Shortlisting with a Limited Capacity*

Sen Geng† Erkut Y. Özbay‡

June 16, 2020

Abstract

A decision maker, who is overwhelmed by the number of available alternatives, limits her consideration. We investigate a model where a decision maker's capacity determines whether she is overwhelmed or not: she considers all the available alternatives if the number of available alternatives does not exceed her capacity; otherwise, she applies the shortlisting heuristic to reduce the number of alternatives to be within her capacity. We show how one can deduce both the decision maker's capacity, preference and the alternatives to which she considers from the observed behavior. Further, we provide the necessary and sufficient conditions for a consideration set to be derived by the choice with limited capacity. (JEL D01)

Keywords: consideration set; shortlisting; limited capacity; revealed preference

*We thank Andrew Caplin, Mark Dean, Emel Filiz Ozbay, Brett Graham, Bart Lipman, Yusufcan Masatlioglu, David Pearce, Ariel Rubinstein, Yongsheng Xu, and seminar attendants at Duke University, New York University, Olin Business School at University of Washington, Tsinghua University, University of Toronto, and Xiamen University for their helpful comments. This paper subsumes Chapter three of Sen Geng's doctoral dissertation. Sen Geng thanks the financial support from Fundamental Research Funds for the Central Universities (Grant No. 20720161035), National Science Foundation of China (Grant No. 71703134), and Natural Science Foundation of Fujian Province (Grant No. 2017J01135).

†Wang Yanan Institute for Studies in Economics (WISE), Xiamen University. E-mail: sg1670@xmu.edu.cn

‡Department of Economics, University of Maryland. E-mail: ozbay@umd.edu.

1. Introduction

Say, you want to purchase a perfume, but it turns out to be there too many alternatives. According to the classical choice theory, a decision maker (DM) considers all the perfumes and chooses the best one. However, especially, the marketing literature provides convincing evidence against the full consideration assumption. Instead consumers use a two-stage procedure where first they decide which alternatives to consider and then they choose the best one among the ones they consider.[1] Theoretically, Manzini and Mariotti (2007) investigates a shortlisting procedure where the DM uses a rationale and she does not consider an alternative that is dominated by another alternative based on that rationale.

In our setup, the DM has a limited capacity that determines the maximal number of alternatives she can consider. When the number of alternatives is within the DM's capacity, she considers all of them as in the classical choice theory. However, when the number of alternatives exceeds her capacity, she uses her rationale to create her shortlist by eliminating the dominated alternatives as in Manzini and Mariotti (2007). Additionally the number of shortlisted alternatives must be within her capacity. In a sense, the capacity is a switch between the classical choice theory and the shortlisting model. We call this procedure as *shortlisting with capacity-k*.

Numerous factors may contribute to the limited capacity of the DM. It may be due to the physical limitation of the DM. For example, the DM may keep in her mind only a limited number of alternatives (Miller, 1956) or the DM may experience an olfactory fatigue after smelling certain number of perfumes. Alternatively, the capacity may be imposed by some exogenous factors. For example, the DM may have a limited time to decide and within this limited time it is impossible to consider all the alternatives (Geng, 2016). Or, simply the capacity may arise as a result of a trade-off between search cost and benefits (Stigler, 1961); and by setting a capacity, the DM may reduce the decision cost (Ergin and Sarver, 2010; Ortoleva, 2013).

Our paper contributes to the literature that incorporates limited capacity into limited

[1] The limited considerations have extensively shown in various markets, such as in investment decisions (Huberman and Regev, 2001), in school choice (Rosen et al., 1998), in job search (Richards et al., 1975), in household grocery consumption (Demuynck and Seel, 2018), or in PC purchases (Goeree, 2008).

consideration models. For example, in Salant and Rubinstein (2008) the capacity can be any number but it is assumed to be observed.[2] In Eliaz et al. (2011), under the consideration set interpretation, the DM narrows down the consideration set to two alternatives. In Bajraj and Ülkü (2015), the DM uses a shortlisting procedure to determine the top two alternatives to consider. To our knowledge, ours is the first paper that identifies the capacity from the choice data. It turns out to be that violation of Weak Axiom of Revealed Preference (WARP) is enough to uniquely identify the capacity. In the classical choice theory, binary choices reveal the preference. In shortlisting with capacity-k, since the full consideration assumption is relaxed, binary choices reveal the preference when WARP is violated. Additionally, we study how the choices reveal the rationale and the shortlisted alternatives.

We also study the link between the heuristics that are used to form consideration sets and the properties of the consideration sets. Observing the consideration sets has been particularly important for the marketing literature especially when the entire choice data is not available (Dardanoni et al., 2018; de Clippel et al., 2014). Recently, the development of the new tools, such as eye tracking, makes it possible to observe the consideration sets (Reutskaja et al., 2011), additionally, in many real-life situations, consideration sets are also observable especially in online stores (Manzini et al., 2019). We provide the necessary and sufficient properties on the consideration sets to be shortlisting with capacity-k. It turns out to be that the competition filter property in Lleras et al. (2017) is quite general and shortlisting with capacity-k refines the competition filter.[3]

We additionally provide a choice based characterization. We relax the existing axioms so that we show the relation between shortlisting with capacity-k and Manzini and Mariotti (2007). In Discussion section, we revisit different choice reversals and compare the degree to which different models can explain these choice reversals. Finally, we investigate a special case of shortlisting with capacity-k, where the DM always exhausts her capacity.[4]

[2]See also Chambers and Yenmez (2018).

[3]The competition filter in Lleras et al. (2017) is based on the idea of competition among alternatives to be considered by an overwhelmed DM: if an alternative is considered in a larger set, then it should be considered as some of alternatives are removed.

[4]Although the DM considers all the alternatives up to her capacity, beyond her capacity, she starts using the shortlisting procedure which may eventually reduce the number of alternatives to be considered well below her capacity.

For example, the DM with capacity-k sorts all the alternatives based on an ordering and she considers the *top-k* alternatives when there are more than k alternatives. It turns out to be that this case imposes a surprisingly tight structure on choice behavior in the sense that at most one violation of WARP is allowed.

The rest of the paper will be structured as follows: in Section 2, we introduce the limited capacity model. In Section 3, we show how the choices reveal capacity, preference, shortlist and conflicting rationale. In Section 4, we provide the characterizations of shortlisting with limited capacity. In Section 5, we discuss the related literature and explore the special case of top-k shortlisting, and Section 6 concludes. All proofs are in the appendix.

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 equal to or less than her capacity.[5]

Motivated by the marketing literature, individuals use some heuristics to reduce the number of options to be considered. The most widely investigated heuristic is the

[5]When a DM faces a choice set beyond her capacity and relies on an exogenous filtering tool to shortlist alternatives, e.g., using filtering tool in online shopping, it is likely that the number of alternatives that are kept in the shortlist is strictly smaller her capacity once she employs the exogenous filtering tool. In settings where a DM may endogenously determine the shortlist of alternatives, while we are not aware of empirical studies that explore how a DM's consideration set varies as the size of choice set increases, there is some indirect evidence suggesting that the possibility exists that a DM facing a larger choice set considers less. For example, Dean et al. (2017) demonstrates experimentally that individuals choose a default alternative more frequently as the choice set expands. In addition, marketing studies about choice overload provide real-world evidence at the store level that larger assortments decrease overall sales (Boatwright and Nunes, 2001; Dreze et al., 1994) .

shortlisting heuristic by Manzini and Mariotti (2007). According to shortlisting, the DM uses a rationale to eliminate some of the alternatives; and the undominated alternatives according to her rationale are shortlisted.[6] 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 $max(S, P) := \{x \in S |$ no $y \in S$ such that $yPx\}$ is the set of undominated options in S with respect to P. Hence, when the number of options in S is greater than her capacity k, the DM eliminates the dominated options in S with respect to a rationale P.

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

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

and $0 < |max(S, P)| \leq k$ if $|S| > k$.[7]

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, $\succ \subseteq X \times X$.[8]

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 2. *A choice function, c, is* ***rationalizable by a shortlisting with capacity-k*** *if*

[6] In Geng and Ozbay (2020), we take an alternative approach to modeling consideration function with capacity-k: we assume that the consideration function satisfies a general property without specifying the function form.

[7] To rule out empty choice, 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, where width(X, P):=$sup\{|Y|$: Y is a choice set in which any two options are not related according to P.$\}$. Alternatively, Dutta and Horan (2015), and Horan (2016) require P to be acyclic.

[8] Manzini and Mariotti (2007) requires this second relation only to be asymmetric, we put the stronger requirement on $\succ$ to be a linear order as in Lleras et al. (2017) in order to highlight that the DM is rational among the options she considers (see also Au and Kawai 2011, Horan 2016 and Yildiz 2016).

there exists a capacity k, a rationale P, and a linear order $\succ$, such that for any $S \in \mathcal{X}$,

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

3. Identification

In the classical choice theory, due to full consideration assumption, x is revealed preferred to y if y is never chosen when x is available. Recall that WARP is the necessary and the sufficient condition for the revealed relation to be a preference relation. In terms of choice functions, WARP is equivalently stated as an alternative can never be chosen in the presence of another alternative that is chosen in a binary choice problem of these two alternatives: $x = c(\{x, y\})$ implies that $y \neq c(S)$, for any $S \supset \{x, y\}$.

Our model consists of three components: capacity-k, rationale, and preference. In this section, we investigate if it is possible to reveal capacity-k, rationale and preference based on the observed choices of the DM.

3.1. Revealed Capacity-k and Revealed Preference

The classical representation theorem can be restated in our setup as a choice function satisfies WARP if and only if it is rationalizable by a shortlisting with capacity-N since the capacity-N corresponds to considering all the feasible alternatives. Interestingly, it is also equivalent to capacity-1 since a DM with capacity-1 shares the same choice function with a DM with capacity-N as long as the rationale employed by the first DM coincides with the second DM's preference relation.

Proposition 1. *For a choice function, c, the following statements are equivalent:*

(i) c satisfies WARP.

(ii) c is rationalizable by a shortlisting with capacity-N.

(iii) c is rationalizable by a shortlisting with capacity-1.

As in any two-stage choice model, the choice function that satisfies WARP is the least informative one (Masatlioglu et al., 2012). In our setup (as shown in Proposition 2), this means that it is rationalizable by a shortlisting with any capacity.

Proposition 2. *If c satisfies WARP, then, c is rationalizable by a shortlisting with capacity-k for any k.*

By Proposition 2, if c satisfies WARP, 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 to the unchosen option. For example, when observing $x = c(\{x, y\})$ and $z = c(\{z, w\})$, we can conclude that either x is preferred to y and z is preferred to w, or x eliminates y and z eliminates w, as long as WARP holds.

We now look at choice functions that violates WARP. A choice function violates WARP if there exist x, y and $S \supset \{x, y\}$ such that $x = c(\{x, y\})$ but $y = c(S)$. We will call such x, y as **WARP-violating choice pair**, such a set S as **WARP-violating choice set**, and such c as **WARP-violating choice function**. Correspondingly, we define **overwhelming set** as a set that includes alternatives no less than a WARP-violating choice set includes:

$$\mathcal{O} = \{T : \exists \text{ a WARP-violating choice set }, S, \text{ such that } |S| \leq |T|\}$$

Note that if a WARP-violating choice function is rationalizable by a shortlisting with capacity-k, the capacity k cannot be 1 or N, i.e. $2 \leq k \leq N-1$. Since the capacity must be at least two, the DM always chooses the option she prefers in any binary choice problem. Hence, a WARP-violating choice function that is rationalizable by a shortlisting with capacity-k directly reveals the preference $\succ$ based on the choice in binary comparisons as in the standard case:

$$x \succ y \text{ iff } x = c(\{x, y\})$$

Additionally, for any WARP-violating choice set, S, the DM should not be able to consider all the alternatives in S. Otherwise, she should have chosen the best one consistent with her choice in binary comparisons. So, the capacity, k, should be less than the cardinality of S. For a WARP-violating choice function, c, let's define the **threshold capacity** as:

$$\mathbf{k_c} := \min\{|S| - 1 : S \text{ is a WARP-violating choice set given the choice function, } c.\}$$

This definition suggests that k_c is exactly one less than the cardinality of the smallest WARP-violating choice set. So, the capacity, k, should be less than or equal to k_c. We show in Proposition 3 that the capacity, k, is unique and it is equal to the threshold capacity.

Proposition 3 (**Uniqueness of Capacity**). *If a WARP-violating choice function is rationalizable by a shortlisting with capacity-k, then k is unique and* $k = k_c$.

3.2. Revealed Shortlist and Revealed Rationale

First, we look at the revealed shortlist. We use a conservative definition for an alternative revealed to be shortlisted in a set as in Masatlioglu et al. (2012). This conservative definition does not depend on the specific elimination rationale.

Definition 3. *Let $\mathcal{R}(c)$ denote the collection of all triples that rationalize c. We say x is revealed to be shortlisted at S if $x \in \Gamma_k^P(S)$ for any $(k, P, \succ) \in \mathcal{R}(c)$.*

Clearly, if c satisfies WARP then only the chosen alternative is revealed to be shortlisted at a choice set. If c violates WARP and it is rationalizable by a shortlisting with limited capacity, the capacity number must be equal to the threshold capacity k_c according to Proposition 3. Hence, the revealed shortlist of a choice set with a size no more than k_c

must be equal to the choice set. However, for a choice set with a size exceeding k_c, i.e., for an overwhelming set, it may have several shortlists that can generate the same choice behavior. We observe that if an alternative is chosen over any other alternative of S in an overwhelming set, the alternative should never be eliminated in S and in turn it is revealed to be shortlisted at S. We show in Proposition 4 that this observation is not only sufficient but also necessary for the alternative revealed to be shortlisted at S. The proposition suggests that at least one alternative is revealed to be shortlisted at S, i.e., $c(S)$. Additionally, there may be multiple alternatives that are revealed to be shortlisted at S. For example, if $x = c(S)$ and $y = c(S')$, where $\{x, y\} \subseteq S \cap S'$ and $S, S' \in \mathcal{O}$, then $\{x, y\}$ are both revealed to be shortlisted at S.

Proposition 4 (**Revealed Shortlist**)**.** *Suppose c is rationalizable by a shortlisting with capacity-k. If c satisfies WARP, then x is revealed to be shortlisted at S if and only if $x = c(S)$. If c violates WARP, then for any $S \notin \mathcal{O}$, x is revealed to be shortlisted at S for any $x \in S$; and for any $S \in \mathcal{O}$, x is revealed to be shortlisted at S if and only if for any $y \in S$ there exists $T \in \mathcal{O}$ such that $\{x, y\} \subset T$ and $x = c(T)$.*

Next, we look at the revealed rationale. Assume S is a WARP-violating set, i.e., $x = c(\{x, y\})$ and $x \in S$ but $y = c(S)$. The revealed preference argument suggests that $x \succ y$ and x should be eliminated in S. The question is which alternative eliminates x. It may sound reasonable that y eliminates x. Formally, define, the trivial rationale, yPx if there exists S such that $x = c(\{x, y\})$ and $x \in S$ but $y = c(S)$. It turns out that the trivial rationale is neither necessary (see Example 1) nor sufficient (see Example 2).

Example 1. *$c(\{x_1, x_2\}) = c(\{x_1, x_3\}) = c(\{x_1, x_4\}) = x_1$, $c(\{x_2, x_3\}) = c(\{x_2, x_4\}) = x_2$, and $c(\{x_3, x_4\}) = x_3$.*

	$\{x_1, x_2, x_3, x_4\}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$
c	x_2	x_2	x_1	x_3	x_2

Since $x_1 = c(\{x_1, x_2\})$ but $x_2 = c(\{x_1, x_2, x_3\})$, based on the trivial rationale x_2Px_1. However, this is inconsistent with $x_1 = c_1(\{x_1, x_2, x_4\})$. Hence, the trivial rationale is not necessary; c is rationalizable by a shortlisting with capacity-2, defining x_3Px_1 and x_2Px_4.

Example 2. *$c(\{x_1, x_2\}) = c(\{x_1, x_3\}) = c(\{x_1, x_4\}) = x_1$, $c(\{x_2, x_3\}) = c(\{x_2, x_4\}) = x_2$, and $c(\{x_3, x_4\}) = x_3$.*

	$\{x_1, x_2, x_3, x_4\}$	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$
c	x_1	x_1	x_1	x_1	x_3

The trivial rationale would only suggest that $x_3 P x_2$ since $\{x_2, x_3, x_4\}$ is the only WARP-violating set. Based on this rationale only x_2 would be eliminated in $\{x_1, x_2, x_3, x_4\}$. But since the revealed capacity is 2, with the unique elimination, the choice function c is not rationalizable by a shortlisting with limited capacity. Hence, the trivial rationale is not sufficient; c is rationalizable by a shortlisting with capacity-2, defining $x_3 P x_2$ and $x_1 P x_4$.

Dutta and Horan (2015) provides a good 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. 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.

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. Again, 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 4. *Let $\mathcal{R}(c)$ denote the collection of all triples that rationalize a WARP-violating choice function c. We define the* ***revealed conflicting rationale*** *P^c by $xP^c y$ if $y \succ x$ and xPy for any $(k_c, P, \succ) \in \mathcal{R}(c)$.*

We show in Proposition 5 that the revealed conflicting rationale in Definition 4 has an appealing behavioral characterization: (1) y is chosen over x in the binary choice of the two options; (2) if y is chosen in an overwhelming set then it is never chosen in an overwhelming set that includes x; if y is never chosen in an overwhelming set then there exist a situation in which x has to eliminate y. The intuition of the characterization proceeds as follows. We observe that if neither x nor y is chosen in an overwhelming set then x eliminates y is not necessary. The reason is that in this case for any overwhelming set including y, we could let its chosen alternative eliminate y. This observation implies that a necessary condition for xP^cy is one of them must be chosen in an overwhelming set. Additionally, whether y is chosen or not in an overwhelming set is a rigid dichotomy. If y is chosen in an overwhelming set, a necessary and sufficient condition for xP^cy is that y is not chosen again in any overwhelming set including x since we already assumed that y is revealed preferred to x. If y is never chosen in an overwhelming set, a necessary condition for xP^cy is that x is chosen in an overwhelming set. There are two sub-cases for this case: (i) x is never chosen in an overwhelming set that includes y; (ii) x is chosen in an overwhelming set that includes y. In sub-case (i), it is perfectly fine to define yPx as shown in the proof of Proposition 7. That says, x eliminates y is not necessary in sub-case (i). In sub-case (ii), a necessary condition for xP^cy is that y cannot be eliminated by any other alternative except x in the overwhelming set. This further implies that for any other alternative z in the overwhelming set, it must be that z is chosen in an overwhelming set but it is never chosen in an overwhelming set including y. Otherwise, either z is never chosen in an overwhelming set or z is chosen in an overwhelming set including y, and in the two situations it is not a must to let y eliminate z so it will be fine to always let the corresponding z eliminate y instead of letting x eliminate y. Thus, if x has to eliminate y in sub-case (ii), it must be that for any other alternative z in the overwhelming set, z is chosen in an overwhelming set but it is never chosen in an overwhelming set including y.

Proposition 5 (**Revealed Conflicting Rationale**). *Suppose that a WARP-violating choice function c is rationalizable by a shortlisting with limited capacity. Then xP^cy if and only if $y = c(\{x, y\})$ and the following two statements hold:*

(1) If there exists $S \in \mathcal{O}$ such that $y = c(S)$, then $y \neq c(T \cup \{x\})$ for any $(T \cup \{x\}) \in \mathcal{O}$,

(2) If there is no $S \in \mathcal{O}$ such that $y = c(S)$, then $x = c(T \cup \{y\})$ for some $(T \cup \{y\}) \in \mathcal{O}$ where for any $z \in T/\{x, y\}$, there exists $T' \in \mathcal{O}$ such that $z = c(T')$ but $z \neq c(T' \cup \{y\})$.

4. Characterization

4.1. Based on Observed Consideration Sets

Instead of applying a heuristic to form the consideration sets, an alternative approach is to impose some properties on the consideration sets without committing to a procedure (Lleras et al., 2017). Marketing literature has devoted considerable efforts to understanding the formation of consideration sets and has developed tools to observe consideration sets (Hauser, 2014; Reutskaja et al., 2011). In this subsection we study if it is possible to identify an unobserved heuristic from observed consideration sets.

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 competition filter that is defined in Lleras et al. (2017) requires a consideration function Γ satisfying that if $x \in \Gamma(T)$ and $x \in S \subset T$, then $x \in \Gamma(S)$.[9] It is straightforward to establish that the competition filter is equivalent to the More is Less property. The competition filter provides a structure as the sets get smaller. However, it is quite flexible as the sets get bigger. Next, we investigate the additional properties on the consideration functions that are consistent with the shortlisting with limited capacity procedure. We introduce below two properties that provide a structure as the sets get bigger.

First, let's define two classes: the class of feasible sets with full consideration $\mathcal{FC}$ and the class of feasible sets with limited consideration $\mathcal{LC}$:

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

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

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

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.[10]

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 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.

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

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

Proposition 6 shows that when a consideration function satisfies the above four properties, it is equivalent to a shortlisting with capacity-k.

Proposition 6. *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.*

4.2. Based on Observed Choices

The purpose of this subsection is to find axioms that we can relate with the other limited consideration models. The overwhelming choice (Lleras et al. (2017)), rationalization (Cherepanov et al. 2013), categorization (Manzini and Mariotti 2012a) are characterized by a single axiom of Weak WARP introduced by Manzini and Mariotti (2007). According Weak WARP, 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:

Axiom 1 (Weak WARP). *Consider $\{x, y\} \subset S \subset T$. $x = c(\{x, y\}) = c(T)$ implies that $y \neq c(S)$.*

In our setup, if there is a WARP-violating set, the capacity 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 2 (No Binary Cycles). *$x = c(\{x, y\})$ and $y = c(\{y, z\})$ implies that $x = c(\{x, z\})$.*

Manzini and Mariotti (2007) also proposes the axiom of Expansion that an option chosen from each of two sets is also chosen from their union. We observe that a choice function satisfies Expansion and No Binary Cycles if and only if it satisfies WARP. 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 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).

We first assume that the property of Expansion holds when all the sizes of the involved choice sets exceed the threshold capacity, i.e., when the involved sets are overwhelming sets. The motivation for this assumption is that when all the sizes of the involved choice sets exceed the threshold capacity, they all exceed the DM's capacity and then the elimination machine is always operating.

Axiom 3 (Limited Expansion). *For any* $S, T \in \mathcal{O}$, $x = c(S) = c(T)$, *then* $x = c(S \cup T)$.

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

Axiom 4 (Limited Weak Expansion). *For any* $S, T \in \mathcal{O}$, $x = c(S), y = c(T)$, *then* $x = c(S \cup \{y\})$ *or* $y = c(T \cup \{x\})$.

It is notable that the axiom of Limited Weak Expansion 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 $c(T) = c(T \cup \{c(S)\})$.

Finally, we highlight a behavioral implication that limited capacity has on elimination rationale. On the one hand, choice by shortlisting with limited capacity requires that the size of any choice set after elimination must fall within a DM's capacity and in turn fall within the threshold capacity. On the other hand, it seems that no elimination relation can exist between two options if both are chosen in the presence of each other in overwhelming choice sets:

Definition 5. *Alternatives x and y are* ***mutually undominated*** *if there exists sets S', S'' with $S', S'' \in \mathcal{O}$, $\{x, y\} \subset S' \cap S''$ such that $x = c(S')$ and $y = c(S'')$. A set M is a* ***set of mutually undominated alternatives*** *if any two distinct elements in the set are mutually undominated.*

One could directly assume that for any set of mutually undominated alternatives, its size falls within the threshold capacity. Nevertheless, the assumption is not minimal given that we have imposed the axiom of No Binary Cycles. More precisely, a choice function satisfies No Binary Cycles and Expansion up to a certain size of choice sets if and only if it satisfies WARP up to the same size of choice sets. The latter one further implies that the size of such choice sets must fall within the threshold capacity. This motivates us to assume that when there exists a set of mutually undominated alternatives, the property of Expansion holds up to the same size of choice sets. In this way, the requirement that the size of such choice sets falls within the threshold capacity is implied instead of being assumed.

Axiom 5 (Conditional Expansion). *Let M be a set of mutually undominated alternatives, then $x = c(S) = c(T)$ and $|S \cup T| \leq |M|$ implies that $x = c(S \cup T)$.*

Our characterization result (i.e., Proposition 7) shows that the axiom of No Binary Cycles, the axiom of Weak WARP, and the axioms of Expansion in the above three domains exactly capture the essence of choice by a shortlisting with limited capacity.

Proposition 7 (Characterization). *A choice function is rationalizable by a shortlisting with capacity-k for some k if and only if it satisfies Weak WARP, No Binary Cycles, Limited Expansion, Limited Weak Expansion and Conditional Expansion.*

The intuition of the proof is as follows. The axiom of No Binary Cycles serves to identify the preference relation. WARP violations along with the definition of threshold capacity serve to identify the DM's capacity. The axiom of Limited Weak Expansion serves to guarantee the asymmetry of elimination rationale. The axiom of Weak WARP and the axiom of Limited Expansion guarantee that the chosen option in a large choice set must be preferred to any other option in the consideration set of the choice set. Finally, the axiom of Conditional Expansion serves to guarantee that the size of any choice set after elimination will fall within the DM's capacity. Hence, separating Expansion into three domains enables to see how the proof works.[11]

5. Discussion

5.1. Comparison of different choice models

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

Table 2 illustrates the difference between the limited capacity model and some of existing choice models in terms of explaining choice behavior when $N = 3$. Masatlioglu et al. (2012), Manzini and Mariotti (2012a), Cherepanov et al. (2013), and Lleras et al. (2017)

[11] Alternatively, we establish that these three axioms can be replaced by one single axiom: Let M be a set of mutually undominated alternatives and O be an overwhelming set. $|M| < |O|$, and if either $S, T \in \mathcal{O}$ or $(S \cup T) \notin \mathcal{O}$, 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 $c(T) = c(T \cup \{c(S)\})$.

Table 1: What can/cannot be explained by shortlisting with capacity-k

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

are not falsifiable when $N = 3$. 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 No Binary Cycles.

Table 2: Comparison of different choice models

label	Masatlioglu et al. (2012), Manzini & Mariotti (2012) Cherepanov et al. (2013), Lleras et al. (2017)	Manzini & Mariotti (2007)	Shortlisting with capacity-k
standard choice	✓	✓	✓
attraction effect	✓	✗	✓
choosing pairwisely unchosen	✓	✗	✓
cyclical binary choice	✓	✓	✗
falsifiable when $N = 3$	✗	✓	✓

We know from Proposition 6 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 (2012a) and the *rationalization* choice model in Cherepanov et al. (2013). Table 2 shows that interestingly there is no nested relationship between the shortlisting with limited capacity model and the *shortlisting* choice model of Manzini and Mariotti (2007).

Finally, next example demonstrates 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 3. *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 $x_4 \neq x^*$.*

5.2. A Special Case: Choice with Top-k Shortlisting

In our main setup, the DM may not need to use her capacity fully. Now, we look at the special case where the DM needs to use her full capacity when she can. The DM sorts the alternative based on a ranking and always consider top-k available alternatives. So, when there are abundance of alternatives, the number of considered alternatives, is equal to the capacity of the DM.

Definition 6. *A consideration function, Γ, is called a **top-k shortlisting** if there exists a rationale P such that for each feasible set $S \in \mathcal{X}$:*

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

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

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

Definition 7. *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 capacity of the top-k shortlist, i.e., the number k, can take only a few values. In fact, the capacity of the shortlist depends on the size of the grand set, N, where $N \geq 3$.

Proposition 8 (Capacity of Top-k Shortlisting). *If 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 $N = 4$.*

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 $N > 4$.

Proposition 8 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 the characterization for a choice function to be rationalizable by a top-k shortlisting in Appendix B.. It turns out that top-k shortlisting imposes a surprisingly tight structure on choice behavior in the sense that at most one choice reversal can be allowed.[12]

Finally, we discuss revealed preference and revealed shortlist for a WARP-violating choice with top-k shortlisting. Clearly, the definition applies that a DM has a revealed preference x over y if $x = c(\{x, y\})$. When $N \neq 4$, 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.

6. Conclusion

We investigate a two-stage procedure where a 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 the shortlisting heuristic of Manzini and Mariotti (2007), the DM limits the number of alternatives to consider to be within her capacity and then chooses the best one.

In the determination of the capacity of a DM, WARP violation is the key. To illustrate this point, say we observe a choice reversal ($x = c(\{x, y\})$ but $y = c(S)$). This reveals that $|S|$ is beyond the capacity of a DM. The largest set in which we don't observe choice reversals determines the threshold capacity. Beyond the threshold capacity, it is too many for a DM so that she will be overwhelmed. We show that the threshold capacity of the choices is the unique capacity of the DM. Additionally, since the capacity of a DM should be at least two, the choice based on binary choice problems reveals the preference as in the classical choice theory.

We also provide two characterizations based on (i) consideration sets and (ii) choice data. With the characterization based on choice data, we demonstrate minimal modifications to well-known axioms which enable to relate our model with the well-known models. The characterization based on consideration sets provides a novel perspective to the literature.

[12] Choosing pairwisely unchosen is now ruled out in top-k shortlisting.

Suppose we cannot observe the entire choice data, but we observe the consideration sets instead. We provide the necessary and sufficient properties of a consideration set to conclude that it is generated by a shortlisting with limited capacity. Additionally, in the environments in which both choice data and consideration set data are available, our characterizations will allow the further test of the validity of the conclusions based only on choice data. One can also check whether the DM is choosing the best alternative from her consideration set; or one can deduce whether the DM has a capacity by looking at the consideration set data even when the choice is consistent with WARP.

Alternatively, when the DM is overwhelmed, the DM may be using an alternative heuristic than the shortlisting such as the rationalization of Cherepanov et al. (2013) or categorization of Manzini and Mariotti (2012a). In Geng and Ozbay (2020), we investigate the limited capacity under these heuristics. It turns to be that although the capacity is uniquely identified under shortlisting, the threshold capacity is an upper limit of the capacity under these heuristics.

An interesting but not trivial extension of our setup is multiple rationales (see Manzini and Mariotti 2012b, Apesteguia and Ballester 2013). 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 studies.

References

Apesteguia, J. and M. A. Ballester (2013): "Choice by Sequential Procedures," *Games and Economic Behavior*, 77, 90 – 99.

Au, P. H. and K. Kawai (2011): "Sequentially Rationalizable Choice with Transitive Rationales," *Games and Economic Behavior*, 73, 608–614.

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

Boatwright, P. and J. C. Nunes (2001): "Reducing Assortment: an Attribute-based Approach," *Journal of Marketing*, 65, 50–63.

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.

Dean, M., Ö. Kibris, and Y. Masatlioglu (2017): "Limited Attention and Status Quo Bias," *Journal of Economic Theory*, 169, 93–127.

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

Dreze, X., S. J. Hoch, and M. E. Purk (1994): "Shelf Management and Space Elasticity," *Journal of Retailing*, 70, 301–326.

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.

GENG, S. AND E. Y. OZBAY (2020): "How Many is Too Many," Unpublished.

GOEREE, M. S. (2008): "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica*, 76, 1017–1074.

HAUSER, J. R. (2014): "Consideration-Set Heuristics," *Journal of Business Research*, 67, 1688–1699.

HORAN, S. (2016): "A Simple Model of Two-Stage Choice," *Journal of Economic Theory*, 162, 372–406.

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.

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

——— (2012b): "Choice by lexicographic semiorders," *Theoretical Economics*, 7, 1–23.

MANZINI, P., M. MARIOTTI, AND L. ÜLKÜ (2019): "Sequential approval: A Model of Likes, Paper Downloads and Other Forms of Click Behaviour," Unpublished.

MASATLIOGLU, Y., D. NAKAJIMA, AND E. Y. 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.

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

Yildiz, K. (2016): "List-Rationalizable Choice," *Theoretical Economics*, 11, 587–599.

Appendix A. Proofs

Proof of Proposition 1

Proof. $(i) \implies (iii)$: Let c satisfy WARP. Then there exists a $\succ$ that rationalizes c. Set $P = \succ$. So, for any non-empty S, $\Gamma_k^P(S) = \{c(S)\}$, hence, $|\Gamma_k^P(S)| = 1$, i.e. $k = 1$.

$(iii) \implies (i)$: Let c be a shortlisting with capacity-1. Let $x = c(\{x, y\})$. Then xPy since $k = 1$. So, if $x \in S$, $y \neq \Gamma_1^P(S)$, hence, $y \neq c(S)$.

$(i) \iff (ii)$ is the standard case. □

Proof of Proposition 2

Proof. Define xPy and $x \succ y$ if $x = c(\{x, y\})$. Then $\Gamma_k^P(S) = \{c(S)\}$ when $|S| > k$ and $\Gamma_k^P(S) = S$ when $|S| \leq k$. Hence, if $|S| > k$, then $\max(\Gamma_k^P(S), \succ) = \max(\{c(S)\}, \succ) = c(S)$. If $|S| \leq k$, then $\max(\Gamma_k^P(S), \succ) = \max(S, \succ) = c(S)$ since WARP guarantees that $c(S) \succ y$ for any $y \in S$. □

Proof of Proposition 3

Proof. If a WARP-violating choice function c is rationalizable by a shortlisting with capacity-k, then $2 \leq k \leq k_c$. In the proof of Proposition 7, we show 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. $\square$

Proof of Proposition 4

Proof. Assume that c violates WARP and $S \in \mathcal{O}$. Suppose for any $y \in S$ there exists $T \in \mathcal{O}$ such that $\{x, y\} \subset T$ and $x = c(T)$. Then for any $(k_c, P, \succ) \in \mathcal{R}(c)$, $x \in \Gamma^P_{k_c}(T)$, i.e., x is revealed to be shortlisted at S.

Now assume that x is revealed to be shortlisted at S, i.e., $x \in \cap_{i=1}^n \Gamma^{P_i}_{k_c}(S)$. We firstly show that this implies that there exists $T \in \mathcal{O}$ such that $x = c(T)$. Suppose by contradiction that $x \neq c(T)$ for any $T \in \mathcal{O}$. Let P be the elimination rationale defined in the proof of Proposition 7. Then on the one hand, $x \in \cap_{i=1}^n \Gamma^P_{k_c}(S)$. On the other hand, $c(T)Px$ according to the definition of P. This is a contradiction.

We then establish that for any $y \in S$ there exists $T \in \mathcal{O}$ such that $\{x, y\} \subset T$ and $x = c(T)$. Suppose by contradiction that there exists a certain $y \in S$ such that for any $T \in \mathcal{O}$ such that $\{x, y\} \subset T$ and $x \neq c(T)$. Then, yPx according to the definition of P in the proof of Proposition 7. This is a contradiction. $\square$

Proof of Proposition 5

Proof. Assume that the conditions hold and call condition $y = c(\{x, y\})$ as A. When there exists $S \in \mathcal{O}$ such that $y = c(S)$, A and (1) implies that $xP^c y$ because otherwise $y = c(S \cup \{x\})$. When there is no $S \in \mathcal{O}$ such that $y = c(S)$, A and (2) implies that $x = c(T \cup \{y\})$ for some $(T \cup \{y\}) \in \mathcal{O}$ where for any $z \in T/\{x, y\}$, there exists $T' \in \mathcal{O}$ such that $z = c(T')$ but $z \neq c(T' \cup \{y\})$. Consider any $(k_c, P, \succ) \in \mathcal{R}(c)$. Clearly, $y \notin \Gamma^p_{k_c}(T \cup \{y\})$. In addition, no $z \in T/\{x, y\}$ such that zPy. So it must be that xPy.

Now assume that $xP^c y$. A comes from Definition 4. We have the following observation:

Observation 1 *If $xP^c y$, then there exists $S \in \mathcal{O}$ such that either $x = c(S)$ or $y = c(S)$.*

The proof of Observation 1 proceeds as follows. Suppose by contradiction that $x \neq c(S)$ and $y \neq c(S)$ for any $S \in \mathcal{O}$. Let P be the elimination rationale defined in the proof of Proposition 7, where we define the rationale P in two specific situations. Clearly, we do not have xPy there. Since we show there $(k_c, P, \succ) \in \mathcal{R}(c)$, this contradicts $xP^c y$.

Additionally, whether y is ever chosen or not in an overwhelming set is a rigid dichotomy.

Case (a): there exists $S \in \mathcal{O}$ such that $y = c(S)$.

In case (a), it must be that $y \neq c(T \cup \{x\})$ for any $(T \cup \{x\}) \in \mathcal{O}$ because xPy for any $(k_c, P, \succ) \in \mathcal{R}(c)$.

Case (b): $y \neq c(S)$ for any $S \in \mathcal{O}$.

In case (b), Observation 1 implies that there exists $S \in \mathcal{O}$ such that $x = c(S)$. We show that $x = c(T \cup \{y\})$ for some $(T \cup \{y\}) \in \mathcal{O}$ where for any $z \in T/\{x, y\}$, there exists $T' \in \mathcal{O}$ such that $z = c(T')$ but $z \neq c(T' \cup \{y\})$. Consider the following dichotomy: (b1) $x \neq c(T \cup \{y\})$ for any $(T \cup \{y\}) \in \mathcal{O}$ and (b2) there exists $(T \cup \{y\}) \in \mathcal{O}$ such that $x = c(T \cup \{y\})$.

In sub-case (b1), let P be the elimination rationale defined in the proof of Proposition 7. Then yPx, which contradicts $xP^c y$. So sub-case (b1) cannot occur given the assumption.

Now consider sub-case (b2). We only need to show that for any $z \in T/\{x, y\}$, there exists $T' \in \mathcal{O}$ such that $z = c(T')$ but $z \neq c(T' \cup \{y\})$. Suppose by contradiction that this statement is not true. That says, for any $(T \cup \{y\}) \in \mathcal{O}$ such that $x = c(T \cup \{y\})$, there exists $z \in T/\{x, y\}$ such that either (i) $z \neq c(S)$ for any $S \in \mathcal{O}$ or (ii) $z = c(T \cup \{y\})$ for some $(T \cup \{y\}) \in \mathcal{O}$. In the first situation, Observation 1 implies that we do not have to let y eliminates z. In the second situation, we do not have to let y eliminates z either because the elimination rationale defined in the proof of Proposition 7 actually suggests we can let z eliminate y. In summary, we do not have to have y eliminate z in either of the two situations. We then let z eliminate y and let x do not eliminate y. It is straightforward to show that a new rationale with this change still rationalizes c. However, this new rationale violates $xP^c y$. So for any $z \in T/\{x, y\}$, there exists $T' \in \mathcal{O}$ such that $z = c(T')$ but $z \neq c(T' \cup \{y\})$. $\square$

Proof of Proposition 6

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 $|T| > k$ because $max(T, P) = \Gamma(T) \neq \emptyset$. $|max(T, P)| \leq k$ 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 competition 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. $\square$

Proof of Proposition 7

Proof. When a choice function c satisfies WARP, the proof of the proposition is trivial. Therefore, we only need to show that when c is a WARP-violating choice function, it is rationalizable by a shortlisting with capacity-k for some k if and only if it satisfies Axioms 1-5. We assume that c is a WARP-violating choice function in the analysis below. Then by definition of threshold capacity, $2 \leq k_c < N$.

"If" direction. Assume that c satisfies Axioms 1-5. We show that c is rationalizable by a shortlisting with capacity-k_c.

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 $S \in \mathcal{O}$ such that $\{x, y\} \subset S$ but $y \neq c(S')$ for any $S' \in \mathcal{O}$ such that $\{x, y\} \subset S'$, or (2) $y = c(T)$ for some $T \in \mathcal{O}$ but $y \neq c(T' \cup \{x\})$ for any $T' \in \mathcal{O}$ satisfying that $y = c(T')$. Define $\Gamma^P_{k_c}(S) = S$ if $|S| \leq k_c$, and $= max(S, P)$ if $|S| > k_c$. We show that c is rationalizable by $(k_c, P, \succ)$.

Obviously, $\succ$ is complete and asymmetric. Axiom 2 guarantees that $\succ$ is transitive.

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 neither of the two situations of defining yPx can occur. If situation 2 of defining xPy occurs, then clearly situation 1 of defining yPx cannot occur. Then the only situation that makes yPx happen is that $x = c(S)$ for some $S \in \mathcal{O}$ but $x \neq c(S' \cup \{y\})$ for any $S' \in \mathcal{O}$ satisfying that $x = c(S')$. This violates Axiom 4. So P defined above is asymmetric.

The definition of P suggests that if an option is ever chosen in an overwhelming choice set, 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 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 $x_1 = c(S_{NP})$. So $S_{NP} \in \mathcal{O}$. Since there is no P relation between x_1 and x_i for any $i \in \{2, \cdots, k_c + 1\}$, there exists $S_i \in \mathcal{O}$ such that $\{x_1, x_i\} \subset S_i$ and $x_i = c(S_i)$. Since $x_i = c(S_i)$ for some $S_i \in \mathcal{O}$, 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 $S_i^{'} \in \mathcal{O}$ and $x_i = c(S_i^{'})$, there exists $S_i^{'}$ such that $x_i = c(S_i^{'} \cup x_j)$, where $j \in \{2, \cdots, k_c + 1\}$. So S_{NP} is a set of mutually undominated alternatives. Since $|S_{NP}| = k_c + 1$, Axiom 2 and Axiom 5 imply that WARP holds up to choice sets with a size of $k_c + 1$, which violates the definition of 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 $z \in S$ ($z \neq y$) such that for any $T \in \mathcal{O}$ such that $\{y,z\} \subset T$, $y \neq c(T)$. Suppose by contradiction that the implication is not true. Then for any $z \in S$ ($z \neq y$) there exists $T_z \in \mathcal{O}$ such that $\{y,z\} \subset T_z$, such that $y = c(T_z)$. Axiom 3 then implies that $y = c(\cup_{z \in S} T_z)$. Since $\{x,y\} \subset S \subseteq \cup_{z \in S} T_z$, Axiom 1 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 $S' \in \mathcal{O}$ such that $\{x, y\} \subset S'$ and $y = c(S')$. Both implications established above suggest that for any $S' \in \mathcal{O}$ such that $y = c(S')$, $y \neq c(S' \cup \{z\})$. 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 1-5.

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

We show that Axiom 1 holds. 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 $y \neq c(T)$. Axiom 2 is satisfied because of the transitivity of $\succ$.

We now show that Axiom 3 holds. Assume that $x = c(S) = c(T)$ and $S, T \in \mathcal{O}$. Then $|S|, |T| > k_c \geq k$, and $c(S) = max(\Gamma^P_k(S), \succ)$ and $c(T) = max(\Gamma^P_k(T), \succ)$. So $x \in \Gamma^P_k(S) \cap \Gamma^P_k(T)$, i.e., there is no $z \in S \cup T$, such that zPx. Then $x \in \Gamma^P_k(S \cup T)$. If there is another option $y \in \Gamma^P_k(S \cup T)$, $y \in \Gamma^P_k(S)$ when $y \in S$ and $y \in \Gamma^P_k(T)$ when $y \in T$. In both cases it must be that $x \succ y$ because $x = c(S) = c(T)$. This establishes that $x = c(S \cup T)$.

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

We finally show that Axiom 5 holds. Let M be a set of mutually undominated alternatives. Consider any two options $x, y \in M$. Then $x = c(S)$ and $y = c(S')$ for some $S \in \mathcal{O}$ and $S' \in \mathcal{O}$ such that $\{x, y\} \subset S \cap S'$. Since $|S|, |S'| > k_c \geq k$, $x = c(S)$ and $y = c(S')$ suggests that neither xPy nor yPx occurs. Then $|M| \leq k$ because otherwise $|max(M, P)| = |M| > k$, which violates the condition that $|max(S, P)| \leq k$ for any $|S| > k$. Assume that $x = c(S) = c(T)$ and $|S \cup T| \leq |M|$. Then $|S \cup T| \leq k$ and in turn $x = c(S_1 \cup S_2)$. This establishes the necessity of Axioms 1-5. $\square$

Proof of Proposition 8

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 B.1 illustrates that $k = 2$ is possible when $N = 4$. □

Appendix B. Characterization of Choice with Top-k Shortlisting

Proposition B.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 WARP-violating choice function 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 Proposition 8 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. Proposition 8 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 $N = 4$, we list in the following example, all the choice patterns that are rationalizable by a top-k shortlisting for some number k.

Example B.1 (Choice with top-k shortlisting when $N = 4$). *Consider $X = \{x_1, x_2, x_3, x_4\}$. Assume without loss of generality that $x_1 = c(\{x_1, x_2\})$, $x_2 = c(\{x_2, x_3\})$, and $x_3 = c(\{x_3, x_4\})$. Then a choice function defined on X is rationalizable by a top-k shortlisting for some k if and only if it satisfies $x_1 = c(\{x_1, x_3\}) = c(\{x_1, x_4\})$, $x_2 = c(\{x_2, x_4\})$, and the choices in the remaining choice sets fit one of the following eight patterns. One may define the preference relation as*

TABLE B.1: Choice with top-k shortlisting when $N = 4$

	$\{x_1, x_2, x_3\}$	$\{x_1, x_2, x_4\}$	$\{x_1, x_3, x_4\}$	$\{x_2, x_3, x_4\}$	X	capacity
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	$k = 2$

$x_1 \succ x_2 \succ x_3 \succ x_4$. 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.