y(S))} = \sum_{t \in N_x(S)} \frac{p(t, N_y(S) \cup t)}{p(y, N_y(S) \cup t)}$$

By Claim 10, for any $t \in N_x(S)$,

$$\frac{p(t, N_y(S) \cup t)}{p(y, N_y(S) \cup t)} = \frac{\beta(t)}{\beta(N_y(S))}$$

Therefore,

$$\frac{p(x, N_x(S) \cup N_y(S))}{p(y, N_x(S) \cup N_y(S))} = \sum_{t \in N_x(S)} \frac{p(t, N_y(S) \cup t)}{p(y, N_y(S) \cup t)} = \sum_{t \in N_x(S)} \frac{\beta(t)}{\beta(N_y(S))} = \frac{\beta(N_x(S))}{\beta(N_y(S))}$$

□

Claim 12. *Let $\{x_1, \dots, x_n\}$ be such that $p(x_i, S) > 0$ for all i and $\sum_i p(x_i, S) = 1$. Then, for any $i, j \in \{1, \dots, n\}$,*

$$\frac{p(x_i, S)}{p(x_j, S)} = \frac{\beta(N_{x_i}(S))}{\beta(N_{x_j}(S))}$$

Proof. By Claim 11,

$$\frac{p(x_i, N_{x_i}(S) \cup N_{x_j}(S))}{p(x_j, N_{x_i}(S) \cup N_{x_j}(S))} = \frac{\beta(N_{x_i}(S))}{\beta(N_{x_j}(S))}$$

By Claim 6, $p(x_i, N_{x_i}(S)) = 1$ for all $i \in \{1, \dots, n\}$. By iteratively applying Axiom 16 and using the fact that $N_{x_1}(S) \cup \dots \cup N_{x_n}(S) = S$ by Claim 8, we get that

$$\frac{p(x_i, S)}{p(x_j, S)} = \frac{\beta(N_{x_i}(S))}{\beta(N_{x_j}(S))}$$

□

Claim 13. *Let $\{x_1, \dots, x_n\}$ be such that $p(x_i, S) > 0$ for all i and $\sum_i p(x_i, S) = 1$.*

(i) If $y \notin \{x_1, \dots, x_n\}$, then $y \neq \operatorname{argmax}(\succ, N_y(S))$,

(ii) If $y \in \{x_1, \dots, x_n\}$, then $y = \operatorname{argmax}(\succ, N_y(S))$ and $p(y, S) = \frac{\beta(N_y(S))}{\beta(S)}$.

Proof. (i) Suppose $y \notin \{x_1, \dots x_n\}$. Then, $p(y, S) = 0$. By Axiom 13, there exists $T \subseteq S$ containing y such that $p(x_i, T) = 1$ for some $x_i \in \{x_1, \dots, x_n\}$. By Claim 4, $x_i \in N_y(T)$ and by A.1 $x_i \in N_y(S)$. By definition $x_i P y$ and hence $x_i \succ y$. Therefore, $y \neq \operatorname{argmax}(\succ, N_y(S)))$.

(ii) Suppose $y = x_i$ for some $x_i \in \{x_1, \ldots, x_n\}$. By Claim 6, $p(x_i, N_{x_i}(S)) = 1$. By definition $x_i P y$ for all $y \in N_{x_i}(S)$, $y \neq x$. Hence $x_i = \text{argmax}(\succ, N_{x_i}(S))$. Since $p(x_1, S) + \cdots + p(x_i, S) + \cdots + p(x_n, S) = 1$ we have that

$$\frac{p(x_1, S) + \cdots + p(x_i, S) + \cdots + p(x_n, S)}{p(x_i, S)} = \frac{1}{p(x_i, S)}$$

By Claim 12, for any $x_j \in \{x_1, \ldots, x_n\}$,

$$\frac{p(x_i, S)}{p(x_j, S)} = \frac{\beta(N_{x_i}(S))}{\beta(N_{x_j}(S))}$$

Then,

$$\frac{\beta(N_{x_1}(S)) + \cdots + \beta(N_{x_i}(S)) + \cdots + \beta(N_{x_n}(S))}{\beta(N_{x_i}(S))} = \frac{1}{p(x_i, S)}$$

By Claim 8, $N_{x_1}(S) \cup \cdots \cup N_{x_n}(S) \cup \cdots \cup N_{x_n}(S) = S$. Therefore, $\beta(N_{x_1}(S)) + \cdots + \beta(N_{x_i}(S)) + \cdots + \beta(N_{x_n}(S)) = \beta(S)$, and hence

$$p(x_i, S) = \frac{\beta(N_{x_i}(S))}{\beta(S)}$$

This completes the proof.

$\square$

Proof of Theorem 5

Let $x \succ y$ if $p_y(x, \{x, y\}) > 0$.

Claim 1. *$\succ$ is complete and transitive*

Proof. By Axiom 18, for any x and y, either $p_x(y, \{x, y\}) > 0$ or $p_y(x, \{x, y\}) > 0$. Therefore, either $x \succ y$ or $y \succ x$.

Suppose $x_1 \succ x_2 \succ x_3 \succ x_1$. Then, $p_{x_1}(x_3, \{x_1, x_3\}) > 0$, $p_{x_2}(x_1, \{x_1, x_2\}) > 0$, and $p_{x_3}(x_2, \{x_2, x_3\}) > 0$. Alternatively, $p_{x_1}(x_1, \{x_1, x_3\}) < 1$, $p_{x_2}(x_2, \{x_1, x_2\}) < 1$, and $p_{x_3}(x_3, \{x_2, x_3\}) < 1$. Then by Axiom 18, $p_{x_1}(x_1, \{x_1, x_2, x_3\}) \leq p_{x_1}(x_1, \{x_1, x_3\}) < 1$, $p_{x_2}(x_2, \{x_1, x_2, x_3\}) \leq p_{x_2}(x_2, \{x_1, x_2\}) < 1$, $p_{x_3}(x_3, \{x_1, x_2, x_3\}) \leq p_{x_3}(x_3, \{x_2, x_3\}) < 1$. But then there is no dominant alternative in $\{x_1, x_2, x_3\}$. This contradicts Axiom 19. Since $\succ$ is complete we have $x_1 \succ x_3$. $\square$

Let $\gamma(x,y) = \max\{p_x(y,\{x,y\}), p_y(x,\{x,y\})\}$ and define $\mathcal{A}_x(D,S)$ as

$$\mathcal{A}_x(D,S) = \prod_{y\in D}\gamma(x,y)\prod_{z\in S\setminus D}(1-\gamma(x,z))$$

Claim 2. *Suppose $z \in S\setminus D$. Then, $\mathcal{A}_x(D,S) = \mathcal{A}_x(D,S\setminus z)\mathcal{A}_x(\{x\},\{x,z\})$.*

Proof. Using the definition of $\mathcal{A}_x(D,S)$ and some algebraic manipulation we get that

$$\begin{aligned}\mathcal{A}_x(D,S) =& \prod_{y\in D}\gamma(x,y)\prod_{t\in S\setminus D}(1-\gamma(x,t))\\ =&\Big(\prod_{y\in D}\gamma(x,y)\prod_{t\in S\setminus(D\cup\{z\})}(1-\gamma(x,t))\Big)(1-\gamma(x,z))\\ =&\mathcal{A}_x(D,S\setminus z)\mathcal{A}_x(\{x\},\{x,z\})\end{aligned}$$

□

Claim 3. $p_x(y,S) = \sum_{y \text{ is } \succ-\text{best in } D} \mathcal{A}_x(D,S)$

Proof. Suppose $S=\{x,y\}$. Firstly, if $x \succ y$, then by definition, $p_x(y,\{x,y\})=0$ or $p_x(x,\{x,y\})=1$. Hence $p_x(x,\{x,y\}) = \mathcal{A}_x(\{x\},\{x,y\}) + \mathcal{A}_x(\{x,y\},\{x,y\})$. Suppose $y\succ x$. Then, by definition, $p_y(x,\{x,y\})=0$ and hence $\mathcal{A}_x(\{x,y\},\{x,y\}) = \gamma(x,y) = p_x(y,\{x,y\})$.

Now suppose the claim is true for all S with $|S|<n$. Pick S with $|S|=n$. Let $z = \text{argmax}(\succ, S)$. Then, we must have $p_z(t,S)=0$ for all $t\in S\setminus z$. Otherwise we would have $t\succ z$. Hence $p_z(z,S)=1$ or z is a dominant alternative in S. By Axiom 20, $p_x(y,S) = p_x(y,S\setminus z)p_x(x,\{x,z\})$. Since $z\succ x$ by induction hypothesis, $p_x(x,\{x,z\}) = \mathcal{A}_x(x,\{x,z\})$. Furthermore, since $|S\setminus z|<n$, $p_x(y,S\setminus z) = \sum_{y \text{ is } \succ-\text{best in } D}\mathcal{A}_x(D,S\setminus z)$.

Therefore,

$$\begin{aligned}p_x(y,S) =& p_x(y,S\setminus z)p_x(x,\{x,z\})\\ =& \sum_{y \text{ is } \succ-\text{best in } D}\mathcal{A}_x(D,S\setminus z)\mathcal{A}_x(x,\{x,z\})\\ =& \sum_{\substack{z\notin D\\ y \text{ is } \succ-\text{best in } D}}\mathcal{A}_x(D,S) && \text{By Claim 2}\\ =& \sum_{y \text{ is } \succ-\text{best in } D}\mathcal{A}_x(D,S) && \text{Since } z\succ y\end{aligned}$$

We proved the claim for all $y\in S\setminus z$. Since $z = 1-\sum_{y\in S\setminus z}p_x(y,S)$, the claim is also satisfied for z. □

References

J. L. Aaker. Dimensions of brand personality. *Journal of marketing research*, pages 347–356, 1997.

J. R. Anderson. A spreading activation theory of memory. *Journal of verbal learning and verbal behavior*, 22(3):261–295, 1983.

J. R. Anderson and G. H. Bower. Human associative memory. 1973.

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

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

R. L. Brady and J. Rehbeck. Menu-dependent stochastic feasibility. *Econometrica*, 84 (3):1203–1223, 2016.

B. J. Bronnenberg, J. B. Kim, and C. F. Mela. Zooming in on choice: How do consumers search for cameras online? *Marketing Science*, 2016.

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

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

Y. Chen and S. Yao. Sequential search with refinement: Model and application with click-stream data. *Available at SSRN 2130646*, 2015.

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

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

M. Dean, O. Kibris, and Y. Masatlioglu. Limited attention and status quo bias. *Available at SSRN 2519242*, 2014.

H. C. Ellis and R. R. Hunt. *Fundamentals of cognitive psychology* . Brown & Benchmark/Wm. C. Brown Publ, 1993.

J. L. L. Garber. The package appearance in choice. *Advances in consumer research*, 22:653–660, 1995.

D. Gentner and A. Stevens. Mental models. *Cognitive science*, 1983.

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

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

E. Honka and P. K. Chintagunta. Simultaneous or sequential? search strategies in the us auto insurance industry. *Search Strategies in the US Auto Insurance Industry (November 20, 2015)*, 2015.

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

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

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

S. Koulayev. Estimating demand in online search markets, with application to hotel bookings. *Federal Reserve Bank of Boston Working Paper*, (09-16), 2010.

G. Linden, B. Smith, and J. York. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing*, 7(1):76–80, 2003.

J. S. Lleras, Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay. When more is less: Limited consideration. *Unpublished, October*, 2010.

R. D. Luce. Individual choice behavior. 1959.

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

P. Manzini and M. Mariotti. Sequentially rationalizable choice. *The American Economic Review*, pages 1824–1839, 2007.

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

P. Manzini and M. Mariotti. Stochastic choice and consideration sets. *Econometrica*, 82(3):1153–1176, 2014.

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

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

Y. Masatlioglu and E. A. Ok. A canonical model of choice with initial endowments. *The Review of Economic Studies*, page rdt037, 2013.

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

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

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

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

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

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

A. Rubinstein and Y. Salant. A model of choice from lists. *Theoretical Economics*, 1 (1):3–17, 2006.

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

A. Tversky and D. Kahneman. Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, pages 1039–1061, 1991.