Choice with Limited Capacity *

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

[1]Xiamen University

[2]University of Maryland

This version: August 22, 2019

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

*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, Bart Lipman, 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. Sen Geng thanks the financial support from the Fundamental Research Funds for the Central Universities (Grant No. 20720161035) and the National Science Foundation of China (Grant No. 71703134).

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: D01

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

[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).

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).[2]

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

[2] Gerasimou (2018) models the implication of limited capacity from a different perspective: a decision maker uses an overload criterion to determine whether she will choose the best alternative in a choice set or whether she will defer choice.

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 No Binary Cycles 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 only conditionally 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.

The rest of the paper will be structured as follows: in Section 2, we define and characterize the limited capacity model, infer shortlist and conflicting rationale from choice, discuss the related literature, and discuss a special case in which the DM considers exactly a certain number of options. Section 3 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).[3] However, as argued in 10 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. Our specific way of modeling has an important aspect at the conceptual level: the DM in Manzini and Mariotti (2007) always uses a rationale to shortlist the alternatives to consider, while in the classical setup, the DM never shortlists the alternatives. In our setup, the DM uses the shortlist only when there are abundant alternatives.

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*

[3] de Clippel et al. (2014) and Dardanoni et al. (2018) look at limited capacity from very different perspectives and they also adopt the assumption of capacity-k in their models.

$S \in \mathcal{X}$:

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

and $0 < |max(S, P)| \leq k$ *if* $|S| > k$.[4]

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, $\succ \subseteq X \times X$.[5] 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 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 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

[4]Formally, P is an asymmetric binary relation with width no more than k and P does not have any cycles of distinct elements of length larger than k. Here width(X, P):=$sup\{|Y| :$ Y is a choice set in which any two options are not related according to P.$\}$

[5]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.

only if it is rationalizable by a shortlisting with capacity-N. As in any two-stage choice model, the choice function that satisfies WARP is the least informative one. In our setup, this means that it is rationalizable by a shortlisting with any capacity. In addition, it is straightforward to establish that being rationalizable by a shortlisting with capacity-1 coincides with being rationalizable by a shortlisting with capacity-N, i.e., coincides with full rationality. In other words, a DM who has a capacity of one shares 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.

We now look at choice functions that do not satisfy WARP. Clearly, these choice functions are not rationalizable by a shortlisting with capacity-1 because otherwise they are rationalizable by a shortlisting with capacity-N. Then, if a choice function that does not satisfy 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 two 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 WARP, exhibiting no cycle in binary choices is an immediate implication of WARP. Thus, it is natural to assume no binary cycles:

Axiom 1 (No Binary Cycles). $x = c(\{x, y\})$ *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 $x = c(S \cup T)$. On the one hand, it is straightforward to establish that the axiom of no binary cycles and the property of Expansion is equivalent to WARP (See Appendix A.1). This suggests that an addition of the property of Expansion makes it impossible to accommodate choice behavior that violates WARP. In other words, given that the axiom of no binary cycles 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 WARP may be violated.

On the other hand, the elimination machine in the present model does operate when the size of choice set exceeds the DM's capacity and the DM is fully rational when the size of choice set falls 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).

Motivated by the above observation, it is appealing to assume that choice sets are categorized into two classes based on sizes of choice sets, and that the property of Expansion holds when all the three choice sets belong to the class of choice sets with a smaller size and when all the three choice sets belong to the class of choice sets with a larger size. We label this property Conditional Expansion and to better formalize it we firstly define:

Definition 4. *Alternatives* x *and* y *are* ***level-*k *mutually undominated*** *if there exists sets* S, S' *with* $|S|, |S'| > k$, $\{x, y\} \subset S \cap S'$ *such that* $x = c(S)$ *and* $y = c(S')$. *A set* T *is a* ***set of level-*k *mutually undominated alternatives*** *if any two distinct elements in* *the set are level-k mutually undominated.*

Clearly, if T is a set of level-k mutually undominated alternatives then it is also a set of level-k' mutually undominated alternatives for $k' < k$. In addition, a set of level-N mutually undominated alternatives is empty set and its size is zero. The size of level-1 mutually undominated alternatives is zero if and only if WARP holds.

Axiom 3 (Conditional Expansion). *There exists a number, k, no less than the size of any set of level-k mutually undominated alternatives satisfying the property that if either $|S|, |T| > k$ or $|S \cup T| \leq k$, then: (1) $c(S) = c(T)$ implies that $c(T) = c(T \cup S)$; and, (2) $c(S) \neq c(S \cup \{c(T)\})$ implies that $c(T) = c(T \cup \{c(S)\})$.*

It is notable that in Axiom 3 claim (1) is exactly the property of Expansion and claim (2) is a variant of the property of Expansion: $c(S) \neq c(S \cup \{c(T)\})$ implies that $c(T) = c(\{c(T), c(S)\})$, which in turn implies that $c(T) = c(T \cup \{c(S)\})$.

One may naturally wonder the difficulty of empirically testing Axiom 3 since it is formulated as "there exists a number, k, $\cdots$" In fact, the procedure of checking Axiom 3 per se is straightforward according to the proof of Theorem 1. Roughly speaking, we firstly check whether Axiom 3 holds for $k = 1$ (or equivalently for $k = N$), and conditional on violation we then check whether Axiom 3 holds for a certain number that will be defined shortly. If Axiom 3 holds for neither of the above two numbers, it can be easily established that Axiom 3 does not hold for any number.

Our representation theorem (i.e., Theorem 1) shows that the axiom of No Binary Cycles, the axiom of Weak WARP, and the axiom of Conditional Expansion exactly capture the essence of choice by a shortlisting with limited capacity.

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 1-3.*

Proof. See the proof in Appendix A.2. □

Theorem 1 shows the sufficient and necessary conditions on a choice function for the existence of the capacity, which is treated as unobservable. A natural question then is when the capacity exists, whether a choice function has a unique capacity representation or not. To provide an intuitive analysis of this question, we introduce below the concepts of WARP-violating choice function and threshold capacity.

Definition 5. *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 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, WARP always holds for choice sets whose size is smaller than the size of the smallest WARP-violating choice sets, and 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 WARP always holds when the size of choice sets falls within the threshold capacity and WARP starts to fail when the size exceeds the threshold capacity. In the degenerate case where no WARP-violating choice sets exists, i.e., WARP always holds, it is natural to define the threshold capacity as N.

Definition 6. *The* ***threshold capacity*** *of a choice function is* $k_c = min\{|S| - 1 : S$ *is an arbitrary WARP-violating choice set given the choice function*$\}$ *if it is a WARP-violating choice function, and is* $k_c = N$ *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.

We further know from the proof of Theorem 1 that a WARP-violating choice function is rationalizable by a shortlisting with capacity-k_c 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. It is also notable that 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 is WARP-violating but rationalizable by a shortlisting with capacity-k, then k is unique and $k = k_c$.*

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

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 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.[6] Indeed, the 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. See the proof in Appendix A.4. □

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

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

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 $\mathcal{LC}$[7]

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

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

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.[8]

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, then either the first option still wins a competition when the second option is added or the second option still wins a competition when the first option is added as long as the two options do not mutually exclude each other from being considered.

No Mutual Exclusion: If $x \in \Gamma(S)$ and $y \in \Gamma(S')$ for some $S, S' \in \mathcal{LC}$, then

[7]It is notable that our characterization of the consideration function in terms of the fixed points of the mapping Γ could be quite helpful for other two-stage models in the literature.

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

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 Appendix A.5. □

2.1 Revealed Preference, Shortlist, and Conflicting Rationale

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 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 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 WARP holds.

We now look at the revealed shortlist of any choice set. If WARP holds for a choice function, the revealed shortlist of a choice set is not interesting because it 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_c according to Theorem 2. In this case, the revealed shortlist of a choice set with a size less than k_c is trivial and is the same as the choice set. Nevertheless, the revealed shortlist of a choice set with a size exceeding k_c 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_c 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 7. *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_c. 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 $|T| > k_c$ 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 $\{x, y\} \subseteq S \cap S'$ and $|S|, |S'| > k_c$, then $\{x, y\} \subseteq \Gamma_c(S)$. In fact, it is evident that $\Gamma_c(S)$ is a set of level-k_c mutually undominated alternatives.

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_c. If $(k_c, P_1, \succ)$, $(k_c, P_2, \succ)$, $\cdots$, and $(k_c, P_n, \succ)$ are all possible combinations of capacity, rationale and preference relation that rationalize the choice function, it must be that $\Gamma_c(S) = \cap_{i=1}^{n} \Gamma_{k_c}^{P_i}(S)$.*

Proof. See the proof in Appendix A.6. □

We finally explore what can be inferred about the underlying rationale if a WARP-violating choice function is rationalizable by a shortlisting with limited capacity. Dutta and Horan (2015) provides a nice methodology guidance on the identification problem by providing a comprehensive analysis of pinning down the underlying two rationales in Manzini and Mariotti (2007)'s *shortlisting* choice model.[9] In our model, among all triples that rationalize a WARP-violating choice function, capacity and preference are both uniquely pinned down from choice. Hence, the interesting question is to identify a rationale which is in contrast with revealed preference, i.e., $x \succ y$ but yPx. Similar to Masatlioglu et al. (2012) and Dutta and Horan (2015), we follow a conservative approach to define such revealed rationale in the sense that only those features which are common in all possible representations are captured.

Definition 8. *Let $\mathcal{R}(c) \equiv \{(k_c, P_1, \succ), \cdots, (k_c, P_n, \succ)\}$ denote the collection of all triples that rationalize a WARP-violating choice function c. We define the* ***revealed conflicting rationale*** *P^c by xP^cy if $y \succ x$ and xP_iy for all $(k_c, P_i, \succ) \in \mathcal{R}(c)$.*

We show that revealed conflicting rationale in Definition 8 has an appealing behavioral characterization: (1) y is chosen over x in the binary choice of the two

[9]They make two observations about choice from two-element choice sets: (1) every revealed preference must be belong to either the first or the second rationale; and, (2) the first rationale is a subset of revealed preference, based on a generalization of which they characterize the relationship between behavior and the underlying rationales. However, in our setup the second "rationale" is pinned down as revealed preference for WARP-violating choice functions and the first rationale may not operate. Thus, while their first observation still applies, their second observation fails to hold in our choice model.

options; (2) when y is ever chosen in a choice set whose size exceeds the threshold capacity, it is never chosen in choice sets that include x and have a size exceeding the threshold capacity; when y is never chosen in a choice set whose size exceeds the threshold capacity, there exist a situation in which x has to eliminate y.

Theorem 5 (**Revealed Conflicting Rationale**). *Suppose that a WARP-violating choice function c is rationalizable by a shortlisting with limited capacity, i.e., by $\mathcal{R}(c)$. Then $xP^c y$ if and only if (1) $y = c(\{x, y\})$; (2) either of the following two conditions holds: (2a) there exists a certain S with $|S| > k_c$ such that $y = c(S)$ and $y \neq c(T)$ for any T satisfying that $|T| > k_c$ and $T \supset \{x, y\}$; (2b) $y \neq c(S)$ for any $|S| > k_c$ and there exists a certain T with $|T| = k_c + 1$ and $T \supset \{x, y\}$ satisfying that $x = c(T)$ and for any $z \in T/\{x, y\}$, $z = c(T')$ for some T' with $|T'| > k_c$ but $z \neq c(T' \cup \{y\})$.*

Proof. See the proof in Appendix A.7. □

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 1.

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

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

explain cyclical binary choice, it is falsifiable for a grand choice set consisting of three options.

Table 2 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.

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 2 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

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 2: Comparison of different choice models

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 1. *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. Masatlioglu et al. (2012) shows that their model can be exactly captured by 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 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 $x_4 \neq x^*$.*

2.3 A Special Case: Choice with Top-k Shortlisting

The focus of our general model is when to use an elimination rationale. Once the elimination stage is triggered, one can imagine that depending on the rationale, it may be the case that the DM sorts alternatives based on a criterion and looks at the first k alternatives, or it may be the case that she looks at the best one, or it may be the case that she looks at more than one but strictly less than k alternatives. We discuss 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).[10] In addition to capturing some realistic examples, the special case has a normatively appealing feature: shortlisting with capacity-k is continuous at the capacity-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.

The special case is similar to Example 3 in Salant and Rubinstein (2008) but the capacity of a DM is assumed to be observable there. In addition, Eliaz et al. (2011) axiomatizes the procedures of choosing the top two, choosing the two extremes, and choosing the top and the top. Chambers and Yenmez (2018) provides an alternative characterization of the procedure of choosing the top q ($q \geq 2$) finalists. The size of finalists is still assumed to be observable in their studies. Instead, we assume unobservable capacity. In Appendix A.8, 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.

[10] Bajraj and Ulku (2015) considers an alternative model according to which a DM applies a linear order to shortlist the top two alternatives and then chooses the winner in this binary shortlist using the second criterion. A characterization of a generalization to larger shortlists in their modelling approach remains an open question.

3 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. For example, a customer, who is deciding what to order among 50 chicken dishes, may use spiciness as a rationale and may consider three least spicy dishes. Alternatively, an insurance buyer may use price as a rationale and may consider five cheapest options. Since for a given grand set of options we determine uniquely the capacity, using different rationales and capacities in different choice domains is consistent with our setup. 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.

An interesting extension of our setup is multiple rationales. In other words, a DM with a limited capacity needs to use multiple rationales to reduce the number of options to be within her capacity. In such environment, it may be fruitful to investigate the effect using multiple rationales simultaneously or sequentially. Another interesting extension of our model is menu-dependent capacity: explicitly modeling a DM's capacity that depends on choice sets, even in the same choice domain (i.e., for a given grand choice set). Our present setup is able to accommodate this type of variation of capacities in some situations by using the smallest capacity among all menu-dependent capacities. Of course, we understand that an explicit characterization of menu-dependent capacity allows more flexibility and we leave the above two extensions for future study.

A Appendix

A.1 Proposition 1

Proposition 1. *A choice function satisfies the axiom of No Binary Cycles and the property of Expansion if and only if it satisfies WARP.*

Proof. Consider a choice function, c, on $X = \{x_1, \cdots, x_N\}$. Assume that it satisfies the axiom of No Binary Cycles and the property of Expansion. We first show that the axiom of No Binary Cycles 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 No Binary Cycles 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., WARP holds.

Now assume that WARP holds. The axiom of No Binary Cycles 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., $y = c(S \cup T)$. 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 Proof of Theorem 1

Proof. "If" direction. Assume that a choice function c satisfies Axioms 1-3. We show that it is rationalizable by a shortlisting with capacity-k for some k.

Case 1: when Axiom 3 holds for $k = 1$ or $k = N$.

We show that in this case Axiom 3 also holds for any other number k and the choice function is rationalizable by a shortlisting with any capacity. Since $|S|, |T| > 1$ and $|S \cup T| \leq N$, the condition of Expansion that is specified in Axiom 3 is trivially satisfied in this case, and in turn the property of Expansion always holds, i.e., $c(S) = c(T)$ implies that $c(T) = c(T \cup S)$. Axiom 1 and Proposition 1 then imply that c satisfies WARP. Consequently, it is straightforward to establish that Axiom 3 also holds for any other number k and c is rationalizable by a shortlisting with any capacity.

Case 2: when Axiom 3 holds only for some k satisfying that $1 < k < N$.

Clearly, WARP must be violated in this case because otherwise Axiom 3 holds for any other number including 1 and N. Let k_c denote the threshold capacity of the WARP-violating choice function c, i.e., the size of the smallest choice sets that make WARP fail minus one.

We now show that it is must be true that the number $k \leq k_c$. Suppose by contradiction that $k > k_c$. Note that Axiom 3 requires that the property of Expansion holds up to choice sets whose size does not exceed k, i.e., $c(S) = c(T)$ implies that $c(T) = c(T \cup S)$ as long as $|S \cup T| \leq k$. Axiom 1 and Proposition 1 then imply that c satisfies WARP up to choice sets whose size does not exceed k, i.e., $x = c(\{x, y\})$ implies that $y \neq c(S')$ for any S' with $|S'| \leq k$ and $S' \supset \{x, y\}$. Thus, WARP holds up to choice sets whose size does not exceed $k_c + 1$. This violates the definition of threshold capacity, k_c. So it must be that $k_c \geq k$.

We then show that c is rationalizable by a shortlisting with capacity-k_c. Specifically, define the preference relation $\succ$ by $x \succ y$ if $x = c(\{x, y\})$. Define the elimination rationale P by xPy if (1) $x = c(S)$ for some $S \supset \{x, y\}$ with $|S| > k_c$ but $y \neq c(S')$ for any $S' \supset \{x, y\}$ with $|S'| > k_c$, or (2) $y = c(T)$ for some T with $|T| > k_c$ but $y \neq c(T' \cup \{x\})$ for any T' satisfying that $|T'| > k_c$ and $y = c(T')$. Define $\Gamma^P_{k_c}(S) = S$ if $|S| \leq k_c$, and $= max(S, P)$ if $|S| > k_c$. We establish that c is rationalizable by $(k_c, P, \succ)$ in the sense that $c(S) = max(\Gamma^P_{k_c}(S), \succ)$ as follows.

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

We firstly 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 $|S| > k_c$ but $y \neq c(S')$ for any $S' \supset \{x, y\}$ with $|S'| > k_c$. 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 $|T| > k_c$ but $y \neq c(T' \cup \{x\})$ for any T' satisfying that $|T'| > k_c$ 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 $|S| > k_c$ but $x \neq c(S' \cup \{y\})$ for any S' satisfying that $|S'| > k_c$ and $x = c(S')$. This implies that $c(T) \neq c(T \cup c(S))$ and $c(S) \neq c(S \cup c(T))$ for $|S|, |T| > k_c \geq k$, which violates Axiom 3. Hence, xPy implies that yPx cannot occur, i.e., P is asymmetric.

The definition of P suggests that if an option is ever chosen in a choice set with more than k_c elements, it can never be eliminated by any element in the choice set according to the rationale. Thus, P does not have any cycles of distinct elements of length larger than k_c, i.e., $|max(S,P)| > 0$ for $|S| > k_c$.

We then show that $|max(S,P)| \leq k_c$ for $|S| > k_c$. Suppose by contradiction there exists a certain S such that $|max(S,P)| > k_c$. Without loss of generality, we assume that $|max(S,P)| = k_c + 1$, i.e., $S_{NP} \equiv max(S,P) = \{x_1, \cdots, x_{k_c}, x_{k_c+1}\}$, and $x_1 = c(S_{NP})$. Since $|S_{NP}| > k_c$ and there is no P relation between x_1 and x_i for any $i \in \{2, \cdots, k_c+1\}$, there must exist a certain $S_i \supset \{x_1, x_i\}$ satisfying that $|S_i| > k_c$ and $x_i = c(S_i)$. Since $x_i = c(S_i)$ for some S_i with $|S_i| > k_c$, 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 $|S_i^{'}| > k_c$ and $x_i = c(S_i^{'})$, there exists a certain $S_i^{'}$ such that $x_i = c(S_i^{'} \cup x_j)$, where $j \in \{2, \cdots, k_c+1\}$. This implies that $x_i = c(S)$ for some $S \supset \{x_i, x_j\}$ with $|S| > k_c$. 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 $|S'| > k_c$ and $x_j = c(S')$. In this way, we essentially construct a set S_{NP} of level-k_c mutually undominated alterantives. Since $k_c \geq k$, S_{NP} is also a set of of level-k mutually undominated alterantives. But $|S_{NP}| = k_c + 1 > k$, which contradicts Axiom 3. Thus, $|max(S,P)| \leq k_c$ for $|S| > k_c$.

We finally show that if $x = c(S)$, then it must be that $x = max(\Gamma^P_{k_c}(S), \succ)$. Assume that $x = c(S)$. When $|S| \leq k_c$, $max(\Gamma^P_{k_c}(S), \succ) = max(S, \succ)$. The definition of k_c 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 $|S| > k_c$, $x = c(S)$ implies that no $y \in S$ such that yPx according to the definition of P. In other words, $x \in \Gamma^P_{k_c}(S)$. We then only need to show that if there is another option $y \in \Gamma^P_{k_c}(S)$ then it must be that $x \succ y$.

Assume that $y \in \Gamma^P_{k_c}(S)$. 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 $|T| > k_c$ 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 $|T_z| > k_c$ and $T_z \supset \{y, z\}$, such that $y = c(T_z)$. Since $k_c \geq k$, 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, $y \in \Gamma^P_{k_c}(S)$ and $x = c(S)$ imply that there must exist a certain $S' \supset \{x, y\}$ with $|S'| > k_c$ such that $y = c(S')$. Both implications established above suggest that for any S' satisfying that $|S'| > k_c$ and $y = c(S')$, $y \neq c(S' \cup \{z\})$. This defines zPy, which contradicts the assumption that $y \in \Gamma^P_{k_c}(S)$. Hence, it must be that $y \succ x$. This establishes the sufficiency of Axioms 1-3.

"Only if" direction. Suppose that c is a choice function that is rationalizable by a shortlisting with capacity-k for some k, i.e., $c(S) = max(\Gamma^P_k(S), \succ)$. When $k = 1$ or $k = N$, c is fully rationalizable and it satisfies the WARP. Then it is straightforward to establish that c satisfies Axioms 1-3. We now consider the case in which c is rationalizable by a shortlisting with capacity-k only for $1 < k < N$.

Axiom 1 is satisfied because of the transitivity of $\succ$. We now 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 $y \neq c(T)$.

We finally show that Axiom 3 is satisfied. Clearly, if there is any set of level-k mutually undominated alternatives, its size does not exceed k. When $|S|, |T| > k$, $c(S) = max(\Gamma^P_k(S), \succ)$ and $c(T) = max(\Gamma^P_k(T), \succ)$. $c(S) = c(T)$ then suggests that $c(T) \in \Gamma^P_k(S) \cap \Gamma^P_k(T)$, i.e., there is no $z \in S \cup T$, such that $zPc(T)$. So $c(T) \in \Gamma^P_k(S \cup T)$. 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 $c(T) \succ y$ because

$c(T) = c(S)$. This establishes that $c(T) = c(S \cup T)$. $c(S) \neq c(S \cup c(T))$ suggests that either $c(T)Pc(S)$ or $c(T) \succ c(S)$ when neither $c(T)Pc(S)$ nor $c(S)Pc(T)$ occurs. In the first case, $c(S)Pc(T)$ cannot occur and then $c(T) = c(T \cup c(S))$. In the second case, $c(T) = c(T \cup c(S))$ regardless of whether $c(S) \in \Gamma_k^P(T \cup c(S))$ or not. When $|S \cup T| \leq k$, $c(S \cup T) = max(S \cup T, \succ)$. It is straightforward to establish that: (1) $c(S) = c(T)$ implies that $c(T) = c(S \cup T)$; and, (2) $c(S) \neq c(S \cup c(T))$ implies that $c(T) = c(T \cup c(S))$. This establishes the necessity of Axioms 1-3. □

A.3 Proof of Theorem 2

Proof. Suppose a choice function is WARP-violating, 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 $2 \leq k \leq k_c$. In the proof of Theorem 1, it is shown that capacity-k_c can rationalize the choice. If $k_c = 2$, then trivially it is only rationalizable by a shortlisting with capacity-2.

Consider now $k_c > 2$ 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_c, $z = c(T)$ for any $T \subset S$ such that $z \in T$ and $|T| \leq k_c$.

Suppose by contradiction that capacity k' also rationalizes the choice function for some $2 \leq k' < k_c$. 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 $k' < |T| \leq k_c$. This is a contradiction. Therefore, the choice function is only rationalizable by a shortlisting with capacity-k_c. □

A.4 Proof of Lemma 1

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. $\square$

A.5 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, it is straightforward to establish that $|max(T, P)| > 0$ for $|T| > k$ through proof by contradiction. $|max(T, P)| \leq k$ 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 $0 < |max(T, P)| \leq k$ for $|T| > k$.

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.6 Proof of Theorem 4

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

Now assume that $x \in \cap^n_{i=1}\Gamma^{P_i}_{k_c}(S)$ in the remaining parts. We firstly show that this implies that there exists a certain T with $|T| > k_c$ such that $x = c(T)$. Suppose by contradiction that $x \neq c(T)$ for any T with $|T| > k_c$. Consider augmenting the elimination rationale defined in the proof of Theorem 1 by additionally defining $c(S)Px$ if it is not included previously. Note that this addition does not generate inconsistency with the choice function. In addition, this addition still keeps the asymmetry of elimination rationale P since $xPc(S)$ cannot occur because of the way we define P in the proof of Theorem 1. Finally, this addition does not create a cycle of distinct elements of length larger than k_c, i.e., $xPx_1Px_2P \cdots Px_{k_c-1}Pc(S)Px$ cannot occur. The last claim is true because the chosen option from the above $k_c + 1$-element set can only come from $\{x_1, x_2, \cdots, x_{k_c-1}, c(S)\}$ and then the chosen option cannot be eliminated according to P. In summary, the choice function is rationalizable by $(k_c, P, \succ)$ with the new elimination rationale P. However, $c(S)Px$ contradicts $x \in \cap^n_{i=1}\Gamma^{P_i}_{k_c}(S)$.

We then establish $x \in \Gamma_c(S)$ by contradiction. Suppose by contradiction that $x \notin \Gamma_c(S)$. Then there exists a certain $y \in S$ such that for any $S' \supset \{x, y\}$ with $|S'| > k_c$, $x \neq c(S')$. Since $x = c(T)$ for some T with $|T| > k_c$, the definition of P in the proof of Theorem 1 implies that yPx, which contradicts $x \in \cap_{i=1}^{n} \Gamma_{k_c}^{P_i}(S)$. Therefore, $x \in \cap_{i=1}^{n} \Gamma_{k_c}^{P_i}(S)$ implies that $x \in \Gamma_c(S)$. □

A.7 Proof of Theorem 5

Proof. Assume that conditions (1) and (2a) hold. Suppose by contradiction that xPy does not hold for some rationale P with $(k_c, P, \succ) \in \mathcal{R}(c)$. Then either yPx or there is no P relation between them. In both cases, (1) and (2a) imply that $y = c(S \cup \{x\})$, which is a contradiction. Assume that conditions (1) and (2b) hold. (1) and (2b) imply that y must be eliminated by some element according to P_i and any $z \in T/\{x, y\}$ cannot eliminate y according to P_i, where $(k_c, P_i, \succ) \in \mathcal{R}(c)$. So it must be that $xP^c y$.

Assume that $xP^c y$. (1) comes from Definition 8. Suppose by contradiction that (2) fails, i.e., neither (2a) nor (2b) holds. This implies that either (i) there exists a certain T with $|T| > k_c$ and $T \supset \{x, y\}$ such that $y = c(T)$; or (ii) $y \neq c(S)$ for any $|S| > k_c$ and there does not exist a certain T with $|T| = k_c + 1$ and $T \supset \{x, y\}$ satisfying that $x = c(T)$ and for any $z \in T/\{x, y\}$, $z = c(T')$ for some T' with $|T'| > k_c$ but $z \neq c(T' \cup \{y\})$. (i) evidently contradicts $xP^c y$. We show that (ii) also contradicts $xP^c y$ as follows.

Consider the elimination rationale P defined in the proof of Theorem 1, in which we show that $(k_c, P, \succ) \in \mathcal{R}(c)$. Consider $T \supset \{x, y\}$ and $|T| > k_c$. We firstly show that if $x = c(T)$, there exists a certain $z_T \in T/\{x, y\}$ such that yPz_T cannot occur. Suppose by contradiction that yPz for any $z \in T/\{x, y\}$. Since $y \neq c(S)$ for any $|S| > k_c$, the only situation that makes yPz happen is that $z = c(T')$ for some T' with $|T'| > k_c$ but $z \neq c(T' \cup \{y\})$. So yPz for any $z \in T/\{x, y\}$ implies that for any $z \in T/\{x, y\}$, $z = c(T')$ for some T' with $|T'| > k_c$ but $z \neq c(T' \cup \{y\})$. This implication cannot be true given (ii) for the reasons below: (ii) implies that when $x = c(T)$, $|T| > k_c + 1$ is the only situation in which it might be true that for any $z \in T/\{x,y\}$, $z = c(T')$ for some T' with $|T'| > k_c$ but $z \neq c(T' \cup \{y\})$. However, if it does happen in this situation then $|T/\{x\}| > k_c$ and $c(T/\{x\}) = \emptyset$, which is a contradiction. Thus, if $x = c(T)$, there exists a certain $z_T \in T/\{x,y\}$ such that yPz_T cannot occur.

We now consider a new elimination rationale P' that inherits P except xPy. Specifically, $T \supset \{x,y\}$ and $|T| > k_c$, we drop the elimination relation xPy and replace it with (1) $c(T)P'y$ when $x \neq c(T)$ and (2) $z_T P' y$ when $x = c(T)$, where $z_T \in T/\{x,y\}$ is such that yPz_T cannot occur. We finally show that $(k_c, P', \succ) \in \mathcal{R}(c)$. P' is obviously asymmetric and $|max(S, P')| \leq k_c$ for $|S| > k_c$. We prove $|max(S, P')| > 0$ for $|S| > k_c$ by contradiction. Suppose that by contradiction replacing xPy with $z_T P' y$ makes $z_T P' y P' x_1 P' \cdots P' x_{k_c - 1} P' z_T$ a cycle of distinct elements with length larger than k_c. Then $yPx_1 P \cdots P x_{k_c-1} P z_T$, which implies that y is chosen in the choice set consisting of these $k_c + 1$ elements. This contradicts $y \neq c(S)$ for any $|S| > k_c$. We now show that $c(S) = max(\Gamma^{P'}_{k_c}(S), \succ)$. Obviously, $c(S) \in \Gamma^{P'}_{k_c}(S)$ since $c(S) = max(\Gamma^{P}_{k_c}(S), \succ)$. Finally, consider any other option $w \in \Gamma^{P'}_{k_c}(S)$, $w \in \Gamma^{P}_{k_c}(S)$ by the definition of P'. So $c(S) \succ w$ and $c(S) = max(\Gamma^{P'}_{k_c}(S), \succ)$. This establishes that $(k_c, P', \succ) \in \mathcal{R}(c)$ but $xP'y$ does not occur, which contradicts the definition of $xP^c y$. Thus, (2) must hold.

A.8 Characterization of Choice with Top-k Shortlisting

Definition 9. *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 $S \in \mathcal{X}$:*

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

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 10. *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 $S \in \mathcal{X}$,*

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

We firstly 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 6 (Size of Top-k Shortlist). *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. Assume that a choice function is rationalizable by a top-k shortlisting for some k. When the choice function satisfies 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 2 illustrates that $k = 2$ is possible when $N = 4$. □

Theorem 6 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 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.

Corollary 1 (Existence and Uniqueness of Top-k Shortlisting). *Consider $N \neq 4$. A choice function is rationalizable by a top-k shortlisting for some number k if and only if either it satisfies WARP or it has a unique WARP-violating choice pair in which X is the WARP-violating choice set. If a choice function violates the WARP but is rationalizable by a top-k shortlisting, then $k = N-1$.*

Proof. The "if" part of the claim about existence comes from the proof of Theorem 6 and the claim about uniqueness is evident. We now establish the "only if" part of the claim about existence. Suppose that a choice function is rationalizable by a top-

k shortlisting. If it is rationalizable by a top-N shortlisting, then it satisfies WARP. If it is not rationalizable by a top-N shortlisting, then it it is not rationalizable by a top-1 shortlisting. Theorem 6 implies 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. □

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 $N = 4$, a characterization of the behavioral implications for a choice function to be rationalizable by a top-k shortlisting is lengthy. Instead, we list below all choice patterns that are rationalizable by a top-k shortlisting for some number k.

Example 2 (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*

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

may define the preference relation 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 WARP holds. In addition, when 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, which is defined as an option that is chosen over any other option in binary choice problems.

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.

Bajraj, G. and L. Ulku (2015): "Choosing Two Finalists and the Winner," *Social Choice and Welfare*, 45, 729–744.

Chambers, C. P. and M. B. Yenmez (2018): "A Simple Characterization of Responsive Choice," *Games and Economic Behavior*, 111, 217–221.

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

Dardanoni, V., P. Manzini, M. Mariotti, and C. Tyson (2018): "Inferring Cognitive Heterogeneity from Aggregate Choices," Unpublished.

de Clippel, G., K. Eliaz, and K. Rozen (2014): "Competing for Consumer Inattention," *Journal of Political Economy*, 122, 1203–1234.

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.

Dutta, R. and S. Horan (2015): "Inferring Rationales from Choice: Identification for Rational Shortlist Methods," *American Economic Journal: Microeconomics*, 7, 179–201.

Eliaz, K., M. Richter, and A. Rubinstein (2011): "Choosing the Two Finalists," *Economic Theory*, 46, 211–219.

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.

Gerasimou, G. (2018): "Indecisiveness, Undesirability and Overload Revealed Through Rational Choice Deferral," *Economic Journal*, 128, 2450–2479.

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.