Global Research Unit Working Paper #2019-002

Choice with Limited Capacity

Sen Geng, Xiamen University
Erkut Y. Ozbay, University of Maryland

Choice with Limited Capacity *

Sen Geng[1] and Erkut Y. Ozbay[2]

[1]Xiamen University

[2]University of Maryland

Abstract

The number of alternatives in a choice problem may have an impact on the consideration of people with limited capacity: a decision maker considers all the available alternatives if the number of alternatives does not exceed her capacity; otherwise, she applies a rationale to reduce the number of alternatives to within her capacity. We provide the necessary and sufficient conditions for a choice function to be rationalizable by a shortlisting with limited capacity and recover the unobserved capacity from the observed choices, which turns out to be unique when choice reversals exist. Secondly, for the settings in which consideration sets are observable, we provide the necessary and sufficient conditions for a consideration function to be generated by the shortlisting with limited capacity procedure. Finally, we investigate a special case in which exactly the same number as the capacity of a decision maker are left when the elimination stage is triggered and show that only certain capacities are consistent with this special case.

Keywords: consideration set, shortlisting, limited capacity

JEL: D11, D81

*Sen Geng: sg1670@xmu.edu.cn, Erkut Y. Ozbay: ozbay@umd.edu. This paper subsumes Chapter three of Sen Geng's doctoral dissertation. We thank Andrew Caplin, Mark Dean, Brett Graham, Yusufcan Masatlioglu, David Pearce, Ariel Rubinstein, Yongsheng Xu, and seminar attendants at New York University, Tsinghua University, University of Toronto, and Xiamen University for their helpful comments.

1 Introduction

A decision maker (DM), while choosing from a set of available alternatives, chooses the best one among the ones she considers. According to the classical choice theory, the DM considers all the available alternatives, however, in reality, a DM may have a *limited capacity*, and hence in the presence of abundance of alternatives, she overlooks some of them.[1]

Numerous factors contribute to the limited capacity of a DM. Miller (1956) argues that each individual has a capacity such that the number of objects one can hold in one's working memory is limited. Hence, one fails to recall or to evaluate all alternatives beyond her cognitive capacity. Secondly, this capacity may be due to some constraints that are not in the control of the DM. For example, the DM may need to decide within a certain time that is not long enough to evaluate all the alternatives (see Geng 2016). Thirdly, the capacity may arise as a result of the DM's trade-off between search costs and benefits (see Stigler 1961). For example, when deciding on a product that is not very important or the differences across the alternatives are negligible, the number of considered alternatives will be limited. Fourthly, the DM may want to set a capacity in order to reduce the decision cost (see Ergin and Sarver 2010, and Ortoleva 2013).

The limited capacity accompanied by the number of alternatives beyond one's capacity makes full consideration impossible, and the DM needs to limit the number of alternatives considered to within her capacity. For example, a consumer shopping for a laptop in an online store usually uses a filtering tool to consider only a few candidate laptops. Similarly, while using a web-search engine, the DM considers only the products in the first page of the search results. A job recruiting committee receiving hundreds of job applications for one position often first selects a small sample of applicants to interview based on their resumes, and then identifies the most competent applicant among those interviewed. In the case of choosing a supplier where price and reliability of the supplier is important, Dulleck et al. (2011) show that consumers use a lexicographic approach such that they shortlist suppliers by using the price variable only (for example n-cheapest suppliers they want to consider). In all these examples, too many alternatives overwhelm the DM, and hence the DM forms a consideration set by eliminating dominated alternatives based on a rationale as in the shortlisting procedure of Manzini and Mariotti (2007).

Two-stage procedures that relax the full consideration assumption are at the core

[1] For example, Hauser and Wernerfelt (1990) report that the consumers on average consider 3 deodorants, 4 shampoos, 2 air fresheners, 4 laundry detergents, and 4 coffee brands. Similar limited number of considerations can be found in investment decisions (e.g., Huberman and Regev 2001), in university choice (e.g., Rosen et al. 1998), in job search (e.g., Richards et al. 1975), and in household grocery consumption (e.g., Demuynck and Seel 2018).

of the recent bounded rationality literature, and the abundance of alternatives is the basic motive for limiting the consideration set (e.g., Lleras et al. 2017). The flip side of this motivation is that when there are not too many alternatives, the DM should not be overwhelmed and should be able to consider all of them, nevertheless no attention is given to the number of alternatives and to the DM's capacity. In this paper, we study a two-stage procedure where the DM has a *limited capacity* for the number of alternatives to consider: For the choice problems where the number of alternatives is within the DM's capacity, the DM considers all the alternatives and chooses the best alternative. However, when the number of alternatives exceeds the DM's capacity, by using a rationale, the DM limits the number of alternatives considered to within her capacity and then chooses the best one. We call this procedure "shortlisting with limited capacity".

An important aspect of our approach is that when characterizing the shortlisting with limited capacity procedure, we do not assume that the capacity is observable. Instead, we provide the necessary and sufficient conditions for a choice function to be rationalizable by a shortlisting with limited capacity and derive the capacity of a DM from her observed choices. Furthermore, we show that the capacity is uniquely derived when the DM exhibits any choice reversals. The behavioral axioms include Acyclic Binary Choice and Manzini and Mariotti (2007)'s Weak WARP, which requires that if an option is chosen over another option both in the binary choice set and in a large choice including the two options, then the unchosen option can never be chosen in any of its subset that includes the chosen option. In addition, the axiom of Expansion that an option chosen from each of two sets is also chosen from their union is assumed in three different domains rather than assumed in all domains as in Manzini and Mariotti (2007).

In addition to the characterization based on choice data, we also provide a characterization based on consideration set data. The development of the new tools, such as eye tracking, makes it possible to observe a non-choice data that complements choice data. For example, Reutskaja et al. 2011 demonstrate that the consideration sets of consumers can be observed. We investigate the properties of consideration sets that are generated by a shortlisting with limited capacity procedure. The characterization based on consideration sets allows us to relate our setup with the properties of the procedure-free consideration sets (e.g., Masatlioglu et al. 2012 and Lleras et al. 2017). It turns out that the consideration set property in Lleras et al. (2017) is quite general by providing a minimal requirement on the consideration set to capture the more is less phenomenon. According to their competition filter property, if an option is considered in a feasible set, then it must be considered in any subset of this feasible set whenever that option is available. The shortlisting with limited capacity procedure refines the competition filter such that all the feasible alternatives are considered in a set if and only if all the feasible alternatives are considered in any set that has the same size.

Finally, we investigate a special case of the model in which exactly the same number alternatives are left when elimination is triggered provided that there are enough number of alternatives. For example, a DM has a criterion such that alternatives can be ranked with a linear order, and she considers top-k alternatives based on this criterion, e.g., considering only the items that are appeared in the first page of the search. This is similar to Example 3 in Salant and Rubinstein (2008) but the capacity of a DM is assumed to be observable there. We provide a characterization of this special case and elicit the capacity based only on observed choices. It turns out that this special case imposes a surprisingly tight structure on choice behavior in the sense that at most one choice reversal can be allowed.

The rest of the paper will be structured as follows: in Section 2, we define and characterize the limited capacity model, and discuss the related literature. In Section 3, we investigate a special case in which the DM considers exactly a certain number of options. Section 4 concludes.

2 Limited Capacity Model

Let X be a grand choice set consisting of finite options, i.e. $|X| = N > 2$, and $\mathcal{X} = 2^X/\{\emptyset\}$ denote the set of all nonempty subsets of $\mathcal{X}$, which is interpreted as the the collection of all the (objective) feasible sets. Let $\Gamma : \mathcal{X} \mapsto \mathcal{X}$ be a consideration function satisfying that $\emptyset \neq \Gamma(S) \subseteq S$, with an interpretation that $\Gamma(S)$ characterizes the consideration set of a DM for any feasible set, S.

In our model, the DM can be overwhelmed by the number of feasible options. She has a limited capacity of k on the number of options she can consider, where k is a natural number less than or equal to N. If the DM sees that the number of options in a feasible set, $|S|$, is less than or equal to her capacity, k, she considers all of them. However, if it exceeds her capacity, the DM eliminates some of the options to render the number of options she considers less than or equal to her capacity:

Definition 1. *A consideration function, $\Gamma : \mathcal{X} \mapsto \mathcal{X}$, is called a* ***consideration function with capacity-k*** *(denoted as Γ_k) if there exists a capacity $k \leq N$ such that for each feasible set $S \in \mathcal{X}$ there exists $S' \subseteq S$ with $|S'| \leq k$:*

$$\Gamma_k(S) = \begin{cases} S & \text{if } |S| \leq k, \\ S' & \text{if } |S| > k. \end{cases}$$

This definition of consideration function with capacity-k captures not only the idea that the DM's consideration is limited by a capacity constraint as widely documented in the psychology literature (e.g., Miller 1956), but also the idea of "more is less" that a DM overlooks some of the options only when the feasible options are more than the DM's capacity (e.g., Schwartz 2005). However, as argued in Lleras et al. (2017), the consideration sets without any restriction has no falsifiable empirical content. Similarly, the consideration function with capacity-k does not provide any structure or limit on the formation of the consideration sets, because an assumption that the DM has a capacity of one and considers only the chosen alternative rationalizes any choice behavior.

Motivated by the marketing literature, we assume that individuals use some tools to reduce the number of options to be considered. For example, an online shopper considers all the options if the total number of the search results is k or less, otherwise she narrows down the size of the search results by filtering the results according to a rationale, such as a certain price range (e.g., Alba et al. 1997). The DM sorts the options based on a rationale, such as price, and considers only top-k options such as k-cheapest options.

Following Manzini and Mariotti (2007), a rationale, $P \subseteq X \times X$, is an asymmetric binary relation defined on X such that $(x, y) \in P$ (xPy for convenience), indicating that option x eliminates or dominates option y, and $max(S, P) := \{x \in S | \text{no } y \in S \text{ such that } yPx\}$ is the set of undominated options in S with respect to P. Hence, when the number of options in S is greater than her capacity k, the DM eliminates the dominated options in S with respect to a rationale P.

Definition 2. *A consideration function with capacity-k, Γ_k, is called a* ***shortlisting with capacity-k*** *(denoted as Γ_k^P) if there exists a rationale P such that for each feasible set $S \in \mathcal{X}$:*

$$\Gamma_k^P(S) = \begin{cases} S & \text{if } |S| \leq k, \\ max(S, P) & \text{if } |S| > k, \end{cases}$$

and $|max(S, P)| \leq k$.[2]

Among the options the DM considers, she chooses the best option based on her preference (linear order) that is a complete, asymmetric, and transitive relation, $\succ \subseteq X \times X$.[3] It is important to note that none of the capacity-k, the rationale P, and the preference $\succ$ are observable in our setup. We first consider the cases in which only the

[2]Formally, P is an asymmetric binary relation with width no more than k. Here width(X, P):=$sup\{|Y| :$ Y is a choice set in which any two options are not related according to P.$\}$

[3]Manzini and Mariotti (2007) requires this second relation only to be asymmetric. We put the stronger requirement on $\succ$ to be complete and transitive as in Masatlioglu et al. (2012) in order to highlight that the DM is rational among the options she considers.

choices of the DM are observable. A choice function $c : \mathcal{X} \mapsto X$ assigns a unique option $c(S) \in S$ for each feasible set $S \in \mathcal{X}$.

Definition 3. *A choice function, c, is **rationalizable by a shortlisting with capacity-k** if there exists a capacity k, a rationale P, and a linear order $\succ$, such that for any $S \in \mathcal{X}$,*

$$c(S) = max(\Gamma_k^P(S), \succ)$$

In classical choice theory, the necessary and sufficient condition for an observed choice function to be rationalizable is the Weak Axiom of Revealed Preference (WARP). In terms of choice functions, WARP is equivalently stated as an option can never be chosen in the presence of another option if the latter one is ever chosen in a binary choice problem of the two options. Formally, $x = c(\{x, y\})$ implies that $y \neq c(S)$, for any S including x and y. Since the DM is assumed to always consider all the feasible options, i.e. the capacity of the DM is N, the classical representation theorem can be restated in our setup as a choice function satisfies the WARP if and only if it is rationalizable by a shortlisting with capacity-N. As in any two-stage choice model, the choice function that satisfies the WARP is the least informative one. In our setup, this means that it can be rationalizable by a shortlisting with any capacity:

Proposition 1. *If a choice function is rationalizable by a shortlisting with capacity-N, then it is rationalizable by a shortlisting with capacity-k for any $k < N$.*

Proof. Suppose that a choice function is rationalizable by a shortlisting with capacity-N, i.e., $c(S) = max(S, \succ)$, where $\succ$ is a preference relation. For any $k < N$, define a rationale $P = \succ$ and a shortlisting with capacity-k, $\Gamma_k^P(S) = c(S)$ if $|S| > k$ and $= S$ if $|S| \leq k$. It is straightforward to establish that $c(S) = max(\Gamma_k^P(S), \succ)$ and the choice function is rationalizable by a shortlisting with capacity-k. □

Interestingly, in this setup, being rationalizable by a shortlisting with capacity-1 coincides with being rationalizable by a shortlisting with capacity-N, i.e., coincides with full rationality. As shown in Proposition 2, a DM who has a capacity of one share the same choice function with a DM who has a capacity of N as long as the rationale employed by the first DM coincides with the second DM's preference.

Proposition 2. *A choice function is rationalizable by a shortlisting with capacity-1 if and only if it is rationalizable by a shortlisting with capacity-N.*

Proof. If a choice function can be rationalized by the pair $(\Gamma_N^P, \succ)$, then clearly it can be rationalized by the pair $(\Gamma_1^{\succ}, \succ')$, where $\succ'$ refers to an arbitrary preference relation.

Now assume that a choice function can be rationalized by the pair $(\Gamma_1^P, \succ)$. Note that P is complete in this case since $|max(\{x,y\}, P)| = 1$ for any $x, y \in X$. We then show that P is transitive. Suppose by contradiction that $xPyPzPx$, then $max(F_1^P(\{x,y,z\}), \succ) = \emptyset$. However, being rationalizable by a shortlisting with capacity-1 implies that $max(\Gamma_1^P(\{x,y,z\}), \succ) = c(\{x,y,z\}) \neq \emptyset$. Hence, P is transitive and then the choice function can be rationalized by the pair $(\Gamma_N^{P'}, P)$, where P' refers to an arbitrary rationale. □

We now look at choice functions that do not satisfy the WARP. Proposition 2 suggests that these choice functions are not rationalizable by a shortlisting with capacity-1. Then, if a choice function that does not satisfy the WARP is rationalizable by a shortlisting with limited capacity, the capacity k must be 2 or higher. In other words, the DM who has such a choice function should fully consider feasible sets of size 2 at the very minimum. Thus, in any binary choice problem, the DM chooses an option according to her preference. Therefore, her choice should not exhibit any cycles in binary choices in order to be consistent with a transitive preference. In addition, for choice functions that satisfy the WARP, exhibiting no cycle in binary choices is an immediate implication of the WARP. Thus, it is natural to assume acyclic binary choices:

Axiom 1 (Acyclic Binary Choice). *$x = c(\{x,y\})$ and $y = c(\{y,z\})$ implies that $x = c(\{x,z\})$.*

Manzini and Mariotti (2007) proposes that if an option is chosen over another option both in the binary choice set and in a large choice set including the two options, then the unchosen option can never be chosen in any of its subset that includes the chosen option. They label this property Weak WARP. In a similar spirit, we impose the same requirement on a DM's choice behavior.

Axiom 2 (Weak WARP). *Consider $\{x,y\} \subset S \subset T$. $x = c(\{x,y\}) = c(T)$ implies that $y \neq c(S)$.*

Manzini and Mariotti (2007) also proposes the property of Expansion that an option chosen from each of two sets is also chosen from their union. Formally, $x = c(S) = c(T)$ implies that $x = c(S \cup T)$. Nevertheless, this property is inconsistent with the spirit of choice with limited capacity. The reason is that in their model the elimination machine always operates regardless of the size of choice set, but in the present model the elimination machine first steps in only when the size exceeds the DM's capacity and the DM is fully rational when within capacity. It then is conceivable that in our setup the property of Expansion holds when all the sizes of the three choice sets fall within capacity or when all the sizes fall beyond capacity, and the property may fail when the elimination machine operates on some of the three choice sets and does not operate on other choice set(s).

More importantly, we show in Proposition 3 that the axiom of acyclic binary choice and the property of Expansion is equivalent to the WARP. This suggests that an addition of the property of Expansion makes it impossible to accommodate choice behavior that violates the WARP. In other words, given that the axiom of acyclic binary choice has been assumed, the axiom of Expansion should not be assumed in all domains, if we would like to investigate interesting choice settings in which the WARP may be violated.

Proposition 3. *A choice function satisfies the axiom of Acyclic Binary Choice and the property of Expansion if and only if it satisfies the WARP.*

Proof. See the proof in the Appendix A.1. □

Motivated by the above observation, we make the assumption of Expansion in three different domains rather than in all domains as in Manzini and Mariotti (2007). Since the domains in which the property of Expansion possibly holds are related to a DM's capacity, we introduce below the concepts of WARP-violating choice function and threshold capacity before we introduce the three axioms of Expansion.

Definition 4. *We say $x = c(\{x, y\})$ and $y = c(S)$ for some $S \supset \{x, y\}$ constitutes a **WARP-violating choice pair**. S is named as a **WARP-violating choice set**. We say a choice function is a **WARP-violating choice function** if it has at least one WARP-violating choice pair.*

This definition deserves a few comments. First, it is obvious that the WARP holds if and only if there exists no WARP-violating choice pair. Hence, a WARP-violating choice function is equivalent to a choice function that is not fully rationalizable. Second, our definition of WARP-violating choice pair restricts itself to the scenarios of a two-option choice set and a multiple-option choice set. One may wonder how we should categorize the scenarios in which seemingly inconsistent choices occur in two multiple-option choice sets, e.g., $x = c(S)$ and $y = c(S^{'})$ for some S and S' such that $\{x, y\} \subset S \cap S'$. In fact, the seemingly inconsistent choice situation has a root in its degenerate version since either $x = c(\{x, y\})$ or $y = c(\{x, y\})$. Thus, the present definition of WARP-violating choice pair suffices. Third, since the multiple-option choice set in a WARP-violating choice pair contains more relevant information than the two-option choice set, we assign to the multiple-option choice set a name of WARP-violating choice set.

Among all WARP-violating choice sets of a WARP-violating choice function, the smallest ones are particularly interesting. By definition, the WARP always holds for choice sets whose size is smaller than the size of the smallest WARP-violating choice sets, and the WARP starts to fail for choice sets whose size is equal to or exceeds the size of the smallest WARP-violating choice sets. If we define the size less than one as a threshold capacity, then it is clear that the WARP always holds when the size of choice sets falls within the threshold capacity and the WARP starts to fail when the size exceeds the threshold capacity. In the degenerate case where no WARP-violating choice sets exists, i.e., the WARP always holds, it is natural to define the threshold capacity as N.

Definition 5. *The* ***threshold capacity*** *of a choice function is* $k^* = min\{|S| - 1 : S$ *is an arbitrary WARP-violating choice set given the choice function*$\}$ *if it is a WARP-violating choice function, and is* $k^* = N$ *if it is not a WARP-violating choice function.*

Clearly, the above definition suggests that the threshold capacity of a choice function may range from 2 to N, and in particular, the threshold capacity of a WARP-violating choice function may range from 2 to $N - 1$. Additionally, since a DM who follows the shortlisting with limited capacity procedure chooses the best alternative within her capacity, WARP violation cannot be observed for choice sets whose sizes fall within her capacity. Hence, a DM's capacity, if it exists, should not exceed the threshold capacity of her choice function.

Our first assumption on Expansion is that the property of Expansion holds when all the sizes of the involved choice sets exceed the threshold capacity. The motivation for this assumption is that when all the sizes of the involved choice sets exceed the threshold capacity, they all exceed the DM's capacity and then the elimination machine is always operating. Hence, it is natural to require the property of Expansion in this case for the same reason as in Manzini and Mariotti (2007).

Axiom 3 (Limited Expansion). *When* $|S|, |T| > k^*$, $x = c(S) = c(T)$ *implies that* $x = c(S \cup T)$.

We now consider the situation in which the chosen option of a first choice set is not chosen any more once the chosen option of a second choice set is added to the first choice set. In a fully rational setup, it suggests that the second chosen option is preferred over the first chosen option. So the second chosen option must still be chosen when the first chosen option is added to the second choice set. In the setup of Manzini and Mariotti (2007), the property of Expansion suggests that the second chosen option must be chosen in the binary choice of the first and second chosen options, which in turn implies that the second chosen option must still be chosen when the first option is added to the second choice set. This observation motivates us to make a second assumption on Expansion:

Axiom 4 (Limited Weak Expansion). *When* $|S|, |T| > k^*$, $c(T) \neq c(T \cup c(S))$ *implies that* $c(S) = c(S \cup c(T))$.

Finally, we highlight a behavioral implication that limited capacity has on elimination rationale. On the one hand, choice by shortlisting with limited capacity requires that the size of any choice set after elimination must fall within a DM's capacity and in turn fall within the threshold capacity. On the other hand, it seems that no elimination relation can exist between two options if both are ever chosen in the presence of each other in choice sets whose size exceeds the threshold capacity, i.e., if $x = c(S)$ and $y = c(S')$ for some S and S' satisfying that $|S|, |S'| > k^*$, and $\{x, y\} \subset S \cap S'$. One could directly assume that for any choice set in which no elimination relation can exist between any two options, the size of the choice set must fall within the threshold capacity. Nevertheless, the assumption is not minimal given that we have imposed other axioms, especially the axiom of acyclic binary choice. More precisely, Proposition 3 implies that a choice function satisfies the axiom of Acyclic Binary Choice and the property of Expansion up to a certain size of choice sets if and only if it satisfies the WARP up to the same size of choice sets. The latter one further implies that the size of such choice sets must fall within the threshold capacity. This motivates us to assume that for any choice set in which no elimination relation can exist between any two options, the property of Expansion holds up to the same size of choice sets. In this way, the requirement that the size of such choice sets falls within the threshold capacity is implied instead of assumed.

Axiom 5 (Conditional Expansion). *If there exists a choice set T satisfying that for any two options $x, y \in T$, $x = c(S)$ and $y = c(S')$ for some S and S' satisfying that $|S|, |S'| > k^*$, and $\{x, y\} \subseteq S \cap S'$, then $x = c(S_1) = c(S_2)$ and $|S_1 \cup S_2| \leq |T|$ implies that $x = c(S_1 \cup S_2)$.*

Our representation theorem (i.e., Theorem 1) shows that the axiom of Acyclic Binary Choice, the axiom of Weak WARP, and the axioms of Expansion in the above three domains exactly capture the essence of choice by a shortlisting with limited capacity. In the Appendix A.2, we provide five examples illustrating the necessity of each of these five axioms. In other words, we provide minimal examples in which a choice function is not rationalizable by a shorting with limited capacity when four of these axioms hold and one of them fails.

The intuition of the proof is as follows. The axiom of Acyclic Binary Choice serves to identify the preference relation. The definition of threshold capacity serves to identify the DM's capacity. The axiom of Limited Weak Expansion serves to guarantee the asymmetry of elimination rationale. The axiom of Conditional Expansion serves to guarantee that the size of any choice set after elimination will fall within the DM's capacity. Finally, the axiom of Weak WARP and the axiom of Limited Expansion guarantee that the chosen option in a large choice set must be preferred over any other option in the consideration set of the choice set.

One may think that the above intuition suggests a very natural definition of preference relations and elimination relations, i.e., define preference relation from binary choices and define elimination relation from WARP-violating choice pairs. In particular, one may find it is appealing to define $x \succ y$ if $x = c(\{x, y\})$ and define xPy if $y = c(\{x, y\})$ but $x = c(S)$ for some $S \supset \{x, y\}$. It turns out that the above definition of preference relations applies but the above definition of elimination relations is neither necessary nor sufficient. We present two choice functions below to highlight the incompetence of such natural definition of elimination relation.

Example 1. *Consider $X = \{x_1, x_2, x_3, x_4\}$ and two choice functions c_i ($i = 1, 2$) satisfying that $c_i(\{x_1, x_2\}) = c_i(\{x_1, x_3\}) = c_i(\{x_1, x_4\}) = x_1$, $c_i(\{x_2, x_3\}) = c_i(\{x_2, x_4\}) = x_2$, and $c_i(\{x_3, x_4\}) = x_3$. In addition, the two choice functions specify the choices from the remaining choice sets as illustrated in Table 1. For the choice function c_1, the above natural definition of elimination relation would suggest that $x_2 P x_1$ since $x_2 = c_1(\{x_1, x_2, x_3\})$. But this is inconsistent with the fact that $x_1 = c_1(\{x_1, x_2, x_4\})$. In fact, the choice function c_1 is rationalizable by a shortlisting with capacity-2, defining $x_3 P x_1$ and $x_2 P x_4$. This demonstrates that the above natural definition of elimination relation is not necessary. For the choice function c_2, the above natural definition of elimination relation would suggest the unique elimination $x_3 P x_2$. With the unique elimination, the choice function c_2 is not rationalizable by a shortlisting with limited capacity. In fact, it is rationalizable by a shortlisting with capacity-2, defining $x_3 P x_2$ and $x_1 P x_4$. This demonstrates that the above natural definition of elimination relation is not sufficient.*

choice set	$\{x_1, x_2, x_3, x_4\}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$
choice by c_1	x_2	x_2	x_1	x_3	x_2
choice by c_2	x_1	x_1	x_1	x_1	x_3

TABLE 1: Incompetence of defining elimination relation from WARP-violating choice pairs

Theorem 1 (Existence of Capacity). *A choice function is rationalizable by a shortlisting with capacity-k for some k if and only if it satisfies Axioms 1-5.*

Proof. See the proof in the Appendix A.3. □

Theorem 1 shows the sufficient and necessary conditions on a choice function for the existence of the capacity, which is treated as unobservable. We know that the capacity, if exists, must be no more than the threshold capacity, k^*, which can be recovered from choice data. Theorem 2 further establishes that the capacity is uniquely determined for a WARP-violating choice function. On the one hand, we know from the proof of Theorem 1 that a WARP-violating choice function is rationalizable by a shortlisting with capacity-k^* if it is rationalizable by a shortlisting with limited capacity. In other words, the capacity can be equal to the threshold capacity if the capacity exists. On the other hand, the capacity cannot be less than the threshold capacity. The reason is that if so, for any choice set whose size falls between the capacity and the threshold capacity, the elimination machine is operating so that a WARP-violating choice pair occurs. In other words, there exists a WARP-violating choice set whose size is no more than the threshold capacity, which is impossible according to the definition of threshold capacity. This intuition suggests that the capacity must be equal to the threshold capacity.

Theorem 2 (**Uniqueness of Capacity**). *If a choice function violates the WARP but is rationalizable by a shortlisting with capacity-k, then k is unique and $k = k^*$.*

Proof. Suppose a choice function violates the WARP, i.e. there exists a WARP-violating feasible set S such that for some $x, y \in S$, $x = c(\{x, y\})$ and $y = c(S)$. Since this choice function is rationalizable by a shortlisting with capacity-k, then $2 \leq k \leq k^*$. In the proof of Theorem 1, it is shown that capacity-k^* can rationalize the choice. If $k^* = 2$, then trivially it is only rationalizable by a shortlisting with capacity-2.

Consider now $k^* > 2$ and assume that S is a WARP-violating choice set. We know that S has one "best" option, i.e., $z = c(\{z, w\})$ for any $w \in S$, since the choice function is rationalizable by a shortlisting with capacity-k. The WARP violation in S implies that the "best" option in S is not chosen in S, i.e., $z \neq c(S)$. This suggests that z has to be eliminated by some option in S, say t, once the elimination procedure is triggered. On the other hand, since the choice function is rationalizable by a shortlisting with capacity-k^*, $z = c(T)$ for any $T \subset S$ such that $z \in T$ and $|T| \leq k^*$.

Suppose by contradiction that capacity k' also rationalizes the choice function for some $2 \leq k' < k^*$. Then the elimination procedure is triggered when the size of a choice set exceeds k', which implies that $z \neq c(T)$, where $\{z, t\} \subset T \subset S$ and $k' < |T| \leq k^*$. This is a contradiction. Therefore, the choice function is only rationalizable by a shortlisting with capacity-k^*. □

The existence and uniqueness of capacity established in Theorems 1 and 2 suggest that we can recover the unobservable capacity from a DM's choice data if the DM's choice function is rationalizable by a shortlisting with limited capacity. Specifically, we identify the threshold capacity from choice data with the size of the smallest WARP-violating choice set less than one, and then identify the DM's capacity with the threshold capacity. In addition, the uniqueness of capacity suggests an important difference between our limited capacity model and any other two-stage choice model. It is well known that existing two-stage choice models inherently admit multiplicity of representations, even for the most interesting cases in which a choice function does not satisfy the WARP. In our limited capacity model, a representation is governed by a capacity number, a rationale, and a preference. As we show in Theorem 2 and Subsection 2.1, the representation of a WARP-violating choice function is unique in terms of capacity number and preference, while multiplicity of rationales may still exist.

Here, we commit to a specific procedure of forming consideration set motivated by the literature. An alternative approach is to impose conditions on the consideration sets without committing to a procedure (e.g., Masatlioglu et al. 2012, Lleras et al. 2017). The marketing literature has devoted considerable efforts to understanding the formation of consideration sets and has developed tools to observe consideration sets. To highlight the link between the alternative modeling approach and our modeling approach, we investigate the consideration functions that are consistent with the shortlisting with limited capacity procedure.

Now, consider a situation in which the DM splits the feasible options into two sets and considers the alternatives in each set separately. For example, an online shopper instead of having all the results in one tab, she may open two tabs and split the results in these tabs. Then the DM forms her consideration set from each subset. If the DM is overwhelmed by the number of alternatives, then this may allow her to consider more options:

More is Less: $\Gamma(S \cup T) \subseteq \Gamma(S) \cup \Gamma(T)$ for any $S, T \in \mathcal{X}$

The More is Less property may be violated by the *attention filter* that is defined in Masatlioglu et al. (2012). The consideration filter requires a consideration function Γ satisfying that if $x \notin \Gamma(S)$, then $\Gamma(S) = \Gamma(S \setminus x)$. In other words, if the DM does not consider an option, then her consideration set does not change if that option becomes unavailable. However, it does not put any restriction on the consideration set when a considered option becomes unavailable. Hence, if the DM whose consideration function satisfies the attention filter considers all of the options in a bigger set and considers only one option in all the subsets, then the consideration function violates the More is Less property.

On the other hand, this More is Less property is consistent with the *competition filter* that is defined in Lleras et al. (2017). The competition filter requires a consideration function Γ satisfying that if $x \in \Gamma(T)$ and $x \in S \subset T$, then $x \in \Gamma(S)$. According to the competition filter, in the sense that the options compete to be considered, and as the feasible set gets larger, it gets harder to be considered. Hence, if an option is considered in a set, it should be considered in any smaller sets that include that option.[4] Indeed, the

[4]The competition filter is the requirement of Sen's α property on the consideration sets.

competition filter is equivalent to the More is Less property:

Lemma 1. *A consideration function* Γ *is a competition filter if and only if it satisfies the More is Less property.*

Proof. Suppose that Γ is a competition filter. If $x \in \Gamma(S \cup T)$, then $x \in S \cup T$. Without loss of generality, assume that $x \in S$. Then $x \in \Gamma(S)$ according to the definition of competition filter. In other words, $\Gamma(S \cup T) \subseteq \Gamma(S) \cup \Gamma(T)$. Now suppose that $\Gamma(S \cup T) \subseteq \Gamma(S) \cup \Gamma(T)$ for any $S, T \in \mathcal{X}$. Assume that $x \in S \subset T$ and $x \in \Gamma(T)$. Then $x \in \Gamma(S \cup (T/S)) \subseteq \Gamma(S) \cup \Gamma(T/S)$. Since $x \notin T/S$ and then $x \notin \Gamma(T/S)$, $x \in \Gamma(S)$. In other words, Γ is a competition filter. $\square$

The competition filter provides a structure as the sets get smaller. However, it is quite flexible as the sets get bigger. We introduce below two properties that provide a structure as the sets get bigger. We firstly define two classes: the class of feasible sets with full consideration $\mathcal{FC}$ and the class of feasible sets with limited consideration $\mathcal{LC}$

$$\mathcal{FC} = \{S \in \mathcal{X} : \Gamma(S) = S\},$$
$$\mathcal{LC} = \{S \in \mathcal{X} : \Gamma(S) \neq S\}.$$

All of the options of a feasible set in $\mathcal{FC}$ wins the competition, or we may say all of them are considered without competing to attract the attention. However, for the options of a feasible set in $\mathcal{LC}$, they strive to be considered and there are winners and losers. In this sense, the competition in $\mathcal{LC}$ is serious. Given that the competition filter is based on the idea that the products are in a competition to get consumers' attention, we introduce a property that is especially appealing under the competition interpretation: If an option is able to get into the consideration set in two serious competitions, it also belongs to the consideration set when the two competitions are combined:[5]

Weak Consideration Dominance: $\Gamma(S) \cap \Gamma(T) \subseteq \Gamma(S \cup T)$ for any $S, T \in \mathcal{LC}$.

In this setup, if an option wins a serious competition in one set and another option wins a serious competition in another set, there should be a set that includes both options such that at least one of them are considered, i.e., they do not mutually exclude each other from being considered.

[5]The weak consideration dominance property is the requirement of Sen's γ property on the consideration sets.

No Mutual Exclusion: If $x \in \Gamma(S)$ and $y \in \Gamma(S')$ for some $S, S' \in \mathcal{LC}$, then there exists $T \in \mathcal{LC}$ satisfying that $\{x, y\} \subset T$ and $\{x, y\} \cap \Gamma(T) \neq \emptyset$.

Finally, the number of alternatives in a choice set is important for their competing to be considered. If the formation of consideration set is triggered by the abundance of alternatives, then the feasible sets in which some options are not considered should be more crowded than the feasible sets in which all the options are considered:

Separability: $max\{|S|, S \in \mathcal{FC}\} < min\{|T|, T \in \mathcal{LC}\}$.

Theorem 3 shows that when a consideration function satisfies the above four properties, it is equivalent to a shortlisting with capacity-k. In addition, since the property of competition filter is the essential property of the limited consideration model in Lleras et al. (2017), Theorem 3 and Lemma 1 imply that the model of shortlisting with limited capacity is a refinement of the limited consideration model.

Theorem 3. *A consideration function $\Gamma : \mathcal{X} \mapsto \mathcal{X}$ can be represented as a shortlisting with capacity-k for some k if and only if it satisfies the properties of More is Less, Weak Consideration Dominance, No Mutual Exclusion, and Separability.*

Proof. See the proof in the Appendix A.4. □

2.1 Revealed Preference and Revealed Shortlist

In the fully rational choice model, we say a DM has a revealed preference of x over y if $x = c(\{x, y\})$. We reexamine the appropriateness of the definition in the setup of choice by shortlisting with limited capacity.

It is of interest to observe that this type of definition may still apply for a WARP-violating choice function. The reason is that for a WARP-violating choice function, if it is rationalizable by a shortlisting with limited capacity, the capacity must be at least two (actually ranging from 2 to $N - 1$). Hence, the DM always chooses the option she prefers in any binary choice problem. In other words, it is reasonable to expect that she has a revealed preference of x over y if $x = c(\{x, y\})$.

However, this type of definition fails when the WARP holds. In this case, the choice function is rationalizable by a shortlisting with any capacity ranging from 1 to N. If choices are indeed made according to a shortlisting with capacity-1, then x being chosen over y reveals that x eliminates y. If choices are indeed made according to a shortlisting with a capacity higher than one, then x being chosen over y reveals that x is preferred over y. So we can not conclude revealed preference or revealed elimination from the observation that x is chosen over y. In this sense, we can not differentiate choice behavior that is generated by a DM who makes a choice decision according to her "real" preference from choice behavior that is generated by a DM who makes a choice decision according to a complete elimination rationale, independent of his "real" preference.

In fact, when the WARP holds, what we can infer from choice behavior is that either the chosen option in a binary choice problem always eliminates the unchosen option, or the chosen option in a binary choice problem is always preferred over the unchosen option. For example, when observing $x = c(\{x, y\})$ and $z = c(\{z, w\})$, we can conclude that either x is preferred over y and z is preferred over w, or x eliminates y and z eliminates w, as long as the WARP holds.

We now look at the revealed shortlist of any choice set. If the WARP holds for a choice function, the revealed shortlist of a choice set is trivial and can be a subset of it with any size depending on the capacity. If a WARP-violating choice function is rationalizable by a shortlisting with limited capacity, the capacity number must be equal to the threshold capacity k^* according to Theorem 2. In this case, the revealed shortlist of a choice set with a size less than k^* is still trivial and is the same as the choice set. Nevertheless, the revealed shortlist of a choice set with a size exceeding k^* may not be uniquely pinned down due to the multiplicity of elimination rationale. For example, consider $X = \{x, y, z\}$ and a choice function satisfying that $c(\{x, y\}) = c(\{x, z\}) = x$, $c(\{y, z\}) = y$, and $c(X) = y$. The choice function is rationalizable by a shortlisting with capacity-2. The elimination relation may be yPx and then the shortlist of X is $\{y, z\}$. It may also be yPx and yPz so that the shortlist of X is $\{y\}$.

While it is impossible to determine the revealed shortlist of a choice set with a size exceeding k^* for a WARP-violating choice function, it is possible to determine the core of revealed shortlist of the choice set, regardless of the specific elimination rationale. Specifically, we know that if an option is ever chosen over another option in a choice set with a size exceeding the threshold capacity, it should never be eliminated by the second option in any case. Therefore, a natural definition of the core of the revealed shortlist of a choice set is that any two options in the core should never be eliminated by each other in the above sense.

Definition 6. *Assume that c is a WARP-violating choice function on X and $S \subseteq X$ is a choice set with a size exceeding the threshold capacity k^*. The* ***core of the revealed shortlist*** *of choice set S is $\Gamma_c(S) = \{x \in S :$ for any $y \in S$, there exists a certain $T \supset \{x, y\}$ with $|T| > k^*$ such that $x = c(T)\}$.*

Since $c(S) \in \Gamma_c(S)$, the core of the revealed shortlist of choice set S includes at least one option. In addition, it is easy to see that the core of the revealed shortlist of a choice set may include multiple options. For example, if $x = c(S)$ and $y = c(S')$, where $\{x, y\} \subset S \cap S'$ and $|S|, |S'| > k^*$, then $\{x, y\} \subset \Gamma_c(S)$.

We show that the core of the revealed shortlist of a choice set is exactly the common part of shortlists of this choice set under all possible elimination rationales.

Theorem 4 (Core of Revealed Shortlist). *Assume that c is a WARP-violating choice function on X and $S \subseteq X$ is a choice set with a size exceeding the threshold capacity k^*. If $(k^*, P_1, \succ)$, $(k^*, P_2, \succ)$, $\cdots$, and $(k^*, P_n, \succ)$ are all possible combinations of capacity, rationale and preference relation that rationalize the choice function, it must be that $\Gamma_c(S) = \cap_{i=1}^{n} \Gamma_{k^*}^{P_i}(S)$.*

Proof. Assume that $x \in \Gamma_c(S)$. Then for any $y \in S$, there exists a certain $T \supset \{x, y\}$ with $|T| > k^*$ such that $x = c(T)$. So yP_ix cannot occur for any i. In other words, $x \in \Gamma_{k^*}^{P_i}(S)$ for any i and then $x \in \cap_{i=1}^{n} \Gamma_{k^*}^{P_i}(S)$.

Now assume that $x \in \cap_{i=1}^{n} \Gamma_{k^*}^{P_i}(S)$. Suppose by contradiction that $x \notin \Gamma_c(S)$. Then there exists a certain $y \in S$ such that for any $T \supset \{x, y\}$ with $|T| > k^*$, $x \neq c(T)$. Consider an elimination rationale P by augmenting P_1 by additionally defining yPx. The augment of P_1 in this way will not generate inconsistency with the choice function because x is never chosen in a choice set that exceeds the capacity and includes y. Therefore, the choice function is rationalizable by $(k^*, P, \succ)$. In other words, $(k^*, P, \succ)$ is one of such combinations of capacity, rationale and preference relation that rationalize the choice function. The assumption that $x \in \cap_{i=1}^{n} \Gamma_{k^*}^{P_i}(S)$ then implies that $x \in \Gamma_{k^*}^{P}(S)$, i.e., no zPx for any $z \in S$. This contradicts the definition of elimination rationale P according to which yPx and $y \in S$. □

2.2 Related Literature

In this subsection, we provide minimal examples to highlight the difference between the limited capacity model and some of existing choice models in the literature. We first look at what can be explained and what cannot be explained by the limited capacity model when the grand choice set consists only of three options, e.g., $X = \{x, y, z\}$. Without loss of generality, we assume that a choice function is characterized by $c(\{x, y\}) = x$, $c(\{y, z\}) = y$, and the remaining parts that are determined in Table 2.

It is clear that the attraction effect is rationalizable by a shortlisting with capacity-2. For example, the real preference may be that $x \succ y \succ z$ and zPx in the first stage when elimination is necessary. It is also clear that choosing the pariwisely unchosen option is rationalizable by a shortlisting with capacity-2. For example, the real preference may be that $x \succ y \succ z$ and $zPyPx$ in the first stage when elimination is necessary. Finally, since the limited capacity model does not explain cyclical binary choice, it is falsifiable for a grand choice set consisting of three options.

choice set	$\{x, z\}$	$\{x, y, z\}$	label	explain
case 1	x	x	fully rational choice	✓
case 2	x	y	attraction effect	✓
case 3	x	z	choosing pairwisely unchosen	✓
case 4	z	x, y, or z	cyclical binary choice	✗

TABLE 2: What can/cannot be explained by the limited capacity model

Table 3 illustrates the difference between the limited capacity model and some of existing choice models in terms of explaining choice behavior when $N = 3$. We know Masatlioglu et al. (2012), Manzini and Mariotti (2012), Cherepanov et al. (2013), and Lleras et al. (2017) are not falsifiable when $N = 3$ from the characterizations of their models. The other two models are falsifiable when $N = 3$: Manzini and Mariotti (2007) rules out the attraction effect and choosing the pairwisely unchosen option since the property of Expansion requires that an option that is chosen in two choice sets must also be chosen in the union of the two choice sets, and our limited capacity model rules out cyclical binary choice due to the axiom of acyclic binary choice.

label	Masatlioglu et al. (2012), Manzini & Mariotti (2012) Cherepanov et al. (2013), Lleras et al. (2017)	Manzini & Mariotti (2007)	This paper
fully rational choice	✓	✓	✓
attraction effect	✓	✗	✓
choosing pairwisely unchosen	✓	✗	✓
cyclical binary choice	✓	✓	✗
falsifiable when $N = 3$	✗	✓	✓

TABLE 3: Comparison of different choice models

We know from Theorem 3 that the limited capacity model is a refinement of the limited consideration model in Lleras et al. (2017), which is behaviorally equivalent to both the *categorization* choice model in Manzini and Mariotti (2012) and the *rationalization* choice model in Cherepanov et al. (2013). In addition, Table 3 shows that there is no nested relationship between the limited capacity model and the *shortlisting* choice model in Manzini and Mariotti (2007), and that the choice model with limited attention in Masatlioglu et al. (2012) is not nested in the limited capacity model. Finally, we provide an example to demonstrate that the limited capacity model is not nested in the choice model with limited attention that is characterized in Masatlioglu et al. (2012) either.

Example 2. *Suppose that a choice function c on $X = \{x_1, x_2, x_3, x_4, x_5\}$ is rationalizable by a shortlisting with capacity-3. The preference relation is $x_1 \succ x_2 \succ x_3 \succ x_4 \succ x_5$ and the elimination relation is $x_5 P x_1 P x_3$ and $x_5 P x_2$ when elimination is necessary. We show that the choice function cannot be rationalized by the model with limited attention. Manzini and Mariotti (2012) shows that their model can be exactly captured by the WARP with Limited Attention, which requires that: for any nonempty set S, there exists $x^* \in S$ such that, for any T including x^*, if $c(T) \in S$ and $c(T) \neq c(T/\{x^*\})$, then $c(T) = x^*$. Therefore, we only need to show that the choice function violates the WARP with Limited Attention. Consider $S = \{x_1, x_3, x_4, x_5\}$. Then $c(S) = x_4$. We show that there is no such $x^* \in S$ satisfying the above requirement. $x_1 = c(\{x_1, x_4, x_5\}) = c(\{x_1, x_3, x_4\}) \neq x_4$ implies that $x_3 \neq x^*$ and $x_5 \neq x^*$. $x_3 = c(\{x_3, x_4, x_5\}) \neq x_4$ implies that $x_1 \neq x^*$. Finally, $c(\{x_2, x_3, x_4, x_5\}) = x_3$ but $c(\{x_2, x_3, x_5\}) = x_2$ implies that $x_4 \neq x^*$.*

3 A Special Case: Choice with Top-k Shortlisting

In this section, we investigate a special case in which a DM considers exactly a certain number of options. Specifically, when the number of options in a choice set exceeds her capacity k, the DM applies an elimination rationale to exclude some options such that the size of the consideration set is exactly k (labeled as top-k shortlisting). In addition to capturing some realistic examples, the special case has the appeal of ruling out the possible discontinuity inherent in the previous shortlisting with capacity-k: when the size of a choice set is k the elimination stage is not triggered and the DM considers k options, but once an additional option is added to the choice set, the elimination machine starts to operate and possibly renders the number of options considered less than k. As a comparison, in the special case the DM considers k options both when the size of a choice set is k and when the size exceeds k. As an illustrating example, when $N = 3$, fully rational choice, the attraction effect, and choosing pairwisely unchosen are rationalizable by a shortlisting with limited capacity, but choosing pairwisely unchosen is now ruled out in the special case in which the DM considers exactly a certain number of options.

Definition 7. *A consideration function with capacity-k, Γ_k, is called a **top-k shortlisting** if there* *exists a rationale P such that for each feasible set $S \in \mathcal{X}$:*

$$\Gamma_k^P(S) = \begin{cases} S & \text{if } |S| \leq k, \\ max(S,P) & \text{if } |S| > k, \end{cases}$$

and $|max(S,P)| = k$.

We then define a choice function to be rationalizable by a top-k shortlisting if it can be represented by a pair of top-k shortlisting and preference relation.

Definition 8. *A choice function, c, is **rationalizable by a top-k shortlisting** if there exists a top-k shortlisting, Γ_k^P, and a linear order, $\succ$, such that for any $S \in \mathcal{X}$,*

$$c(S) = max(\Gamma_k^P(S), \succ)$$

Given this definition, it is clear that if a choice function is rationalizable by a top-k shortlisting, then it is rationalizable by a shortlisting with capacity-k. In addition, a choice function is rationalizable by a top-N shortlisting if and only if it is rationalizable by a shortlisting with capacity-N. Similarly, a choice function is rationalizable by a top-1 shortlisting if and only if it is rationalizable by a shortlisting with capacity-1.

3.1 Size of Top-k Shortlist

We first observe that if a choice function is rationalizable by a top-k shortlisting for some k, the size of the top-k shortlist, i.e., the number k, can take only a few values. In fact, the size of the shortlist depends on the size of the grand set, N, where $N \geq 3$. Specifically, the size of the shortlist can only be $N, N-1$, or 1 when $N \neq 4$, and can be 4, 3, 2, or 1 when $N = 4$.

This essentially suggests that a choice function cannot be rationalized by a top-k shortlisting for $N-1 > k > 1$ when the grand choice set includes more than four options. The intuition proceeds as follows. If the size of a top-k shortlist is less than $N-1$, a top-k shortlisting of the grand choice set requires that at least two options are eliminated according to a rationale P. Then when $k > 2$, we can find a choice set with $k+1$ options including two eliminated options and the one or two options that eliminate the two options. For this choice set, the top-k shortlist according to the rationale P has a size less than k. Hence, it is impossible for a choice function to be rationalizable by a top-k shortlisting for $N-1 > k > 2$. In addition, we need $N-2$ options to be eliminated from the grand choice set if the size of a top-k shortlist is equal to two, and since $2(N-2) > N$

when $N > 4$, it is impossible that all the $2(N-2)$ options that involve elimination relation are completely distinct options. In other words, we can find a choice set including three options in which two options are eliminated by some option(s) in the set once the elimination stage is triggered. Then the top-2 shortlist of the choice set according to the rationale P has a size of only one. Hence, it is also impossible for a choice function to be rationalizable by a top-2 shortlisting when $N > 4$.

Theorem 5. *If there exists a certain k such that a choice function is rationalizable by a top-k shortlisting, then k may only be $N, N-1$, or 1 when $N \neq 4$ and may be 4, 3, 2 or 1 when $N = 4$.*

Proof. See the proof in the Appendix A.5. □

Similar to Propositions 1 and 2, we observe that a choice function that satisfies the WARP may be rationalizable by a top-k shortlisting for a few ks while it may not be rationalizable for any $k \leq N$.

Proposition 4. *If a choice function is rationalizable by a top-N shortlisting, then it is rationalizable by a top-$N-1$ shortlisting and rationalizable by a top-1 shortlisting. It is also rationalizable by a top-2 shortlisting when $N = 4$. A choice function is rationalizable by a top-N shortlisting if and only if it is rationalizable by a top-1 shortlisting.*

Proof. Note that a choice function is rationalizable by a top-N shortlisting if and only if it is rationalizable by a shortlisting with capacity-N. Similarly, a choice function is rationalizable by a top-1 shortlisting if and only if it is rationalizable by a shortlisting with capacity-1. Proposition 2 suggests that a choice function is rationalizable by a shortlisting with capacity-N if and only if it is rationalizable by a shortlisting with capacity-1. Thus, a choice function is rationalizable by a top-N shortlisting if and only if it is rationalizable by a top-1 shortlisting. The proof of the remaining claims can be found in the proof of Theorem 5 and Example 3. □

3.2 Characterization of Choice with Top-k Shortlisting

Theorem 5 suggests that when a DM applies top-k shortlisting to make a choice decision, her choice function is fully rationalizable, i.e., rationalizable by a top-N shortlisting, or is close to fully rationalizable, i.e., rationalizable by a top-$N-1$ shortlisting. Thus, her behavioral implications should be very similar to the WARP that captures the fully rational choice behavior. We provide below a representation theorem for a choice function to be rationalizable by a top-k shortlisting.

Theorem 6 (**Existence of Top-k Shortlisting**). *When $N \neq 4$, a choice function is rationalizable by a top-k shortlisting for some number k if and only if either it satisfies the WARP or it has a unique WARP-violating choice pair in which X is the WARP-violating choice set.*

Proof. We have established in the proof of Theorem 5 that if a choice function satisfies the WARP or it has a unique WARP-violating choice pair in which X is the WARP-violating choice set, then it is rationalizable by a top-k shortlisting for some number k. Thus, it sufficies to establish the "only if" part of the claim.

"only if" direction. Suppose now that a choice function is rationalizable by a top-k shortlisting. If it is rationalizable by a top-N shortlisting, then it satisfies the WARP. If it is not rationalizable by a top-N shortlisting, Theorem 5 and Proposition 4 suggest that it must be rationalizable by a top-$N-1$ shortlisting, i.e., $c(S) = max(\Gamma^P_{N-1}(S), \succ)$, where Γ^P_{N-1} is a top-$N-1$ shortlisting. Firstly, note that there exists a certain option x such that $x \succ c(X)$ because otherwise the choice function must be also rationalizable by a top-N shortlisting. In this way, we find a WARP-violating choice pair. Secondly, x is the unique option such that $x \succ c(X)$. Suppose by contradiction that $y \succ c(X)$. Then both x and y must be eliminated by some options according to P, and in turn the size of the top-$N-1$ shortlist of X is less than $N-1$, which violates the definition of top-$N-1$ shortlisting. In addition, since $c(S) = max(\Gamma^P_{N-1}(S), \succ) = max(S, \succ)$ if $S \neq X$, there is no other WARP-violating choice pair. Thus, the choice function has a unique WARP-violating choice pair in which X is the WARP-violating choice set. □

Since Theorem 6 suggests a WARP-violating choice function that is rationalizable by a top-k shortlisting has a unique WARP-violating choice set X, we know that a WARP-violating choice function must be rationalizable by a top-$N-1$ shortlisting if it is rationalizable by a top-k shortlisting for some k.

Proposition 5 (**Uniqueness of Top-k Shortlisting**). *If a choice function violates the WARP but is rationalizable by a top-k shortlisting, then $k = N-1$ when $N \neq 4$.*

The above characterization of choice with top-k shortlisting essentially says that compared to the fully rational choice model, the model of choice with top-k shortlisting only marginally increases the flexibility of rationalizing a choice function: in addition to the situation in which the "best" option is always chosen in any choice set, only one more situation in which the second "best" option is chosen in the grand choice set and the "best" option is always chosen in any other choice set is treated as rationalizable choice behavior.

When the size of the grand choice set is four, a characterization of the behavioral implications for a choice function to be rationalizable by a top-k shortlisting is lengthy.

For this reason, we list all choice patterns that are rationalizable by a top-k shortlisting for some number k.

Example 3 (**Choice with a top-k shortlisting when** $N = 4$)**.** *Consider $X = \{x_1, x_2, x_3, x_4\}$. Assume without loss of generality that $x_1 = c(\{x_1, x_2\})$, $x_2 = c(\{x_2, x_3\})$, and $x_3 = c(\{x_3, x_4\})$. Then a choice function defined on X is rationalizable by a top-k shortlisting for some k if and only if it satisfies $x_1 = c(\{x_1, x_3\}) = c(\{x_1, x_4\})$, $x_2 = c(\{x_2, x_4\})$, and the choices in the remaining choice sets fit one of the following eight patterns. One may define the preference relation*

choice pattern	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$	X	size of shortlist
pattern 1	x_1	x_1	x_1	x_2	x_1	$k = 4, 3, 2, 1$
pattern 2	x_1	x_1	x_1	x_2	x_2	$k = 3$
pattern 3	x_1	x_2	x_3	x_3	x_3	$k = 2$
pattern 4	x_2	x_1	x_3	x_3	x_3	$k = 2$
pattern 5	x_2	x_2	x_1	x_2	x_2	$k = 2$
pattern 6	x_2	x_1	x_3	x_2	x_2	$k = 2$
pattern 7	x_1	x_2	x_3	x_2	x_2	$k = 2$
pattern 8	x_1	x_1	x_1	x_3	x_1	$k = 2$

TABLE 4: Choice with a top-k shortlisting when $N = 4$

as $x_1 \succ x_2 \succ x_3 \succ x_4$. In addition, one may define elimination rationales in choice patterns 2-8, respectively: (1) $x_2 P x_1$ for choice pattern 2; (2) $x_4 P x_1$ and $x_3 P x_2$ for choice pattern 3; (3) $x_3 P x_1$ and $x_4 P x_2$ for choice pattern 4; (4) $x_2 P x_1$ and $x_3 P x_4$ for choice pattern 5; (5) $x_3 P x_1$ and $x_2 P x_4$ for choice pattern 6; (6) $x_4 P x_1$ and $x_2 P x_3$ for choice pattern 7; and (7) $x_3 P x_2$ and $x_1 P x_4$ for choice pattern 8. It is straightforward to establish that the corresponding choice functions are rationalizable by a top-k shortlisting.

Finally, we discuss revealed preference and revealed top-k shortlist in the model of choice with top-k shortlisting when $N \neq 4$. Clearly, the definition that a DM has a revealed preference x over y if $x = c(\{x, y\})$ may apply for a WARP-violating choice function but fails to apply when the WARP holds. In addition, when the WARP holds, the revealed top-k shortlist of any choice set except the grand choice set is either the singleton set of the chosen option or the choice set itself, and the revealed top-k shortlist of the grand choice set could be the singleton set of the chosen option, the grand set itself, or any $N - 1$-option subset including the chosen option. Finally, for a WARP-violating choice function, the revealed top-k shortlist of any choice set except the grand choice set is the choice set itself, and the revealed top-k shortlist of the grand choice set is the $N - 1$-option subset that excludes the "best" option that is chosen in all binary choice sets including the option.

4 Conclusion

This paper is the first one that pays specific attention to the number of available alternatives. We study a two-stage procedure where the DM has a *limited capacity* for the number of alternatives to consider: For the choice problems where the number of alternatives is within the DM's capacity, the DM considers all the alternatives and chooses the best alternative. However, when the number of alternatives exceeds the DM's capacity, by using a rationale, the DM limits the number of alternatives to consider to be within her capacity and then chooses the best one.

We provide two characterizations based on (i) choice data and (ii) consideration sets. As in any theoretical model based on choices, the behavioral axioms can be tested by using choice data. While it is easy to acquire choice data, the entire choice data may not be available. Recently, the emerging new tools, such as eye tracking, make it possible to observe consideration sets. Hence, testing the properties of consideration sets (e.g., the property of Separability) will allow one to determine the capacity of a DM in situations where the entire choice data is not available. In addition, in the environments in which both choice data and consideration set data are available, our characterizations will allow one to investigate to what extent the conclusions based on choice data are valid when non-choice data are taken into account. Specifically, one can distinguish different models based on the consideration set data when those models are consistent with the observed choice behavior. Additionally, one can check whether the DM is choosing the best alternative from her consideration set. Even when the choice is consistent with the WARP, we can deduce whether the DM has a capacity by looking at the consideration set data.

References

Alba, J., J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, and A. Sawyer (1997): "Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces," *Journal of Marketing*, 61, 38–53.

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

Demuynck, T. and C. Seel (2018): "Revealed Preference with Limited Consideration," *American Economic Journal: Microeconomics*, 10, 102–131.

Dulleck, U., R. Kerschbamer, and M. Sutter (2011): "The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition," *American Economic Review*, 101, 526–555.

Ergin, H. and T. Sarver (2010): "A Unique Costly Contemplation Representation," *Econometrica*, 78, 1285–1339.

Geng, S. (2016): "Decision Time, Consideration Time, and Status Quo Bias," *Economic Inquiry*, 54, 433–449.

Hauser, J. R. and B. Wernerfelt (1990): "An Evaluation Cost Model of Consideration Sets," *The Journal of Consumer Research*, 16, 393–408.

Huberman, G. and T. Regev (2001): "Contagious Speculation and a Cure for Cancer: A Non-Event that Made Stock Prices Soar," *Journal of Finance*, 56, 387–396.

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

Manzini, P. and M. Mariotti (2007): "Sequentially Rationalizable Choice," *American Economic Review*, 97, 1824–1839.

——— (2012): "Categorize then Choose: Boundedly Rational Choice and Welfare," *Journal of the European Economic Association*, 10, 1141–1165.

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

Miller, G. A. (1956): "The Magical Number Seven, Plus or Minus two: Some Limits on our Capacity for Processing Information," *Psychological Review*, 63, 81–97.

Ortoleva, P. (2013): "The Price of Flexibility: Towards a Theory of Thinking Aversion," *Journal of Economic Theory*, 148, 903–934.

Reutskaja, E., R. Nagel, C. F. Camerer, and A. Rangel (2011): "Search Dynamics in Consumer Choice under Time Pressure: An Eye-tracking Study," *American Economic Review*, 101, 900–926.

Richards, M. D., J. E. Sheridan, and J. W. Slocum (1975): "Comparative Analysis of Expectancy and Heuristic Models of Decision Behavior," *Journal of Applied Psychology*, 60, 361–368.

Rosen, D. R., J. M. Curran, and T. B. Greenlee (1998): "College Choice in a Brand Elimination Framework: The Administrator's Perspective," *Journal of Marketing for Higher Education*, 8.

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

Schwartz, B. (2005): *The Paradox of Choice : Why More Is Less*, Harper Perennial.

Stigler, G. J. (1961): "The Economics of Information," *Journal of Political Economy*, 69, 213–225.

A Appendix

A.1 Proof of Proposition 3

Proof. Consider a choice function, c, on $X = \{x_1, \cdots, x_N\}$. Assume that it satisfies the axiom of Acyclic Binary Choice and the property of Expansion. We first show that the axiom of Acyclic Binary Choice implies that for any $S \subseteq X$, there must exist a certain $x \in S$ such that $x = c(\{x,y\})$ for any $y \in S$. Suppose by contradiction that there exists a certain $S \subseteq X$ such that for any $x \in S$, $z = c(\{x,z\})$ for some $z \in S$. Since $|S|$ is finite, there must exist some options from S such that they are chosen cyclically, e.g., $x_1 = c(\{x_1,x_2\}), x_2 = c(\{x_2,x_3\}), \cdots, x_{n-1} = c(\{x_{n-1},x_n\})$, and $x_n = c(\{x_1,x_n\})$. The axiom of Acyclic Binary Choice then implies that $x_1 = c(\{x_1,x_n\})$, which contradicts $x_n = c(\{x_1,x_n\})$. Hence, for any $S \subseteq X$, there must exist a certain $x \in S$ such that $x = c(\{x,y\})$ for any $y \in S$. The property of Expansion then implies that $x = c(S)$ if and only if $x = c(\{x,y\})$ for any $y \in S$. Therefore, $x = c(\{x,y\})$ implies that $y \neq c(S)$ for any $S \supset \{x,y\}$, i.e., the WARP holds.

Now assume that the WARP holds. The axiom of Acyclic Binary Choice must be satisfied because otherwise no option can be chosen in the union set of these binary choice sets. We then show that the property of Expansion must hold. Assume that $x = c(S) = c(T)$. Suppose by contradiction that $x \neq c(S \cup T)$, i.e., $y = c(S \cup T)$. The WARP is violated regardless of whether $c(\{x,y\}) = x$ or $c(\{x,y\}) = y$. Hence, the property of Expansion must be satisfied. □

A.2 Illustration: Examples on the Necessity of Axioms 1-5

Example 4 (Necessity of Axiom 1). *Consider a choice function on $X = \{x_1,x_2,x_3\}$: $c(\{x_1,x_2\}) = x_1, c(\{x_2,x_3\}) = x_2, c(\{x_1,x_3\}) = x_3$, and $x_1 = c(X)$. The choice function satisfies Axioms 2-5 trivially and violates Axiom 1. The choice function is not rationalizable by a shortlisting with limited capacity.*

Example 5 (Necessity of Axiom 2). *Consider a choice function on $X = \{x_1,x_2,x_3,x_4\}$: $c(\{x_1,x_i\}) = x_1$ for $2 \leq i \leq 4$, $c(\{x_2,x_3\}) = c(\{x_2,x_4\}) = x_2$, and $c(\{x_3,x_4\}) = x_3$. The choices of the remaining choice sets are specified in Table 5.*

choice set	X	$\{x_1,x_2,x_3\}$	$\{x_1,x_2,x_4\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_3,x_4\}$
choice	x_1	x_3	x_1	x_1	x_2

Table 5: Necessity of Axiom 2

The threshold capacity of the choice function is $k^* = 2$. It is straightforward to establish that c satisfies Axiom 1, Axiom 3, and Axiom 4. Since $\{x_1, x_3\}$ and $\{x_2, x_3\}$ are the two choice sets satisfying the condition of Axiom 5, Axiom 5 is trivially satisfied. Axiom 2 is violated because $x_1 = c(\{x_1, x_3\}) = c(X)$ but $x_3 = c(\{x_1, x_2, x_3\})$.

Finally, it is straightforward to establish that c is not rationalizable by a shortlisting with capacity-4, capacity-3, or capacity-1. It is not rationalizable by a shortlisting with capacity-2 because if so, $x_3 = c(\{x_1, x_2, x_3\})$ implies that $x_1 \neq c(X)$, which contradicts the assumption.

Example 6 (Necessity of Axiom 3). *Consider a choice function on $X = \{x_1, x_2, x_3, x_4\}$: $c(\{x_1, x_i\}) = x_1$ for $2 \leq i \leq 4$, $c(\{x_2, x_3\}) = c(\{x_2, x_4\}) = x_2$, and $c(\{x_3, x_4\}) = x_3$. The choices of the remaining choice sets are specified in Table 6.*

choice set	X	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$
choice	x_3	x_2	x_2	x_3	x_2

Table 6: Necessity of Axiom 3

The threshold capacity of the choice function is $k^* = 2$. It is straightforward to establish that c satisfies Axiom 1, Axiom 2, and Axiom 4. Since $\{x_2, x_3\}$ is the only choice set satisfying the condition of Axiom 5, Axiom 5 is trivially satisfied. Axiom 3 is violated because $x_2 = c(\{x_1, x_2, x_3\}) = c(\{x_1, x_2, x_4\})$ but $x_3 = c(X)$.

Finally, it is straightforward to establish that c is not rationalizable by a shortlisting with capacity-4, capacity-3, or capacity-1. It is not rationalizable by a shortlisting with capacity-2 because otherwise $x_2 = c(\{x_1, x_2, x_3\}) = c(\{x_1, x_2, x_4\})$ implies that none of x_1, x_2, and x_3 eliminates x_2 when the elimination is triggered, and then $x_3 \neq c(X)$ since $x_2 = c(\{x_2, x_3\})$.

Example 7 (Necessity of Axiom 4). *Consider a choice function on $X = \{x_1, x_2, x_3, x_4\}$: $c(\{x_1, x_i\}) = x_1$ for $2 \leq i \leq 4$, $c(\{x_2, x_3\}) = c(\{x_2, x_4\}) = x_2$, and $c(\{x_3, x_4\}) = x_3$. The choices of the remaining choice sets are specified in Table 7.*

choice set	X	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$
choice	x_3	x_3	x_4	x_1	x_2

Table 7: Necessity of Axiom 4

The threshold capacity of the choice function is $k^* = 2$. It is straightforward to establish that c satisfies Axioms 1-3. Since $\{x_1, x_3\}$, $\{x_1, x_4\}$, $\{x_2, x_3\}$, and $\{x_2, x_4\}$ are the choice sets satisfying the condition of Axiom 5, Axiom 5 is trivially satisfied. Axiom 4 is violated because $x_1 = c(\{x_1, x_3, x_4\})$ and $x_2 = c(\{x_2, x_3, x_4\})$ but $x_3 = c(X)$.

In addition, it is straightforward to establish that c is not rationalizable by a shortlisting with capacity-4, capacity-3, or capacity-1. Finally, we show that it is not rationalizable by a shortlisting with capacity-2. Suppose by contradiction that it is rationalizable by a shortlisting with capacity-2, and the elimination relation is P and the preference relation is $\succ$. Since $x_1 = c(\{x_1, x_2\})$, $x_1 \succ x_2$. If $x_1 P x_2$, then $x_2 P x_1$ cannot occur and $x_1 = c(\{x_1, x_3, x_4\})$ implies that $x_1 = c(X)$, which is a contradiction. If $x_2 P x_1$, then $x_1 P x_2$ cannot occur and $x_2 = c(\{x_2, x_3, x_4\})$ implies that $x_2 = c(X)$, which is a contradiction. If neither $x_1 P x_2$ nor $x_2 P x_1$, then $x_1 = c(\{x_1, x_3, x_4\})$ and $x_1 \succ x_2$ imply that $x_1 = c(X)$, which is a contradiction.

Example 8 (Necessity of Axiom 5). *Consider a choice function on $X = \{x_1, x_2, x_3, x_4, x_5\}$: $c(\{x_1, x_i\}) = x_1$ for $2 \leq i \leq 5$, $c(\{x_2, x_i\}) = x_2$ for $3 \leq i \leq 5$, $c(\{x_3, x_i\}) = x_3$ for $4 \leq i \leq 5$, and $c(\{x_4, x_5\}) = x_4$. The choices of the remaining choice sets are specified in Table 8.*

choice set	choice	choice set	choice	choice set	choice
X	x_5	$\{x_1, x_2, x_3\}$	x_1	$\{x_2, x_3, x_4\}$	x_4
$\{x_1, x_2, x_3, x_4\}$	x_1	$\{x_1, x_2, x_4\}$	x_1	$\{x_2, x_3, x_5\}$	x_2
$\{x_1, x_2, x_3, x_5\}$	x_5	$\{x_1, x_3, x_4\}$	x_1	$\{x_2, x_4, x_5\}$	x_4
$\{x_1, x_2, x_4, x_5\}$	x_5	$\{x_1, x_2, x_5\}$	x_5	$\{x_3, x_4, x_5\}$	x_3
$\{x_1, x_3, x_4, x_5\}$	x_3	$\{x_1, x_3, x_5\}$	x_3		
$\{x_2, x_3, x_4, x_5\}$	x_4	$\{x_1, x_4, x_5\}$	x_5		

Table 8: Necessity of Axiom 5

The threshold capacity of the choice function is $k^* = 2$. Axiom 1 is trivially satisfied. Axiom 2 is satisfied because none of x_2, x_3, x_4 is chosen in any three-option subset of $\{x_1, x_2, x_3, x_4\}$ that includes $\{x_1\}$, neither x_4 nor x_5 is chosen in any three-option subset of $\{x_1, x_3, x_4, x_5\}$ that includes $\{x_3\}$, and x_5 is not chosen in any three-option subset of $\{x_2, x_3, x_4, x_5\}$ that includes $\{x_4\}$.

Axiom 3 is satisfied because the property of Expansion holds for choice sets whose size exceeds k^*. Specifically, $c(\{x_1, x_2, x_3\}) = c(\{x_1, x_2, x_4\}) = c(\{x_1, x_3, x_4\}) = c(\{x_1, x_2, x_3, x_4\}) = x_1$, $c(\{x_1, x_2, x_5\}) = c(\{x_1, x_4, x_5\}) = c(\{x_1, x_2, x_4, x_5\}) = x_5$, $c(\{x_1, x_3, x_5\}) = c(\{x_3, x_4, x_5\}) = c(\{x_1, x_3, x_4, x_5\}) = x_3$, $c(\{x_2, x_3, x_4\}) = c(\{x_2, x_4, x_5\}) = c(\{x_2, x_3, x_4, x_5\}) = x_4$, and $c(\{x_1, x_2, x_3, x_5\}) = c(\{x_1, x_2, x_4, x_5\}) = c(X) = x_5$.

Axiom 4 requires that for any two choice sets whose size exceeds k^*, either the chosen option in the first choice set must still be chosen once the chosen option in the second choice set is added or the chosen option in the second choice set must still be chosen once the chosen option in the first choice set is added. If the chosen option in one choice set is in another choice set, the above requirement is trivially satisfied. Therefore, we only need to check that for any two choice sets whose size exceeds k^* and in which neither of the two chosen options is in another *choice set, the above requirement is satisfied. In fact, we only need to check the following eight cases. They are: (1) $c(\{x_1, x_2, x_3\}) = x_1$ and $c(\{x_2, x_3, x_4\}) = x_4$; (2) $c(\{x_1, x_2, x_3\}) = x_1$ and $c(\{x_2, x_4, x_5\}) = x_4$; (3) $c(\{x_1, x_2, x_4\}) = x_1$ and $c(\{x_3, x_4, x_5\}) = x_3$; (4) $c(\{x_1, x_3, x_4\}) = x_1$ and $c(\{x_2, x_3, x_5\}) = x_2$; (5) $c(\{x_1, x_2, x_5\}) = x_5$ and $c(\{x_2, x_3, x_4\}) = x_4$; (6) $c(\{x_1, x_3, x_5\}) = x_3$ and $c(\{x_2, x_4, x_5\}) = x_4$; (7) $c(\{x_1, x_2, x_3, x_5\}) = x_5$ and $c(\{x_2, x_3, x_4\}) = x_4$; and (8) $c(\{x_2, x_3, x_4, x_5\}) = x_4$ and $c(\{x_1, x_2, x_3\}) = x_1$. The requirement is satisfied for the first four cases and case (8) since $c(\{x_1, x_2, x_3, x_4\}) = x_1$. The requirement is satisfied for cases (5) and (6) since $c(\{x_2, x_3, x_4, x_5\}) = x_4$. The requirement is satisfied for case (7) since $c(X) = x_5$.*

Nevertheless, Axiom 5 is violated. The reason is as follows. It can be shown that $\{x_3, x_4, x_5\}$ is a choice set satisfying the condition of Axiom 5. $x_1 = c(\{x_1, x_2\}) = c(\{x_1, x_5\})$ but $x_1 \neq c(\{x_1, x_2, x_5\})$. Therefore, the axiom of conditional Expansion fails.

In addition, it is straightforward to establish that c is not rationalizable by a shortlisting with capacity-5, capacity-4, capacity-3, or capacity-1. Finally, we show that it is not rationalizable by a shortlisting with capacity-2. Suppose by contradiction that it is rationalizable by a shortlisting with capacity-2 in the sense that $c(S) = max(\Gamma_2^P, \succ)$. It can be shown that there is no P relation among x_3, x_4, and x_5 because any of them is ever chosen in a choice set exceeding the capacity and including another of them. Then $|\Gamma_2^P(\{x_3, x_4, x_5\})| = 3$ is greater than the capacity, which is a contradiction.

A.3 Proof of Theorem 1

Proof. Consider a choice function, c, on X. If c satisfies the WARP, then its threshold capacity is $k^* = N$ and c is always rationalizable by a shortlisting with capacity-N. In addition, Axioms 1-2 are trivially satisfied and Axioms 3-5 are void. In this case, it is straightforward to establish the claim that c is rationalizable by a shortlisting with capacity-k for some k if and only if it satisfies Axioms 1-5. Therefore, we only need to show that when c is a WARP-violating choice function, it is rationalizable by a shortlisting with capacity-k for some k if and only if it satisfies Axioms 1-5. By definition, its threshold capacity k^* may range from 2 to $N - 1$.

"If" direction. Assume that c is a WARP-violating choice function that satisfies Axioms 1-5.

Define $x \succ y$ if $x = c(\{x, y\})$. Define xPy if (1) $x = c(S)$ for some $S \supset \{x, y\}$ with $|S| > k^*$ but $y \neq c(S')$ for any $S' \supset \{x, y\}$ with $|S'| > k^*$, or (2) $y = c(T)$ for some T with $|T| > k^*$ but $y \neq c(T' \cup \{x\})$ for any T' satisfying that $|T'| > k^*$ and $y = c(T')$. Define $\Gamma_{k^*}^P(S) = S$ if $|S| \leq k^*$, and $= max(S, P)$ if $|S| > k^*$. We show that c is rationalizable by $(k^*, P, \succ)$ in the sense that $c(S) = max(\Gamma_{k^*}^P(S), \succ)$.

Firstly, it is obvious that $\succ$ is complete and asymmetric. Axiom 1 guarantees that $\succ$ is transitive. Therefore, $\succ$ is a linear order.

Secondly, we show that P is asymmetric. Assume that xPy, then either situation 1 of defining xPy applies or situation 2 of defining xPy applies. If situation 1 of defining xPy occurs, then $x = c(S)$ for some $S \supset \{x, y\}$ with $|S| > k^*$ but $y \neq c(S')$ for any $S' \supset \{x, y\}$ with $|S'| > k^*$. In this situation, it is clear that neither of the two situations of defining yPx can occur. If situation 2 of defining xPy occurs, then $y = c(T)$ for some T with $|T| > k^*$ but $y \neq c(T' \cup \{x\})$ for any T' satisfying that $|T'| > k^*$ and $y = c(T')$. In this situation, it is clear that situation 1 of defining yPx cannot occur. Then the only situation that makes yPx happen is that $x = c(S)$ for some S with $|S| > k^*$ but $x \neq c(S' \cup \{y\})$ for any S' satisfying that $|S'| > k^*$ and $x = c(S')$. This implies that $c(T) \neq c(T \cup c(S))$ and $c(S) \neq c(S \cup c(T))$, which violates Axiom 4. Hence, xPy implies that yPx cannot occur, i.e., P is asymmetric.

Thirdly, we show that $|\Gamma^P_{k^*}(S)| \leq k^*$. We only need to show that $|max(S, P)| \leq k^*$ for $|S| > k^*$. Suppose by contradiction there exists a certain S such that $|max(S, P)| > k^*$. Without loss of generality, we assume that $|max(S, P)| = k^* + 1$ and label it as $S_{NP} = \{x_1, \cdots, x_{k^*}, x_{k^*+1}\}$. Assume that $x_1 = c(S_{NP})$. Since $|S_{NP}| > k^*$ and there is no P relation between x_1 and x_i for any $i \in \{2, \cdots, k^* + 1\}$, there must exist a certain $S_i \supset \{x_1, x_i\}$ satisfying that $|S_i| > k^*$ and $x_i = c(S_i)$. Since $x_i = c(S_i)$ for some S_i with $|S_i| > k^*$, then to make situation 2 of defining $x_j P x_i$ does not apply, it must be that among all $S^{'}_i$s satisfying that $|S^{'}_i| > k^*$ and $x_i = c(S^{'}_i)$, there exists a certain $S^{'}_i$ such that $x_i = c(S^{'}_i \cup x_j)$, where $j \in \{1, \cdots, k^* + 1\}$. This implies that $x_i = c(S)$ for some $S \supset \{x_i, x_j\}$ with $|S| > k^*$. To make situation 1 of defining $x_i P x_j$ does not apply, there must exist a certain $S' \supset \{x_i, x_j\}$ satisfying that $|S'| > k^*$ and $x_j = c(S')$. In this way, we essentially construct a set $T = S_{NP}$ that satisfies the condition of Axiom 5. Axiom 5 then implies that if $x = c(S_1) = c(S_2)$ and $|S_1 \cup S_2| \leq k^* + 1$, then $x = c(S_1 \cup S_2)$. Proposition 3 and Axiom 1 then imply that the WARP holds for choice sets with a size no more than $k^* + 1$. But the definition of the threshold capacity k^* requires that $x = \{x, y\}$ and $y = c(S)$ for some $S \supset \{x, y\}$ with $|S| = k^* + 1$, i.e., the WARP fails for S. This is a contradiction. Therefore, the claim that $|max(S, P)| \leq k^*$ for $|S| > k^*$ is established.

We finally show that if $x = c(S)$, then it must be that $x = max(\Gamma^P_{k^*}(S), \succ)$. Assume that $x = c(S)$. When $|S| \leq k^*$, $max(\Gamma^P_{k^*}(S), \succ) = max(S, \succ)$. The definition of k^* suggests that S is not a WARP-violating choice set and then $x = c(\{x, y\})$ for any $y \in S$. This implies that $x \succ y$ for any $y \in S$, i.e., $x = max(S, \succ)$. When $|S| > k^*$, $x = c(S)$ implies that no $y \in S$ such that yPx according to the definition of P. In other words, $x \in \Gamma^P_{k^*}(S)$. We then only need to show that if there is another option $y \in \Gamma^P_{k^*}(S)$ then it must be that $x \succ y$.

Assume that $y \in \Gamma^P_{k^*}(S)$. Suppose by contradiction that $y \succ x$, i.e., $y = c(\{x, y\})$. We show that $y = c(\{x, y\})$ and $x = c(S)$ imply that there exists a certain $z \in S$ ($z \neq y$) such that for any T satisfying that $|T| > k^*$ and $T \supset \{y, z\}$, $y \neq c(T)$. Suppose by contradiction that the implication is not true. Then for any $z \in S$ ($z \neq y$) there exists a certain T_z satisfying that $|T_z| > k^*$ and $T_z \supset \{y, z\}$, such that $y = c(T_z)$. Axiom 3 then implies that $y = c(\cup_{z \in S} T_z)$. Since $\{x, y\} \subset S \subseteq \cup_{z \in S} T_z$, Axiom 2 implies that $x \neq c(S)$. This is a contradiction. So the implication must be true. On the other hand, $y \in \Gamma^P_{k^*}(S)$ and $x = c(S)$ imply that there must exist a certain $S' \supset \{x, y\}$ with $|S'| > k^*$ such that $y = c(S')$. Both implications established above suggests that for any S' satisfying that $|S'| > k^*$ and $y = c(S')$, $y \neq c(S' \cup \{z\})$. This defines zPy, which contradicts the assumption that $y \in \Gamma^P_{k^*}(S)$. Hence, it must be that $y \succ x$. This establishes the sufficiency of Axioms 1-5.

"Only if" direction. Suppose that c is a WARP-violating choice function that is rationalizable by a shortlisting with capacity-k for some k, i.e., $c(S) = max(\Gamma^P_k(S), \succ)$. The definition of threshold capacity k^* suggests that $2 \leq k \leq k^* \leq N - 1$.

Firstly, we show that Axiom 1 is satisfied. $x = c(\{x, y\})$ and $y = c(\{y, z\})$ imply that $x \succ y$ and $y \succ z$. The transitivity of $\succ$ guarantees that $x \succ z$, which in turn implies that $x = c(\{z, x\})$.

Secondly, we show that Axiom 2 is satisfied. Assume that $\{x, y\} \subset T \subset S$ and $x = c(\{x, y\}) = c(S)$. $x = c(\{x, y\})$ implies that $x \succ y$. If $|T| \leqslant k$, $c(T) = max(T, \succ)$. $x \succ y$ implies $y \neq c(T)$. If $|T| > k$, then $c(T) = max(\Gamma^P_k(T), \succ)$ and $c(S) = max(\Gamma^P_k(S), \succ)$. Note that $x = c(S)$ implies that $x \in \Gamma^P_k(S)$, which implies that $x \in \Gamma^P_k(T)$. When $y \in \Gamma^P_k(T)$, $y \neq c(T)$ because $x \succ y$. When $y \notin \Gamma^P_k(T)$, it is obvious that $y \neq c(T)$.

Thirdly, we show that Axiom 3 is satisfied. Assume that $x = c(S) = c(T)$ and $|S|, |T| > k^*$. Then $|S|, |T| > k$, and $c(S) = max(\Gamma^P_k(S), \succ)$ and $c(T) = max(\Gamma^P_k(T), \succ)$. $x = c(S) = c(T)$ then suggests that $x \in \Gamma^P_k(S) \cap \Gamma^P_k(T)$, i.e., there is no $z \in S \cup T$, such that zPx. So $x \in \Gamma^P_k(S \cup T)$. If there is another option $y \in \Gamma^P_k(S \cup T)$, $y \in \Gamma^P_k(S)$ when $y \in S$ and $y \in \Gamma^P_k(T)$ when $y \in T$. In both cases it must be that $x \succ y$ because $x = c(S) = c(T)$. This establishes that $x = c(S \cup T)$.

Axiom 4 is satisfied because of the asymmetry of P. Assume that $c(T) \neq c(T \cup c(S))$, where $|S|, |T| > k^*$. Since $|S|, |T| > k^* \geq k$, $c(T) \neq c(T \cup c(S))$ suggests that either $c(S)Pc(T)$ or $c(S) \succ c(T)$ when neither $c(T)Pc(S)$ nor $c(S)Pc(T)$ occurs. In the first case, $c(T)Pc(S)$ cannot occur and then $c(S) = c(S \cup c(T))$. In the second case, $c(S) = c(S \cup c(T))$ regardless of whether $c(T) \in \Gamma^P_k(S \cup c(T))$ or not.

Finally, we show that Axiom 5 is satisfied. Assume that there exists a choice set T satisfying that for any two options $x, y \in T$, $x = c(S)$ and $y = c(S')$ for some S and S' satisfying that $|S|, |S'| > k^*$, and $\{x, y\} \subset S \cap S'$. Since $k^* \geq k$, $x = c(S)$ and $y = c(S')$ suggests that neither xPy nor yPx occurs. Then $|T| \leq k$ because otherwise

$|max(T,P)| = |T| > k$, which violates the condition that $|max(S,P)| \leq k$ for any $|S| > k$. Suppose that $x = c(S_1) = c(S_2)$ and $|S_1 \cup S_2| \leq |T|$. Then $|S_1 \cup S_2| \leq k$ and $x \succ y$ for any $y \in S_1 \cup S_2$. Therefore, $x = c(S_1 \cup S_2)$. This establishes the necessity of Axioms 1-5. □

A.4 Proof of Theorem 3

Proof. "If" direction. Define $k = max\{|S|, S \in \mathcal{FC}\}$. The property of Separability implies that $T \in \mathcal{LC}$ ($T \in \mathcal{FC}$) if and only if $|T| > k$ ($|T| \leqslant k$).

We first show that $\Gamma(T) \in \mathcal{FC}$ for any $T \in \mathcal{X}$. When $T \in \mathcal{FC}$, $\Gamma(T) = T \in \mathcal{FC}$. When $T \in \mathcal{LC}$, let $D = \Gamma(T)$. We first show that $D \subseteq \Gamma(D)$. To see this, let $x \in D = \Gamma(T)$. Then $x \in \Gamma(T) \subseteq T$, and the property of competition filter (i.e., More is Less) implies that $x \in \Gamma(\Gamma(T)) = \Gamma(D)$. Now consider a binary partition of D: D_1 and D_2. $D \subseteq \Gamma(D) = \Gamma(D_1 \cup D_2) \subseteq \Gamma(D_1) \cup \Gamma(D_2) \subseteq D_1 \cup D_2 = D$, where the second inclusion relation comes from the property of More is Less. Hence, it must be that $\Gamma(D) = D$, i.e., $\Gamma(T) = D \in \mathcal{FC}$.

We then consider the following three cases, respectively.

Case 1: $1 < k < N$.

Define a binary relation P as follows: xPy if (1) $x \in \Gamma(T)$ for some $T \in \mathcal{LC}$ satisfying that $T \supset \{x, y\}$, but $y \notin \Gamma(T)$ for any $T \in \mathcal{LC}$ satisfying that $T \supset \{x, y\}$; or (2) $y \in \Gamma(S)$ for some $S \in \mathcal{LC}$ but for any T' satisfying that $T' \in \mathcal{LC}$ and $y \in \Gamma(T')$, $y \notin \Gamma(T' \cup \{x\})$. Define $max(T, P) = \{x \in T : \text{no } y \in T \text{ such that } yPx\}$.

Firstly, we show that P is asymmetric. Suppose by contradiction that both xPy and yPx. The definition of P suggests that if $x \in \Gamma(T)$ for some $T \in \mathcal{LC}$ satisfying that $T \supset \{x, y\}$, then yPx can not occur. Hence, xPy and yPx imply that for any $T \in \mathcal{LC}$ satisfying that $T \supset \{x, y\}$, $x \notin \Gamma(T)$ and $y \notin \Gamma(T)$. Then the only situation that makes xPy and yPx happen must be that $y \in \Gamma(S_1)$ for some $S_1 \in \mathcal{LC}$ but for any $S_1^{'}$ satisfying that $S_1^{'} \in \mathcal{LC}$ and $y \in \Gamma(S_1^{'})$, $y \notin \Gamma(S_1^{'} \cup \{x\})$, and also that $x \in \Gamma(S_2)$ for some $S_2 \in \mathcal{LC}$ but for any $S_2^{'}$ satisfying that $S_2^{'} \in \mathcal{LC}$ and $x \in \Gamma(S_2^{'})$, $x \notin \Gamma(S_2^{'} \cup \{y\})$. This implies that $y \in \Gamma(S_1)$ for some $S_1 \in \mathcal{LC}$ and $x \in \Gamma(S_2)$ for some $S_2 \in \mathcal{LC}$, and for any $T \in \mathcal{LC}$ including options x and y, $\{x, y\} \cap \Gamma(T) = \emptyset$, which contradicts the property of No Mutual Exclusion. Therefore, P defined above is asymmetric.

Secondly, we show that $\Gamma(T) = max(T, P)$ when $|T| > k$, and $\Gamma(T) = T$ when $|T| \leqslant k$. The second part of the claim is established due to the property of Separability, as provided in the beginning of the proof. Now consider $|T| > k$, i.e., $T \in \mathcal{LC}$. We first establish that $\Gamma(T) \subseteq max(T, P)$. Consider any $x \in \Gamma(T)$ and another $y \in T$. According to the definition of P, yPx cannot occur. So $\Gamma(T) \subseteq max(T, P)$. We then establish that $max(T, P) \subseteq \Gamma(T)$. Assume that $x \in max(T, P)$ and we need to show that $x \in \Gamma(T)$. Consider any $y \in \Gamma(T)$. Then $y \in max(T, P)$, which implies that neither xPy nor yPx. The fact that $y \in \Gamma(T)$

and yPx cannot occur implies that there must exist a certain $S \in \mathcal{LC}$ satisfying that $x \in \Gamma(S)$ and $S \supset \{x, y\}$. Now consider an arbitrary $z \in T$ and note that zPx cannot occur. Since $x \in \Gamma(S)$ for some $S \in \mathcal{LC}$, the fact that zPx cannot occur imply that among those S's satisfying that $S' \in \mathcal{LC}$ and $x \in \Gamma(S')$, there exists at least one S' such that $x \in \Gamma(S' \cup \{z\})$. This suggests that $x \in \Gamma(S_z)$ for some $S_z \in \mathcal{LC}$ satisfying that $S_z \supset \{x, z\}$. The property of Weak Consideration Dominance then implies that $x \in \Gamma(\cup_{z \in T} S_z)$. Since $T \subseteq \cup_{z \in T} S_z$, the property of competition filter (i.e., More is Less) implies that $x \in \Gamma(T)$. Hence, $max(T, P) \subseteq \Gamma(T)$. This shows that $\Gamma(T) = max(T, P)$ when $|T| > k$.

Finally, $|max(T, P)| \leq k$ when $|T| > k$ because $|max(T, P)| = |\Gamma(T)|$ and $\Gamma(T) \in \mathcal{FC}$. This shows that Γ can be represented as a shortlisting with capacity-k.

Case 2: $k = 1$.

In this case, for any two-element set $S = \{x, y\}$, $\Gamma(S) \neq S$, i.e., $S \in \mathcal{LC}$. Then $\Gamma(S) = \{x\}$ or $\{y\}$. Define xPy if $\Gamma(\{x, y\}) = \{x\}$. It is obvious that P is asymmetric. We then show that $\Gamma(T) = max(T, P)$ when $|T| > 1$. Since $\Gamma(T) \in \mathcal{FC}$, we assume that $\{x\} = \Gamma(T)$ without loss of generality. The property of competition filter then implies that $\{x\} = \Gamma(\{x, y\})$ for any other $y \in T$. So xPy and then $\{x\} = max(T, P)$. Now assume that $x \in max(T, P)$. Consider any other $y \in T$. Since no yPx, $\Gamma(\{x, y\}) \neq \{y\}$, which implies that $\Gamma(\{x, y\}) = \{x\}$. The property of Weak Consideration Dominance implies that $x \in \Gamma(\cup_{y \in T}\{x, y\}) = \Gamma(T)$. Therefore, $\Gamma(T) = max(T, P)$ when $|T| > 1$. Finally, the fact that $\Gamma(T) = max(T, P)$ and $|\Gamma(T)| = 1$ implies that $|max(T, P)| = 1$. This shows that Γ can be represented as a shortlisting with capacity-1.

Case 3: $k = N$.

In this case, $\Gamma(T) = T$ for any $T \in \mathcal{X}$. Let P be any rationale, including the empty set. It is trivial to establish that Γ can be represented as a shortlisting with capacity-N.

"Only if" direction. Suppose Γ can be represented as a shortlisting with capacity-k, i.e., $\Gamma(T) = max(T, P)$ when $|T| > k$, and $\Gamma(T) = T$ when $|T| \leq k$, where P is asymmetric and $|max(T, P)| \leq k$.

Suppose that $x \in S \subset T$ and $x \in \Gamma(T)$. If $|S| \leqslant k$, then $x \in S = \Gamma(S)$. If $|S| > k$, then $|T| > k$ and $\Gamma(T) = max(T, P)$. This implies that no yPx for any $y \in T$. So no yPx for any $y \in S$ and then $x \in max(S, P) = \Gamma(S)$. This shows that Γ satisfies the property of compeition filter, i.e., the property of More is Less.

By the assumption, $T \in \mathcal{FC}$ if $|T| \leq k$. In addition, when $|T| > k$, $|\Gamma(T)| = |max(T, P)| \leq k$, which implies that $\Gamma(T) \neq T$, i.e., $T \in \mathcal{LC}$. This essentially establishes that $T \in \mathcal{FC}$ if and only if $|T| \leq k$, and that $T \in \mathcal{LC}$ if and only if $|T| > k$. Therefore, the consideration function Γ satisfies the property of Separability.

Consider $T_1, T_2 \in \mathcal{LC}$. Assume that $x \in \Gamma(T_1) \cap \Gamma(T_2)$. Then $x \in max(T_1, P)$ and $x \in max(T_2, P)$, which implies that $x \in max(T_1 \cup T_2, P) = \Gamma(T_1 \cup T_2)$. This shows that the consideration function satisfies the property of Weak Consideration Dominance.

Now suppose that $x \in \Gamma(T_1)$ and $y \in \Gamma(T_2)$, where $T_1, T_2 \in \mathcal{LC}$. If $x \notin \Gamma(T_1 \cup \{y\}) = max(T_1 \cup \{y\}, P)$, then it must be that yPx. Since the relation P is asymmetric, xPy cannot occur and then $y \in max(T_2 \cup \{x\}, P)$=$\Gamma(T_2 \cup \{x\})$. In other words, we find a $T \in \mathcal{LC}$ such that $\{x, y\} \cap \Gamma(T) \neq \emptyset$. This shows that the consideration function satisfies the property of No Mutual Exclusion. □

A.5 Proof of Theorem 5

Proof. Assume that a choice function is rationalizable by a top-k shortlisting for some k. When the choice function satisfies the WARP, apparently the size of the top-k shortlist can be N or 1. We also show that the size of the top-k shortlist can be $N-1$ in this case. Suppose that c is top-N-focus rationalizable, i.e., $max(S, \succ) = c(S)$. Let $x = c(X)$ and y is another option in X. So y is never chosen when x is available. Let xPy and (x, y) be the unique pair of options that has the relation P. Define $\Gamma^P_{N-1}(S) = S$ if $S \neq X$ and $\Gamma^P_{N-1}(X) = X/\{y\}$. Then obviously $max(\Gamma^P_{N-1}(S), \succ) = c(S)$. In other words, c is rationalizable by a top-$N-1$ shortlisting.

When the choice function is a WARP-violating choice function that has a unique WARP-violating choice pair in which X is the WARP-violating choice set, there exists a unique option y that is "better" than the chosen option in the grand set X. In other words, $y = c(y, c(X))$ and there is no other WARP-violating choice pair. Then define $z \succ w$ if $z = c(\{z, w\})$, $c(X)Py$, $\Gamma^P_{N-1}(X) = X/\{y\}$, and $\Gamma^P_{N-1}(S) = S$ if $S \neq X$. It is straightforward to establish that $c(S) = max(\Gamma^P_{N-1}(S), \succ)$, i.e., it is rationalizable by a top-$N-1$ shortlisting.

Therefore, it remains to show that $N-1 > k > 1$ is impossible when $N > 4$ and k may be 2 when $N = 4$.

We first show that $N-1 > k > 2$ is impossible. Since the size of the top-k shortlist of X is $k < N-1$ in this case, there must be at least two options that are eliminated according to a rationale P. Three situations exist for two eliminated options: (1) xPy and wPz, (2) xPy and xPz, or (3) xPy and yPz. In each of the three situations, we can find a choice set that includes $k+1$ options and includes the corresponding three or four options. For this choice set, the size of the top-k shortlist is less than k, which violates the definition of top-k shortlisting. Thus, $N-1 > k > 2$ is impossible.

We then show that $k = 2$ is impossible when $N > 4$. In the case of $k = 2$, $N-2$ options are eliminated according to a rationale P because the size of the top-2 shortlist of X is two. Since $2(N-2) > N$ for $N > 4$, it is impossible that the eliminated options and the options that do the elimination are $2(N-2)$ distinct options. In other words, either of the following two situations must occur: (1) xPy and yPz, or (2) xPy and xPz. Then the top-2 shortlist of $\{x, y, z\}$ contains only one option, which violates the definition of top-2 shortlisting. Thus, $k = 2$ is impossible either.

Finally, Example 3 illustrates that $k = 2$ is possible when $N = 4$. □