Revealed Attention*

Yusufcan Masatlioglu Daisuke Nakajima Erkut Y. Ozbay

September 2009

Abstract

One of the underlying assumptions of classical choice theory is that the presented choice set is the same as what the agent indeed considers while deciding. However, marketing and psychology literatures provide well-established evidence that consumers do not consider all brands in a given market before making a purchase decision. Instead, they pay attention to only a subset of it and then undertake a more detailed analysis of the reduced sets of alternatives. Building on this idea, we provide a choice theoretical foundation for maximizing a single preference relation under limited attention. In our model, the DM picks her most preferred item from the alternatives she pays attention, not from the entire feasible set. Most importantly, we illustrate how one can deduce both her preference and the alternatives to which she pays attention and inattention from the observed behavior.

JEL Classification: D11, D81.

Keywords: Revealed Preferences, Awareness, Attention, Consideration Set

1 Introduction

Revealed preference is one of the most influential ideas in economics, which has been applied in a number of areas of economics such as consumer theory.[1] According to

*We would like to thank Tilman Börgers, Emel Filiz-Ozbay, Stephan Lauermann, Paola Manzini, Marco Mariotti and Neslihan Uler for their helpful comments.

[1]Varian (2006) is a nice survey of revealed preference analysis.

the standard revealed preference theory, x is revealed to be preferred to y if and only if x is chosen when y is also available (Gul and Pesendorfer (2008)) and any choice reversals, observed both empirically and experimentally, have been attributed as irrational since they cannot be expressed as a maximization of a preference.

The revealed preference argument relies on an implicit assumption of the choice theory that a decision maker (DM) considers all the presented alternatives. Although DM prefers x to y, she may still choose y when x is present simply because she does not realize that x is also available (Hausman (2008)). In such a situation, although the decision maker behaves seemingly irrational, the DM may be rational within her bounded understanding of what is available.[2] For example, while using a search engine, a DM may only pay attention to the alternatives appearing in the first page of the results since it takes too much time to consider all the search results. If her most preferred item, x, is not in the first page, then she picks the best alternative only within those in the first page, y. However, she will choose x when some of the items become unavailable and x moves to the first page.

In this paper, we provide a choice theoretical foundation for maximizing a single preference relation under limited attention. In our model, the DM picks her most preferred item from the alternatives she pays attention to, not from the entire feasible set.[3] Most importantly, we illustrate how one can deduce both her preference and the alternatives to which she pays attention and she does not.

The notion of limited attention has been extensively studied in marketing and finance literatures. The marketing literature calls the set of alternatives that a DM pays attention in her choice process as the *consideration set* (Wright and Barbour (1977)). It has been demonstrated that individuals are faced with an immense amount of product information and/or a relatively large number of alternative. Since most individuals have neither the cognitive abilities nor resources required to effectively process all available information relevant to a particular choice occasion, they do not

[2]As argued in Aumann (2005), this behavior is still considered rational since she is picking the best alternative under her limited information of what is available (Simon (1957)).

[3]Of course, if the DM pays attention to all the alternatives among which she is going to choose, then she acts rational in the sense of the classical choice theory.

pay attention to all the available alternatives. For example, a shopper in a typical supermarket needs to select from 285 varieties of cookies, 85 flavors and brands of juices, 230 different soups and 275 varieties of cereal (Schwartz (2005)). Abundance of choice also creates difficulty in financial decisions such as how to invest 401(k) plans among hundreds of funds or deciding on health insurance provider, premium, deductible and coverage amounts. Due to cognitive limitations, the DM cannot pay attention to all the available alternatives. As Simon (1957) pointed out being able to consider all possible alternatives is as hard as comparing them for decision makers. Therefore, a DM with a limited cognitive capacity, such as unawareness as demonstrated in Goeree (2007)[4], considers only a small fraction of the objects present in the associated market (Stigler (1961); Pessemier (1978); Chiang et al. (1998)).[5] Lavidge and Steiner (1961) is presented awareness of an item as a necessary condition to be in the consideration set. Furthermore, an individual may not consider all items even when she is aware of them so her consideration set can be even smaller than her awareness set. This is empirically supported by Jarvis and Wilcox (1977), which shows that while the average number of known products may vary a lot for different class of products, the average size of the consideration set is three to eight products.

A sizable amount of research has been conducted on the formation of the consideration sets in behavioral research in marketing (Hauser and Wernerfelt (1990); Roberts and Lattin (1991)). The common property is that when an alternative that is not paid attention becomes unavailable, the consideration set remains the same. This basic property of the attention is also documented in the psychology literature (Broadbent (1958)). We interpret this property as the minimal condition and call the consideration sets satisfying this property as *attention filter*. As Broadbent (1958) points out that this filter prevents the information-processing system from becoming overloaded because we have only a limited capacity to process information.

[4]How unawareness alters the behavior of the DM has been studied in various contexts such as game theory (Heifetz et al. (2007); Ozbay (2008)), and contract theory (Filiz-Ozbay (2008)).

[5]Also in financial economics, it is shown that investors reach a decision within their limited attention (Huberman and Regev (2001)). Similar examples can be found in job search (Richards et al. (1975)), university choice (Dawes and Brown (2005); Laroche et al. (1984); Rosen et al. (1998)), and airport choice (Basar and Bhat (2004)).

While this property is trivially satisfied in the classical choice theory, it is normatively appealing especially when (i) a DM pays attention to all of items she is aware of, and (ii) she is unaware that she is unaware of other items. If she is not only unaware of a particular alternative but also she is not at all aware that she overlooks some items. Then, even when this alternative is removed, such a change will not be recognized by her, so her consideration set must stay the same.

In addition, the property is also satisfied when the formation is based on some decision heuristics, such as paying attention only to N-most advertised or the products appearing in the first page of search results. Interestingly, the property is also appealing in the other extreme case where the DM is perfectly rational in the sense that she chooses the set of alternatives which are worthwhile to consider by calculating the expected values including the cost of search and investigation. Assume she chooses a particular set, A, as her consideration set and x is some alternative outside of A. Then, for any alternative, the marginal contribution to A is not affected when x becomes unavailable, so is the consideration set.

Under this structure, one can elicit the DM's preference whenever a choice reversal is observed.[6] To see this, assume that she chooses x but removing y changes the choice. This can happen only when her attention span has changed. This is impossible if she had not paid attention to y, hence y must have been considered (*Revealed Attention*). Given the fact that x is chosen while y draws her attention, we conclude that she prefers x over y (*Revealed Preference*). In sum, whenever her choices changes as a consequence of removing an alternative, the initially chosen alternative is preferred to the removed one.

Choice with attention filter accommodates several frequently observed behaviors that cannot be captured by the standard choice theory. Attraction Effect, Cyclical Choice, Choosing Pairwisely Unchosen are some of the examples in that regard.[7] We will discuss them in turn.

[6]Without any structure on the formation of the consideration sets, of course any choice behavior can be rationalized by any preference (Hausman (2008)).

[7]These anomalies are also studied in Cherepanov et al. (2008).

Attraction Effect

The attraction effect refers to a phenomenon where adding an inferior alternative to the choice set affects the choice.[8] Lehmann and Pan (1994) experimentally show that introducing new products causes the attraction effect particularly by affecting the composition of consideration sets. Our model provides an explanation for the attraction effect in line with their findings.

For illustration purposes, suppose that there are two options x and y where x is better than y. Now imagine that a third option z that is clearly dominated by y but not by x becomes available. The presence of an inferior element, z, attracts the DM to pay attention to y and prevents the DM from paying attention to x. Thus, she chooses y from $\{x, y, z\}$, which will generate a typical attraction effect choice pattern with three alternatives:

$$c(xyz) = y, \quad c(xy) = x, \quad c(yz) = y, \quad c(xz) = x.$$

Furthermore, the attraction effect can be generalized to accommodate more than one inferior but attracting alternative:[9]

$$c(xyzt) = x, \quad c(xyz) = y, \quad c(xy) = x.$$

Here, as in the previous example, the presence of z attracts the DM to pay attention to y and prevents the DM to pay attention to x. However, adding t to the alternative set attracts the attention of the DM to x again. Therefore, when $\{x, y, z, t\}$ is presented, the DM pays attention to all the alternatives and chooses x as she was choosing x from x, y.[10]

[8] This phenomenon is well-documented and robust in behavioral research in marketing (Huber et al. (1982); Huber and Puto (1983); Tversky and Simonson (1993); Lehmann and Pan (1994)), including in choice among monetary gambles (Wedell (1991)), political candidates (Pan and Pitts (1995)), and job candidates (Highhouse (1996)), environmental issues (Bateman et al. (2005)), medical decision making (Schwartz and Chapman (1999)). Advertising inferior alternatives is commonly used as a marketing strategy to invoke the attraction effect on the customers.

[9] While Ok et al. (2008) propose a reference-dependent choice model that accounts for the attraction effect, de Clippel and Eliaz (2009) provide a characterization for both the attraction and compromise effects which are derived from an intra-personal bargaining problem among different selves of an individual.

[10] The generalization of the attraction effect lies outside the rationalization model (Cherepanov et al. (2008)) and shortlisting (Manzini and Mariotti (2007)) since it does not satisfy the axiom called Weak WARP. We will discuss these models in detail later.

Cyclical Choice

May (1954) provides the first experiment where cyclical choice patterns are observed. Choice cycles have been demonstrated in many different choice environments (e.g. Tversky (1969); Loomes et al. (1991); Manzini and Mariotti (2009b); Mandler et al. (2008)). In a cyclical choice pattern:

$$c(xyz) = y, \quad c(xy) = x, \quad c(yz) = y, \quad c(xz) = z.$$

In our framework, this choice pattern reveals that that x is the best and z is the worst alternative and the DM does not pay attention to x whenever z is present Therefore, the best alternative x id not chosen from $\{x, z\}$, so causes the cyclical behavior.

Choosing Pairwisely Unchosen[11]

In this choice pattern, the DM chooses an alternative that is never chosen from pairwise comparisons:

$$c(xyz) = z, \quad c(xy) = x, \quad c(yz) = y, \quad c(xz) = x.$$

In our framework, this happens when z is the best alternative but the DM realizes it only when both x and y are present (while x are y are always considered).

In addition to providing a characterization of choice with limited attention, we study further structures on the attention filter. In many real-world markets, every product is competing with each other for the space in the consideration set of DM, who has cognitive limitations. In these situations, if an alternative attracts attention when there are many others, then it must be considered when some of them become unavailable. If a product is able to attract attention in a crowded supermarket shelf, the same product will be noticed when there are fewer alternatives (Lleras et al. (2009)). We call this class of attention filters as *strong attention filter*. We then provide a characterization and investigate the question of revealed preferences for this special class.

Finally, we study two independent special cases of our model. The first model assumes that the decision maker pays attention to both alternatives in any binary budget set. In the second one, the attention filter is generated by a transitive order,

[11] Cherepanov et al. (2008) call this choice pattern as the "difficult choice".

which might conflict with the preferences. This order is used, independently of the budget set, to remove dominated alternatives within the available ones.

Related Literature

There are several papers closely related to ours. Manzini and Mariotti (2007) proposes a model in which the same fixed set of rationales is intentionally applied to remove inferior alternatives in a sequential manner.[12] Cherepanov et al. (2008) offers a formal model of rationalization. In this model, a decision maker has well-defined preferences over alternatives. However, she cannot choose an alternative unless it can be rationalized. That is, it must be optimal according to at least one of her rationales. Manzini and Mariotti (2008) also considers a model where a decision maker first eliminates alternatives belonging to some inferior category and picks the best one among remaining alternatives. Unlike the limited attention model, these models cannot accommodate the generalized attraction effect. Having said that, our model is not a generalization of these models since one can always find examples that violates our axiom nevertheless can be captured by these models. Another difference is that the DM in these models is implicity assumed to consider all available alternatives in the first stage and she *intentionally* eliminates several alternatives before she proceeds to the next stage.

Inability to consider all alternatives at the beginning of the decision process is also present in Masatlioglu and Nakajima (2007). Unlike our model, there the consideration sets evolve depending on an exogenously given starting point, and the choice is determined through an iterative search process. Eliaz and Spiegler (2007) analyzed a market where firms face consumers who make choices through consideration sets which change depending on firms' costly marketing devices. Eliaz et al. (2009) studies a very concrete and reasonable ways to construct a consideration set. Indeed, some of consideration sets we shall present as examples are within their models. Contrary

[12]Other papers studying sequential rationality are Apesteguia and Ballester (2008); Houy (2007); Houy and Tadenuma (2009).

to ours, in their paper, the decision maker's consideration set (called *finalists*) is observed and is directly investigated. On the other hand, in our model it is an object that must be inferred from her final choice.

2 Choice with Limited Consideration

Let X be an arbitrary non-empty finite set and $\mathcal{X}$ be the set of all non-empty subsets of X. A choice or plan assigns a unique chosen element to every non-empty feasible set. This choice can be represented by a choice function on $\mathcal{X}$, $c : \mathcal{X} \to X$, such that $c(S) \in S$ for every $S \in \mathcal{X}$. Let $\succ$ be a strict linear order on X. We denote the best element in S with respect to $\succ$ by $\max_{\succ} S$.

We propose a model to capture the idea of limited consideration: the DM pays attention to only a subset of all available alternatives and picks the best alternative among them:

Definition. *A choice function c is a* ***choice with limited consideration (CLC)*** *if there exists a strict linear order $\succ$ and a consideration set mapping Γ such that*

$$c(S) = \max_{\succ} \ \Gamma(S)$$

where $\emptyset \neq \Gamma(S) \subset S$ is the consideration set that consists of alternatives to which the DM considers under choice problem S.

Occasionally, we say that $(\Gamma, \succ)$ represents c. We also mention that $\succ$ represents c, which means that there exists some consideration set mapping Γ such that $(\Gamma, \succ)$ represents c.

2.1 Attention Filter

As we argue in the introduction, the formation of consideration sets and their properties are the main focus of the marketing literature. Let us consider some typical examples of the consideration sets formation:

Example 1. *(i) A consumer pays attention only to N-most advertised or N-safest cars in the market. In case of choosing a supplier for a particular product,* *Dulleck et al. (2008) showed that consumers select a shortlist of suppliers by using the price variable only (for example, three cheapest suppliers).*

(ii) An economics department pays attention only to the top job candidate in each field for hiring one assistant professor.

(iii) A web-search engine user pays attention only to the products appearing in the first page of search results and/or sponsored links (Hotchkiss et al. (2004)).

(iv) A consumer visits only the store which offers the most variety. Zyman (1999) provides a real-world evidence for such behavior where each store now corresponds to a category and consumers consider products within the most popular category. Sprite was a well-liked drink among target consumers but it suffered from being positioned in the lemon-lime category. After Sprite was repositioned simply as a soda (popular category) with a lemon-lime flavor, its sales increased dramatically.

A common property across all of these examples is that removing an alternative which does not attract attention does not change the set of alternatives that are paid attention. For example, in Example 1(ii), when a non-top candidate of a field withdraws from the job market, the department would still consider the same top candidates. Similarly in (iv), the store offering the most variety remain the same when the product she does not consider (i.e. not sold at the store she visits) becomes unavailable so she stays with the store and her attention span is not affected. We call this property of the consideration sets formation as *attention filter*.[13] Formally,

Definition. *A consideration set mapping* Γ *is an* ***attention filter*** *if for any* S, $\Gamma(S) = \Gamma(S \setminus x)$ *whenever* $x \notin \Gamma(S)$.[14]

Clearly, the DM who considers all of the feasible alternatives in any decision problem (as it is the case in the standard theory) can be viewed as an identity attention

[13]Without any restriction on the formation of the consideration set, Γ, any choice can be rationalized as a CLC by allowing Γ to include only the chosen alternative.

[14]Here, we abuse notation and write $S \setminus x$ instead of $S \setminus \{x\}$. Throughout the paper, we use this notation.

filter. As we discuss in the introduction, the DM's consideration set is an attention filter if inclusion of an alternative in a consideration set is based on a trade-off between costs and benefits. These include costs of information search and thinking about and evaluating alternatives and the evaluation of the benefits or value from including an alternative in the consideration set for a particular decision problem. Another example for attention filter is the one when the source of her limited attention is due to the complete unawareness. In addition to these, as we demonstrate, heuristics presented above also generates an attention filter.

First, let us illustrate how to infer (1) the DM's preference and (2) what she pays (and does not pay) attention from her observed choice that is a CLC with an attention filter. The standard theory concludes that x is preferred to y immediately when x is chosen while y is available. To justify such an inference, one must implicity assume that she has paid attention to y. Without this hidden assumption, we cannot make any inference because she may prefer y but overlook it. Therefore, eliciting the DM's preference is no longer trivial because her choice can be attributed to her preference or to her inattention.

This observation suggests that multiple pairs of a preference and an attention filter can generate the same choice behavior. To illustrate this, consider the choice function with three elements exhibiting a cycle:

$$c(xyz) = x, \quad c(xy) = x, \quad c(yz) = y, \quad c(xz) = z.$$

One possibility is that the DM's preference is $z \succ_1 x \succ_1 y$ and she overlooks z both at $\{x, y, z\}$ and $\{y, z\}$. Another possibility is that her preference is $x \succ_2 y \succ_2 z$ and she does not pay attention to x only at $\{x, z\}$ (see Table 1 for the corresponding attention filters).

Preference	Attention Filter				
		xyz	xy	yz	xz
$z \succ_1 x \succ_1 y$	Γ_1	xy	xy	y	xz
$x \succ_2 y \succ_2 z$	Γ_2	xyz	xy	yz	z

Table 1: Two possible representations for the cyclical choice

We cannot identify which of them is her true preference. Nevertheless, if only these two pairs represents c, we can unambiguously conclude that she prefers x to y because both of them rank x above y. For the same reason, we can infer that she pays attention to both x and y at $\{x, y, z\}$ (Table 1). This example makes it clear that we need to define the revealed preference when multiple representations are possible.

Definition 1. *Assume c is a choice by limited attention and there are k different pairs of preference and attention filter representing c, $(\Gamma_1, \succ_1), (\Gamma_2, \succ_2), \ldots, (\Gamma_k, \succ_k)$. We say*

- *x is revealed to be preferred to y if $x \succ_i y$ for all i,*
- *x is revealed to attract attention at S if $\Gamma_i(S)$ includes x for all i,*
- *x is revealed **not** to attract attention at S if $\Gamma_i(S)$ excludes x for all i.*

If one wants to know whether x is revealed to be preferred to y, it seems to be necessary to check for every $(\Gamma_i, \succ_i)$ whether it represents her choice or not, which is not practical especially when there are many alternatives. Instead we shall now provide a handy method to obtain the revealed preference, attention and inattention completely.

In the example above, when Γ is an attention filter, it is possible to determine the relative ranking between x and y. To see this, first, note that if the DM pays attention to x and z at both $\{x, z\}$ and $\{x, y, z\}$, then we should not observe choice reversal. If there is a choice reversal, then this means that her attention set changes when y is removed from $\{x, y, z\}$. This is possible only when she pays attention to y at $\{x, y, z\}$ (*revealed attention*). Given the fact that x is chosen from $\{x, y, z\}$ we conclude that the DM prefers x over y (*revealed preference*). This observation can be easily generalized. Whenever the choices change as a consequence of removing an alternative, the initially chosen alternative is preferred to the removed one. Formally, for any distinct x and y, define:

$$xPy \text{ if there exists } T \text{ such that } c(T) = x \neq c(T \setminus y). \tag{1}$$

By the argument analogous to the above, the revealed preference must include P. In addition, we also conclude that she prefers x to z if xPy and yPz for some y, even when xPz does not hold. Therefore, the transitive closure of P, denoted by P_R, must be also the part of her revealed preference. One may wonder whether some revealed preference is overlooked by P_R. The next proposition states that the answer is no: P_R is the revealed preference.

Proposition 1. *(Revealed Preference) Suppose c is a CLC with an attention filter. Then, x is revealed to be preferred to y if and only if xP_Ry.*

To understand the importance of this result, consider the following example. Assume that there are two alternatives x and y such that x is never chosen in the presence of y and y is sometimes chosen in the presence of x. Bernheim and Rangel (2009) proposes this notion as a strict welfare improvement in their model-free approach. Hence they regard y as being strictly better than x. Contrary to that, this information, in our model, alone does not provide us any revealed preference. More strikingly, x may be revealed to be preferred to y. Suppose her choice include:

- $c(S) = x$ and $c(S \setminus z) \neq x$ where $y \notin S$
- $c(T) = z$ and $c(T \setminus y) \neq z$

for some S and T, and an alternative $z \in S \cap T$. These observations reveal that x is preferred to z and z is preferred to y. Therefore, we conclude that x is preferred to y even though x is never chosen when y is present. This observation highlights the importance of knowledge about the underlying model when we do welfare analysis.[15]

Next, we investigate when we can unambiguously conclude that the DM pays (or does not pay) attention to an alternative. Consider the choice reversal above, from which we have concluded that she prefers x to y. Therefore, whenever y is chosen, she must not have paid attention to x (revealed inattention).

[15] For detail discussion on this subject, see Manzini and Mariotti (2009a).

As we illustrate, we infer that x is revealed to attract attention at S whenever x is chosen from S or removing x from S causes a choice reversal. Furthermore, it is possible to reach the same conclusion even when removing x from S does not cause a choice reversal. Imagine that the DM chooses the same item, say $\alpha \neq x$, from S and T and removing x from T causes a choice reversal, so we know $x \in \Gamma(T)$ for sure. Now collect all items that belong to either S or T but not to both. Suppose all of those items are revealed to be preferred to α. Then, those items cannot be in $\Gamma(S)$ or $\Gamma(T)$. Therefore, removing those items from S or T cannot change her consideration set. Hence, we have

$$\Gamma(S) = \Gamma(S \cap T) = \Gamma(T)$$

and can conclude that x is considered at S. The following proposition summarizes this observation and also provides the full characterization of the revealed attention and inattention.

Proposition 2. *(Revealed (In)Attention) Suppose c is a CLC with an attention filter. Then,*

(1) x is revealed not to attract attention at S if and only if $xP_R c(S)$,

(2) x is revealed to attract attention at S if and only if there exists T (possibly equal to S) such that:

(i) $c(T) \neq c(T \setminus x)$,

(ii) $yP_R c(S)$ for all $y \in S \setminus T$,
$zP_R c(T)$ for all $z \in T \setminus S$.

So far, we provide two propositions to characterize revealed preference and revealed (in)attention. These two propositions suppose that observed choice behavior is a CLC with an attention filter. Without this assumption, propositions are not applicable. Therefore, the question is: how can we test whether choice data is consistent with CLC with an attention filter? Surprisingly, it turns out that CLC with an attention filter can be simply characterized by only one observable property of choice.

Before we state our property, let's recall the sufficient and necessary condition for observed behavior to be consistent with the preference maximization under the full attention assumption: the Weak Axiom of Revealed Preference (WARP). WARP is equivalent to that every set S has the "best" alternative x^* in the sense that it must be chosen from any set T whenever x^* is available and the choice from T lies in S. Formally,

> **WARP:** For any nonempty S, there exists $x^* \in S$ such that for any T including x^*,
> $$c(T) = x^* \text{ whenever } c(T) \in S.$$

Because of the full attention assumption, being feasible is equal to attracting attention. However, this is no longer true when we allow the possibility of limited attention. To conclude that x^* is chosen we need to make sure not only that the chosen element is from S and x^* is available but also that x^* attracts attention. As we have discussed, we can infer it when removing x^* from T changes her choice, which is the additional requirement for x^* to be chosen from T. This discussion suggests the following axiom, which is a weakening of WARP:

> **(A1) WARP with Limited Attention (WARP(LA))** For any nonempty S, there exists $x^* \in S$ such that, for any T including x^*,
> $$c(T) = x^* \text{ whenever } \quad \begin{array}{l} (i)\ c(T) \in S,\ \textit{and} \\ (ii)\ c(T) \neq c(T \setminus x^*) \end{array}$$

WARP with Limited Attention indeed guarantees that the binary relation P defined in (1) is acyclic and it fully characterizes the class of choice functions generated by an attention filter. The next lemma makes it clear that WARP(LA) is equivalent to the fact that P is has no cycle.

Lemma 1. *P is acyclic if and only if c satisfies WARP with Limited Attention.*

Proof. (The if-part) Suppose P has a cycle: $x_1 P x_2 P \cdots P x_k P x_1$. Then for each $i = 1, \ldots, k-1$ there exists T_i such that $x_i = c(T_i) \neq c(T_i \setminus x_{i+1})$ and $x_k = c(T_k) \neq$

$c(T_k \setminus x_1)$. Consider the set $\{x_1, \ldots, x_k\} \equiv S$. Then, for every $x \in S$, there exists T such that $c(T) \in S$ and $c(T \setminus x) \neq c(T)$ but $x \neq c(T)$, so WARP with Limited Attention is violated.

(The only-if part) Suppose P is acyclic. Then every S has at least one element x such that there is no $y \in S$ with yPx, which means that there is no $y \in S$ with $y = c(T) \neq c(T \setminus x)$. Equivalently, whenever $c(T) \in S$ and $c(T) \neq c(T \setminus x)$, it must be $x = c(T)$, which is WARP with Limited Attention. □

Theorem 1. *(Characterization) c satisfies WARP with Limited Attention if and only if c is a CLC with an attention filter.*

Proof. The proof is given in the appendix. □

Theorem 1 shows that a CLC with an attention filter is captured by a single behavioral postulate. This makes it possible to test our model non-parametrically by using the standard revealed-preference technique a la Samuelson and to derive decision maker's preferences and attention filter based on Proposition 1 and 2 from observed choice data. Before we move to the next section, we revisit Choosing Pairwisely Unchosen to illustrate how one can use the results in this section.

$$c(xyz) = z, \quad c(xy) = x, \quad c(yz) = y, \quad c(xz) = x.$$

First, let us illustrate that this choice behavior satisfies WARP(LA). Take $S = \{x, y, z\}$. Since removing x or y from S changes the choice and $c(S) = z$, neither of them can serve as x^*. So the only candidate is z, hence we need to check that decision problems where removing z changes the choice. This happens only when $T = S$, so we must have $c(T) = z$, which is the case. Hence our axiom is satisfied for $S = \{x, y, z\}$. Notice that only $c(S)$ can serve the role of x^* for $S = \{x, y, z\}$. However, this is not always the case.[16] Let S be $\{x, z\}$. $c(S)(= x)$ cannot be x^* for this set. Take $T = \{x, y, z\}$. While $c(T)$ lies in S and removing x from T changes the choice, $c(T)$ is not equal to x. On the other hand, z does the job for $\{x, z\}$. It is routine to check the rest.

[16]Note that x^* in WARP is always equal to $c(S)$.

Since the choice behavior satisfies WARP(LA), it is a CLC with an attention filter by Theorem 1, so Proposition 1 and 2 are applicable: removing x or y from $\{x, y, z\}$ changes the choice, they must have drawn the attention (revealed attention). Combining this with the fact that $c(xyz) = z$ reveals z is the best alternative (revealed preference). Given the revealed preference and choices from binary sets, the DM must not not pay attention to z in any binary comparisons (revealed inattention). Hence the consideration set mapping is *uniquely* pinned down except $\Gamma(xy)$[17]:

$$\Gamma(xyz) = xyz, \quad \Gamma(yz) = y, \text{ and } \Gamma(xz) = x.$$

2.2 Strong Attention Filter

In the last section, we study the class of CLC with an attention filter where removal of an overlooked alternative does not change the consideration set. Since there is no requirement on consideration sets when alternative that are paid attention becomes unavailable, this class is rather general. However, this was intentional so that we can provide revealed preference and revealed (in)attention results even without imposing too much structure.

Nevertheless, we argue that there is another plausible requirement one might like to impose on people's attention filters. In many real-world markets, every product is competing with each other for the space in the consideration set of the DM, who has cognitive limitations. In these situations, if an alternative attracts attention when there are many others, then it is easier to be considered when some of them become unavailable. If a product is able to attract attention in a crowded supermarket shelf, the same product will be noticed when there are fewer alternatives.

This suggests that we should study the case where an attention filter satisfies the following extra property:[18]

Definition. *An attention filter* Γ *is called* ***strong attention filter*** *if for any* S *and* T, $x \in \Gamma(T)$ *implies* $x \in \Gamma(S)$ *whenever* $x \in S \subset T$.

[17]While x belongs to $\Gamma(xy)$, we cannot determine whether y is in $\Gamma(xy)$.

[18]The companion paper, Lleras et al. (2009), extensively investigates consideration sets that satisfies only this property.

Indeed, all of heuristic ways of generating a consideration set discussed in the previous subsection are strong attention filters except the last one.

First let us characterize the reveled preference when Γ is known to be a strong attention filter. To do this, we revisit the cyclical choice behavior in the perivous subsection where we know that while x is revealed to preferred to y, there is no other revealed preference (see Proposition 1). Interestingly, we can uniquely pin down the preference for this example when Γ is a strong attention filter.

To see this, first note that $c(xyz) = x$ implies that the DM pays attention to x at $\{x, y, z\}$ so does she at $\{x, z\}$ (revealed attention). Since she picks z from $\{x, z\}$, we can conclude that she prefers z over x (revealed preference). Since any strong attention filter is an attention filter, we must have x is revealed to preferred to y. Therefore, her preference is uniquely pinned down: $z \succ x \succ y$.

Now we generalize this observation. Suppose $c(T) \neq c(T \setminus y)$. Then we conclude that y must be paid attention to at T. Since Γ is a *strong* attention filter, y must attract attention not only at T but also at any decision problem S smaller than T including y. Therefore, if $c(S) \neq y$, $c(S)$ is revealed to be preferred to y. Formally, for any distinct pair of x and y define:

$$xP'y \text{ if there exist } S \text{ and } T \text{ such that} \quad \begin{array}{l} (i) \;\; \{x, y\} \subset S \subset T \text{ and } x = c(S) \\ (ii) \; c(T) \neq c(T \setminus y) \end{array}$$

Notice that $c(T) \neq c(T \setminus c(T))$. This implies that $c(T)$ must have been considered not only at T but also at any decision problem S smaller than T including $c(T)$ since Γ is a *strong* attention filter. Therefore, whenever $\{c(T)\} \subset S \subset T$ and $c(T) \neq c(S)$, we have $c(S) \succ c(T)$. Indeed this is the way we infer z is better than x in the previous paragraph.

As before, if $xP'y$ and $yP'z$ for some y, we also conclude that she prefers x to z even when $xP'z$ does not hold. The following proposition states that the transitive closure of P', denoted by P'_R is the revealed preference.

Proposition 3. *Suppose c is a CLC with a strong attention filter. Then, x is revealed to be preferred to y if and only if xP'_Ry.*

Proof. The if-part has been already demonstrated. The only-if part can be shown paralleled with Theorem 2, where we shall show that any $\succ$ including P'_R represents c by choosing Γ properly. $\square$

Strong attention filter captures the idea that an alternative that is not paid attention in a smaller set cannot attract attention when there are more alternatives. Hence, situations where presence of some alternatives reminds the DM the existence of some other alternatives are compatible with the attention filter but not with the strong attention filter.

The example of "choosing pairwisely unchosen", which was studied earlier, perfectly highlights this distinction between attention filter and strong attention filter structures. Recall that we uniquely identify the attention filter for $\{x, y, z\}$, $\{x, z\}$, and $\{y, z\}$;

$$\Gamma(xyz) = xyz, \quad \Gamma(yz) = y, \text{ and } \Gamma(xz) = x.$$

Here, z attracts attention only when both x and y are present. In other words, while z draws the attention from a big selection, it is not considered from a restrictive selection. Hence this is not a strong attention filter. Since Proposition 2 requires that $z \in \Gamma(xyz)$ but $z \notin \Gamma(yz)$, we can immediately conclude that this choice behavior cannot be explained by a strong attention filter. We can also reach the same conclusion by using revealed preference: the DM's choice exhibits two choice reversals: (1) between $\{x, y, z\}$ and $\{x, z\}$ and (2) between $\{x, y, z\}$ and $\{y, z\}$. Based on Proposition 3, the first one implies that her preference must be $x \succ z \succ y$ and the second reveals $y \succ z \succ x$, which are contradicting.

Next, we provide a characterization for CLC with a strong attention filter. The axiom we propose is a stronger version of WARP(LA). Remember that WARP(LA) requires that every set S has the "best" alternative x^* and it must be chosen from any other decision problem T as long as it attracts an attention. Remember that, with an attention filter, an alternative, say x^*, attracts attention at a choice set, T, when if removing it changes the choice, i.e., $c(T) \neq c(T \setminus x^*)$. Now that we assume the attention filter is strong, we can also conclude it when we know x^* is paid attention to at some bigger decision problem $T' \supset T$ by observing $c(T') \neq c(T' \setminus x^*)$. Therefore, we need to modify the requirement in WARP(LA): if the removal of x^* changes the choice in some super set of T, then it attracts attention at T.

(A2) For any nonempty S, there exists $x^* \in S$ such that for any $T \ni x^*$,

$$c(T) = x^* \text{ whenever } \quad (i)\ c(T) \in S, \textit{ and}$$
$$(ii)\ c(T') \neq c(T' \setminus x^*) \text{ for some } T' \supset T$$

It turns out that A2 is the necessary and sufficient condition for CLC with a strong attention filter. Indeed, it is equivalent to the acyclicity of the revealed preference, P'_R.

Theorem 2. *(Characterization) A choice function satisfies A2 if and only if it is a CLC with a strong attention filter.*

Theorem 2 characterizes a special of class of choice behaviour studied we studied earlier. Similar to Theorem 1, the characterization involves a single behavioral postulate which is stronger that WARP with Limited Attention. We show that while this model has higher predictive power, which comes with diminishing explanatory power: "choosing pairwisely unchosen" is no longer within the model.

We finalize this section by revisiting the Attraction Effect:

$$c(xyz) = y, \quad c(xy) = x, \quad c(yz) = y, \quad c(xz) = x.$$

It is routine to verify that this choice behavior satisfies A2.[19] Hence Theorem 2 implies that it is consistent with a CLC with a strong attention filter. The choice reversal between $\{x, y, z\}$ and $\{x, y\}$ yields that her preference must be $x \succ y \succ z$.

In addition to that one can derive the unique consideration set mapping. To see this, consider the set $\{x, y, z\}$. First of all, the choice, which is y, must be in the consideration set. Since removing z changes the choice, therefore z is also in it (attention filter). Finaly, we know x is better than the choice from above discussion, x does not belong the consideration set of $\{x, y, z\}$. Hence $\Gamma(xyz) = yz$. In addition, the strong attention filter assumption requires that y and z attract attention whenever

[19] One can show that x serves the role of x^* for $\{x, y, z\}$. For the rest, $c(S)$ does the job.

they are available, which pins down the consideration set mapping uniquely for this example.

$$\Gamma(xyz) = yz, \quad \Gamma(xy) = xy, \quad \Gamma(yz) = yz, \text{ and } \Gamma(xz) = xz.$$

3 Additional Discussion

In this section we provide two independent special cases of our model. The first one assumes that the decision maker has no limited attention problem in a binary set, hence she pays attention to both alternatives. In the second one, the attention filter is generated by a transitive order, which might conflict with the preferences.

3.1 Perfect Sight in Binary Comparisons

The limited attention framework allows even for an extreme form of inattention when the DM does not pay attention to one of the alternatives in binary choice problems. Therefore, abundance of options may not be only source of inattention. This kind of inattention makes sense especially when the DM is unaware of some alternatives. However, if the source of her inattention is abundance of alternatives, she will consider more likely all the alternatives in small choice problems. As a benchmark case, let us consider a decision maker who has an attention filter but pays attention to both alternatives in every binary decision problem. That is, $\Gamma(S) = S$ whenever $|S| = 2$. We first provide a characterization for this class of choice function and then show that the revealed attention and inattention characterized in Section 2.2 are naturally extendable to this case.

In the classical choice theory, where the DM is assumed to pay attention to all the alternatives, under WARP it is possible to infer her preference by asking the choice from the sets of two alternatives:

$$xP^*y \text{ if } c(xy) = x$$

Now, the question is which assumptions are needed in order to have P^* is the revealed preference in limited attention model with perfect sight in binary comparisions.

First, it is needed to assume that

(A3) Pairwise Consistency: P^* is a strict linear order.

As we discussed earlier, an alternative, x, is revealed to attract attention at a set S whenever removing x from S causes a choice reversal. If x is not the chosen one, then it is strictly worse than $c(S)$. This information should not conflict with binary data, that is, $c(S)$ must be the choice from $\{x, c(S)\}$.

(A4) Weak Contraction: If $c(S) \neq c(S \setminus x)$ then $c(S) = c(\{x, c(S)\})$.

This axiom is trivially satisfied in the standard theory, where $c(S) = c(\{x, c(S)\})$ for every alternative x in S. Here, we require a weaker version of it, particularly we need to know whether the alternative attracts attention to reach the same conclusion.

Theorem 3. *A choice function satisfies A3 and A4 if and only if it is a CLC with an attention filter where there is perfect sight in binary sets.*

Our Proposition 2 can be easily extended to this model by replacing the original revealed preference P_R by P^*. Since we have the full revelation in preferences, the following result is stronger.

Proposition 4. *Suppose c is a CLC with an attention filter where there is perfect sight in binary sets. Then,*

(1) x is revealed not to attract attention at S if and only if $xP^*c(S)$,

(2) x is revealed to attract attention at S if and only if there exists T (possibly equal to S) such that:

(i) $c(T) \neq c(T \setminus x)$,

(ii) $yP^*c(S)$ for all $y \in S \setminus T$,
$zP^*c(T)$ for all $z \in T \setminus S$.

3.2 Attention Filters Generated by a Transitive Order

Here, we also consider a natural special case whereby the decision maker overlooks or disregards an alternative because it is dominated by another item in some aspect. Imagine Maryland's economics department is hiring one tenure-track theorist. Since there are too many candidates in the market, the department asks other departments to recommend their best theory student. Therefore, a candidate from Michigan is ignored if and only if there is another Michigan candidate who is rated better by Michigan. In this case, Maryland's filter is represented by a irreflexive and transitive order as long as each department's ranking over its students is rational. However, the order does not compare any two candidates from different schools so it is not complete.[20] Notice that this order may not be consistent with the preference of Maryland. It is possible that Michigan evaluates its job candidates differently than Maryland, in which case Maryland may eliminate its preferred candidate. Therefore, the order and the preference may be inconsistent.

Formally, let $\rhd$ be an irreflexive and transitive order over X and $\Gamma_\rhd$ be an attention filter generated by $\rhd$, that is:

$$\Gamma_\rhd(S) = \{x \in S |\ \nexists y \in S \ \text{ s.t. } \ y \rhd x\},$$

for all $S \in \mathcal{X}$. Here, the decision maker does not pay attention to x at decision problem S if and only if there is another alternative $y \in S$ that dominates x according to the transitive order. It is easy to see that $\Gamma_\rhd$ is indeed a special class of attention filters.

Here we illustrate that $\Gamma_\rhd$ is a strong attention filter. First note that $\Gamma_\rhd(T) \subset \Gamma_\rhd(S)$ for all $S \subset T$. To see this, assume $z \in \Gamma_\rhd(T)$. Then there exists no alternative in T $\rhd$-dominates z, which implies that z is $\rhd$-undominated in any subset of T, so $z \in \Gamma_\rhd(S)$ for all $S \subset T$. Particularly, we have $\Gamma_\rhd(S) \subset \Gamma_\rhd(S \setminus x)$ for any $x \in S$. Now we need to show that $\Gamma_\rhd(S \setminus x) \subset \Gamma_\rhd(S)$ when $x \notin \Gamma_\rhd(S)$. Suppose $x, y \in S \setminus \Gamma_\rhd(S)$. Then, there must exist $z \in S \setminus x$ such that $z \rhd x$. If $x \rhd y$, then by the transitivity,

[20] The special case in which the rationale always yields a unique maximal element corresponds to the standard model of rationality.

$z \rhd y$ as well so $y \notin \Gamma_\rhd(S \setminus x)$. If it is not $x \rhd y$, then what eliminates y at S is also included in $S \setminus x$ so $y \notin \Gamma_\rhd(S \setminus x)$. Therefore, $\Gamma_\rhd$ is a strong attention filter.

Since $\Gamma_\rhd$ is a strong attention filter, A2 is a necessary condition for CLC with an attention filter generated by a transitive order. In addition to that there is another necessary condition:

(A5) Expansion: If $x = c(S) = c(T)$, then $x = c(S \cup T)$.

Manzini and Mariotti (2007) dub this property Expansion, and it directly rules out Attraction Effect type of anomalies. It says that an alternative chosen from each of two sets is also chosen from their union. To see that it is necessary, assume $(\Gamma_\rhd, \succ)$ represents c and $x = c(S) = c(T)$. The latter implies that x is the $\succ$-best element in both $\Gamma_\rhd(S)$ and $\Gamma_\rhd(T)$. Hence x is $\rhd$-undominated in both S and T, so x is in $\Gamma_\rhd(S \cup T)$. Since $\Gamma_\rhd(S \cup T) \subset \Gamma_\rhd(S) \cup \Gamma_\rhd(T)$, x is also the $\succ$-best $\Gamma_\rhd(S \cup T)$. Hence $x = c(S \cup T)$.

Therefore, if an attention filter of our decision maker is generated by a transitive order, then her choice must satisfy Expansion, as well as A2. Indeed its converse is true so these two axioms characterize such choice functions.

Theorem 4. *A choice function satisfies A2 and A5 if and only if it is a CLC with attention filter which is generated by a transitive order.*

The proof of Theorem 4 is given in the Appendix.

Now, we discuss the revealed preference and the revealed order. Notice that this is a special case of CLC with a strong attention filter, P'_R, which is the revealed preference for the strong attention filter must be a part of the revealed preference for this model and it turns out that there is no extra inference of DM's preference.

However, we can now obtain the revealed order: if all of attention filters generated by some transitive order that can represent the choice agree on $x \rhd y$, we call it the revealed order. The revealed order can be obtained in a simple way. If she picks x from $\{x, y\}$ but reveals that she prefers y over x, it must be the case that y is disregarded at $\{x, y\}$. Therefore, we can conclude $x \rhd y$.

Proposition 5. *Suppose c is a CLC with an attention filter that is generated by a transitive order.*

- *x is revealed to be preferred to y if and only if $xP'_R y$.*
- *$x \rhd y$ is revealed if and only if yP'_R but $x = c(xy)$.*

Finally, we argue that the class of choice behavior characterized in Theorem 4 is a specific subclass of Manzini and Mariotti (2007)'s rational shortlist method. Similar to our model, the shortlist method operates through two binary relations P_1 (acyclic) and P_2 (asymmetric): the decision maker filters out P_1-dominated alternatives and selects P_2-best among them. Unlike our model, P_2 is allowed to be cyclic. Hence, our model is a rational shortlist method where P_2 is a preference and P_1 is transitive. Indeed, it is a strict subset of shortlist method. Before we illustrate this, we remind the second axiom, Weak WARP, used in the characterization of the rational shortlist method. The axiom says that if an alternative x is chosen both when only y is also available and when y and other set of alternatives, T, are available, then y is not chosen from any subset of T whenever x is available. Formally,

> **Weak WARP:** Suppose $\{x, y\} \subset S \subset T$. If $x = c(xy) = c(T)$, then $y \neq c(S)$.

Example 2. *The following example satisfies Weak WARP and Expansion but violates A2. There are five alternatives: a, b, c, d, x. The decision maker has two rationales: one is acyclic $P_1 = \{(c, a), (d, b), (a, x)\}$ and the second one is asymmetric including $\{(a, b), (b, c), (c, d), (d, a), (x, d)\}$. Note that P_2 is cyclical. The decision maker sequentially applies P_1 and P_2 to make a choice as in the shortlisting method.*[21] *Now we show that this choice behavior violates A2 at $S = \{a, b, c, d\}$. In other words, there is no alternative in S which serves the role of x^* in the axiom. For example, the alternative a changes the choice ($c(dx) \neq c(adx)$) but it is not chosen ($c(ad) \neq a$ and $\{a, d\} \subset \{a, d, x\}$).*

[21] One can define P_2 completely so that there is a unique survivor of this two-stage elimination, hence this is a well-defined choice function that is a rational shortlist method. Since our aim is to show that it violates A2, we define necessary part of P_2.

4 Conclusion

This paper illustrates when and how one can deduce both the preference and consideration sets of a decision maker who follows a CLC with an attention filter. For instance, if a product is not popular in a market, it is very important for a firm to know the reason, which can be either it is not liked by consumers or it does not attract attentions of consumers. Our model provides a theoretical ground to distinguish these two possibilities and help it choose a right remedy. Similarly, the social planner can find a proper strategy to make sure that people chooses the right option in 401(K), health insurance and so on. Limited attention has been widely studied in economics: neglecting the nontransparent taxes (Chetty and Looney (2009)), inattention to released information (DellaVigna and Pollet (2007)), costly information acquisition (Gabaix et al. (2006)), and rational inattention in macroeconomics (Sims (2003)).

We demonstrate how we can make the identification stronger when we know extra knowledge about her consideration set formation such as a strong attention filter or a transitive order. Since our identification strategy relies on the assumption that a decision maker follows a CLC with an attention filter or its special cases, we provide the characterization for each of our model by weakening the weak axiom of the revealed preference to make it our model behaviorally testable.

References

Apesteguia, J. and M. Ballester, "A characterization of sequential rationalizability," Economics Working Papers, Universitat Autonoma de Barcelona 2008.

Aumann, R. J., "Musings on Information and Knowledge," *Econ Journal Watch, Atlas Economic Research Foundation*, 2005, *2(1)*, 88–96.

Basar, Gzen and Chandra Bhat, "A parameterized consideration set model for airport choice: an application to the San Francisco Bay Area," *Transportation Research Part B: Methodological*, 2004, *38* (10), 889–904.

Bernheim, B. D. and A. Rangel, "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *Quarterly Journal of Economics*, 2009, *124(1)*, 51–104.

Broadbent, Donald E., *Perception and communication*, Pergamon Press, New York, 1958.

Cherepanov, V., T. Feddersen, and A. Sandroni, "Rationalization," 2008. mimeo,Nothwestern University.

Chetty, Raj and Kory Looney Adamand Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, *99*, 114577.

Chiang, Jeongwen, Siddhartha Chib, and Chakravarthi Narasimhan, "Markov chain Monte Carlo and models of consideration set and parameter heterogeneity," *Journal of Econometrics*, November 1998, *89* (1-2), 223–248.

Dawes, P. L. and J. Brown, "The Composition of Consideration and Choice Sets in Undergraduate University Choice: An Exploratory Study," *Journal of Marketing for Higher Education*, 2005, *14(2)*, 37–59.

de Clippel, Geoffroy and Kfir Eliaz, "Reason-Based Choice: A Bargaining Rationale for the Attraction and Compromise Effects," Working Papers, Brown University 2009.

DellaVigna, Stefano and Joshua Pollet, "Demographics and Industry Returns," *American Economic Review*, 2007, *97*, 1167–1702.

Dulleck, Uwe, Franz Hackl, Bernhard Weiss, and Rudolf Winter-Ebmer, "Buying Online: Sequential Decision Making by Shopbot Visitors," Economics working papers 2008-10, Department of Economics, Johannes Kepler University Linz, Austria August 2008.

Eliaz, Kfir and Ran Spiegler, "Consideration Sets and Competitive Marketing," 2007. Mimeo.

— , **Michael Richter, and Ariel Rubinstein**, "An Etude in Choice Theory: Choosing the Two Finalists," 2009. Mimeo.

Filiz-Ozbay, E., "Incorporating Awareness into Contract Theory," 2008. mimeo.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg, "ostly Information Acquisition: Experimental Analysis of a Boundedly Rational Model," *American Economic Review*, 2006, *96* (4), 1043–1068.

Goeree, M., "Limited information and advertising in the US personal computer industry," *Econometrica*, 2007, *76*, 1017–1074.

Gul, F. and W. Pesendorfer, "The Case for Mindless Economics," in A. Caplin and A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, Oxford University Press, USA, 2008, pp. 3–42.

Hauser, John R. and Birger Wernerfelt, "An Evaluation Cost Model of Evoked Sets," *Journal of Consumer Research*, March 1990, *16*, 383–408.

Hausman, D., "Mindless or Mindfull Economics: A Methodological Evaluation," in A. Caplin and A. Schotter, eds., *The Foundations of Positive and Normative Economics: A Handbook*, Oxford University Press, USA, 2008, pp. 125–155.

Heifetz, Aviad, Martin Meier, and Burkhard C. Schipper, "Unawareness, Beliefs and Games," *SSRN eLibrary*, 2007.

Highhouse, S., "Context-dependent selection: The effects of decoy and phantom job candidates," *Organizational Behavior and Human Decision Processes*, 1996, *65*, 68–76.

Hotchkiss, G., S. Jensen, M. Jasra, and D. Wilson, "The Role of Search in Business to Business Buying Decisions A Summary of Research Conducted," 2004. Enquiro White Paper.

Houy, Nicolas, "Rationality and Order-Dependent Sequential Rationality," *Theory and Decision*, 2007, *62* (2), 119–134.

— **and Koichi Tadenuma**, "Lexicographic compositions of multiple criteria for decision making," *Journal of Economic Theory*, 2009, *144* (4), 1770–1782.

Huber, J. and C. Puto, "Market boundaries and product choice: Illustrating attraction and substitution effects," *The Journal of Consumer Research*, 1983, *10* (1), 31–43.

Huber, Joel, John W. Payne, and Christopher Puto, "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *The Journal of Consumer Research*, June 1982, *9* (1), 90–98.

Huberman, Gur and Tomer Regev, "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar," *Journal of Finance*, 02 2001, *56* (1), 387–396.

Jarvis, L.P. and J.B. Wilcox, "A model for predictive measurements of advertising effectiveness.," *Industrial Marketing Management*, 1977, *6*, 9–14.

Kahneman, D. Bateman I., A. Munro, B. Rhodes, C. Starmer, and R. Sugden, "Testing competing models of loss aversion: an adversarial collaboration," *Journal of Public Economics*, 2005, *89*, 1561–1580.

Laroche, Michael, Jerry Rosenblatt, and Ian Sinclair, "Brand Categorization Strategies in an Extensive Problem Solving Situation: A Study of University Choice.," *Advances in Consumer Research*, 1984, *11*, 175–179.

Lavidge, Robert J. and Gary A. Steiner, "A model for predictive measurements of advertising effectiveness.," *Journal of Marketing*, 1961, *25*, 59–62.

Lehmann, Donald R. and Yigang Pan, "Context Effects, New Brand Entry, and Consideration Sets," *Journal of Marketing Research*, 1994, *31* (3), 364–374.

Lleras, J., Y. Masatlioglu, D. Nakajima, and E. Ozbay, "When More is Less: Choice by Limited Consideration," 2009. Mimeo.

Loomes, Graham, Chris Starmer, and Robert Sugden, "Observing Violations of Transitivity by Experimental Methods," *Econometrica*, 1991, *59* (2), 425–39.

Mandler, Michael, Paola Manzini, and Marco Mariotti, "A Million Answers to Twenty Questions: Choosing by Checklist," Working Papers 622, Queen Mary, University of London, Department of Economics 2008.

Manzini, Paola and Marco Mariotti, "Sequentially Rationalizable Choice," *American Economic Review*, December 2007, *97* (5), 1824–1839.

_ **and** _ , "Categorize Then Choose: Boundedly Rational Choice and Welfare," 2008. Queen Mary Working paper n. 561/2006.

_ **and** _ , "Choice based welfare economics for boundedly rational agents," 2009. Mimeo.

_ **and** _ , "Choice Over Time," in "Oxford Handbook of Rational and Social Choice," Oxford University Press, 2009, chapter 10.

Masatlioglu, Yusufcan and Daisuke Nakajima, "A Theory of Choice by Elimination," 2007. Mimeo, Department of Economics, University of Michigan.

May, K. O., "Intransitivity, Utility, and the Aggregation of Preference Patterns," *Econometrica*, 1954, *22*, 1–13.

O'Curry, S. Pan Y. and P. Pitts, "The attraction effect and political choice in two elections," *Journal of Consumer Psychology*, 1995, *4*, 85–101.

Ok, E. A., P. Ortoleva, and G. Riella, "Revealed (P)Reference Theory," 2008. mimeo.

Ozbay, E. Y., "Unawareness and strategic announcements in games with uncertainty," 2008. mimeo.

Pessemier, Edgar A, "Stochastic Properties of Changing Preferences," *American Economic Review*, 1978, *68* (2), 380–85.

Richards, Max D., John E. Sheridan, and John W. Slocum, "Comparative analysis of expectancy and heuristic models of decision behavior," *Journal of Applied Psychology*, 1975, *60* (3), 361–368.

Roberts, J. H. and M. J. Lattin, "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 1991, *28* (4), 429–440.

Rosen, Deborah E., James M. Curran, and Timothy B. Greenlee, "College Choice in a Brand Elimination Framework: The Administrator's Perspective.," *Journal of Marketing for Higher Education*, 1998, *8* (4).

Schwartz, Barry, *The Paradox of Choice : Why More Is Less*, Harper Perennial, January 2005.

Schwartz, Janet A. and Gretchen B. Chapman, "Are More Options Always Better?: The Attraction Effect in Physicians' Decisions about Medications," *Med Decis Making*, 1999, *19* (3), 315–323.

Simon, H. A., *Models of man: social and rational: mathematical essays on rational human behavior in a social setting*, Wiley, 1957.

Sims, Christopher A., "Implications of rational inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Stigler, George J., "The Economics of Information," *Journal of Political Economy*, 1961, *69* (3), 213–225.

Tversky, Amos, "Intransitivity of Preferences," *Psychological Review*, 1969, *76*, 31–48.

_ **and Itamar Simonson**, "Context-dependent preferences," *Management Science*, 1993, *39* (10), 1179–1189.

Varian, Hal, "Revealed Preference," in Michael Szenberg, Lall Ramrattan, and Aron A. Gottesman, eds., *Samuelsonian Economics and the Twenty-First Century*, Oxford University Press, USA, 2006, pp. 99–115.

Wedell, Douglas H., "Distinguishing among models of contextually induced preference reversals," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1991, *17* (4), 767–778.

Wright, P. and F. Barbour, "Phased decision strategies: Sequels to an initial screening," in M. Starr and M. Zeleny, eds., *Studies in Management Sciences, Multiple Criteria Decision Making*, North-Holland, Amsterdam, 1977, pp. 91–109.

Zyman, S., *The End of Marketing as We Know It*, Harper Collins, New York, 1999.

5 Appendix

The Proofs of Proposition 1, Proposition 2 and Theorem 1

The if-parts of Propositions 1, 2(1) and 2(2) have been shown in the main text. The if-part of Theorem 1 is immediate from Lemma 1 because if c is a CLC with an attention filter represented by $(\Gamma, \succ)$ then $\succ$ must include P, which cannot be cyclical so c must satisfy WARP with Limited Attention.

Next, we prove the only-if part of Theorem 1. Suppose c satisfies WARP with Limited Attention. By Lemma 1, P is acyclic so there is a preference $\succ$ that includes P and define

$$\Gamma(S) = \{x \in S \ : \ c(S) \succ x\} \cup \{c(S)\}$$

Then, it is clear that $c(S)$ is the unique $\succ$-best element in $\Gamma(S)$ so all we need to show is that Γ is an attention filter. Suppose $x \in S$ but $x \notin \Gamma(S)$ (so $x \neq c(S)$). By construction, $x \succ c(S)$ so it cannot be $c(S)Px$. Hence, it must be $c(S) = c(S \setminus x)$ so we have $\Gamma(S) = \Gamma(S \setminus x)$.

This also proves the only-if part of Proposition 1 and 2(1). This is because the choice of $\succ$ is arbitrary as long as it includes P (automatically it includes P_R as well), so we can represent c by $(\Gamma, \succ)$ such that $y \succ x$ unless yP_Rx and with $x \in \Gamma(S)$ unless $xP_Rc(S)$.

Finally, we show the only-if part of Proposition 2(2). Suppose there is no T that satisfies the condition. We shall prove that if c is a CLC with an attention filter then it can be represented by some attention filter Γ with $x \notin \Gamma(S)$. If $c(S)P_Rx$ does not hold, we have already shown that c can be represented with $x \succ c(S)$ and $x \notin \Gamma(S)$ so x is not revealed to attract attention at S, so we focus on the case when $c(S)P_Rx$.

Now construct a binary relation, $\tilde{P}$, where $a\tilde{P}b$ if and only if "aP_Rb" or "$a = c(S)$ and not $bP_Rc(S)$." That is, $\tilde{P}$ puts $c(S)$ as high as possible as long as it does not contradict P_R. Since P_R is acyclic and c is represented by an attention filter, one can show that $\tilde{P}$ is also acyclic. Given this, take any preference relation $\succ$ that includes $\tilde{P}$, which includes P_R as well. We have already shown that $\tilde{\Gamma}(S) \equiv \{z \in S : c(S) \succ z\} \cup \{c(S)\}$ is an attention filter and $(\tilde{\Gamma}, \succ)$ represents c. Now define Γ as follows :

$$\Gamma(S') = \begin{cases} \tilde{\Gamma}(S') & \text{for } S' \notin \mathcal{D} \\ \tilde{\Gamma}(S') \setminus x & \text{for } S' \in \mathcal{D} \end{cases}$$

where $\mathcal{D}$ is a collections of sets such that

$$\mathcal{D} = \left\{ S' \subset X : \begin{array}{cc} c(S') = c(S) & \text{and} \\ zP_Rc(S) \text{ for all } z \in (S \setminus S') \cup (S' \setminus S) & \end{array} \right\}$$

That is, Γ is obtained from $\tilde{\Gamma}$ by removing from x for any budget set S' where $c(S) = c(S')$ and any item that belongs to S or S' but not to both is revealed to be better than $c(S)$. Notice that x cannot be $c(S)$ because if so the condition of the statement is satisfied for $T = S$. Hence, $\Gamma(S') \subset \tilde{\Gamma}(S')$ always includes $c(S')$. Furthermore the proof of Theorem 1 shows that $(\tilde{\Gamma}, \succ)$ represents c. Therefore, $(\Gamma, \succ)$ also represents c so we only need to show that Γ is an attention filter.

To do that, it is useful to notice that $\tilde{\Gamma}$ is an attention filter and $c(T') = c(T'')$ whenever $\tilde{\Gamma}(T') = \tilde{\Gamma}(T'')$ because $(\tilde{\Gamma}, \succ)$ represents c.

Suppose $y \notin \Gamma(T)$. We shall prove $\Gamma(T) = \Gamma(T \setminus y)$.

Case I: $y = x$

If $T \notin \mathcal{D}$, then we have $\Gamma(T) = \tilde{\Gamma}(T) = \tilde{\Gamma}(T \setminus x) = \Gamma(T \setminus x)$. If $T \in \mathcal{D}$, then it must be $c(T) = c(T \setminus x)$ (otherwise, the condition of the statement is satisfied) so by construction of $\tilde{\Gamma}$ and Γ, we have $\Gamma(T) = \tilde{\Gamma}(T) \setminus x = \tilde{\Gamma}(T \setminus x) = \Gamma(T \setminus x)$.

Case II: $T \in \mathcal{D}$ and $y \neq x$

Since $y \notin \Gamma(T)$ is equivalent to $y \notin \tilde{\Gamma}(T)$, we have $\tilde{\Gamma}(T) = \tilde{\Gamma}(T \setminus y)$. Therefore, $c(T \setminus y) = c(T) = c(S)$. By construction of Γ and $\tilde{\Gamma}$, it must be $y \succ c(S)$, which implies $yP_R c(S)$ by construction of $\succ$. Therefore, $T \setminus y \in \mathcal{D}$. Therefore, $\Gamma(T) = \tilde{\Gamma}(T) \setminus x = \tilde{\Gamma}(T \setminus y) \setminus x = \Gamma(T \setminus y)$.

Case III: $T \notin \mathcal{D}$ and $y \neq x$

If $T \setminus y \in \mathcal{D}$, analogously to the previous case, we have $c(T) = c(T \setminus y) = c(S)$ and $yP_R c(S)$ so it must be $T \in \mathcal{D}$, which is a contradiction. Hence, $T \setminus y \notin \mathcal{D}$ so we have $\Gamma(T) = \tilde{\Gamma}(T) = \tilde{\Gamma}(T \setminus y) = \Gamma(T \setminus y)$.

The Proof of Theorem 2

Define $xP''y$ if and only if there exist T and T' with $x, y \in T \subset T'$ such that

$$x = c(T) \text{ and } c(T') \neq c(T' \setminus y)$$

Lemma 2. *P'' is acyclic if and only if c satisfies Strong Consistency.*

The proof of Lemma 2 is completely analogous to earlier Lemma, so we skip it.

Let P''_R be the transitive closure of P'' and let $\succ$ be any arbitrary completion of P''_R. For every S, we call $B \subset S$ is a minimum block of S if and only if $c(S) \neq c(S \setminus B)$ but $c(S) = c(S \setminus B')$ for any $B' \subsetneq B$. Given this, define Γ recursively as follows:

1. $\Gamma(X)$ consists of the $\succ$-worst element of each of X's minimum block.
2. Suppose Γ has been already defined for all proper supersets of S. Then, define $\Gamma(S)$
   (a) First, put $x \in S$ into $\Gamma(S)$ if $x \in \Gamma(T)$ for some $T \supsetneq S$.
   (b) If there is a minimum block of S that does not have an element in $\Gamma(S)$ according to the above, pick the $\succ$-worst element into $\Gamma(S)$.

Lemma 3. *For any S,*

(i) *$\{c(S)\}$ is a minimum block of S. There is no other minimum block that includes $c(S)$.*

(ii) *If B is a minimum block of S other than $\{c(S)\}$, then $c(S) \succ x$ for all $x \in B$.*

(iii) *If $c(T) \neq c(S)$ and $T \supsetneq S$, then T has a minimum block that is a subset of $T \setminus S$.*

Proof. Part (i) and (iii) are obvious so only prove Part (ii). Let $B' = B \setminus x$ (it may be empty). Then we have

$$c(S) = c(S \setminus B') \neq c((S \setminus B') \setminus x)$$

Therefore, we have $c(S)P''x$ so it must be $c(S) \succ x$. □

Claim 1. Γ *is a strong attention filter.*

Proof. Γ is an attention filter by construction so we shall prove that Γ is a strong attention filter. Suppose $x, y \in S$, $x, y \notin \Gamma(S)$, but $y \in \Gamma(S \setminus x)$. Then there exists $T \supset S$ such that (i) $T \setminus x$ has a minimum block B and y is the worst element in B and (ii) none of elements in B is included in $\Gamma(T')$ for any $T' \supsetneq T \setminus x$.

Then, we must have $c(T) = c(T \setminus x)$. Otherwise $\{x\}$ is a minimum block of T' so we have $x \in \Gamma(T')$ that implies $x \in \Gamma(S)$. Therefore, we have

$$c(T) = c(T \setminus x) \neq c((T \setminus x) \setminus B) = c(T \setminus (\{x\} \cup B))$$

Therefore, by Lemma 3 (iii), T has a minimum block that is a subset of $x \cup B$ so at least one element in $x \cup B$ must be in $\Gamma(T)$, which is a contradiction. □

Now we want to show that $(\succ, \Gamma)$ represents c. Since Lemma 3 (i) implies that $c(S) \in \Gamma(S)$, all we need to show is that $c(S) \succ y$ for all $y \in \Gamma(S) \setminus c(S)$.

Claim 2. *If $y \in \Gamma(S)$ and $y \neq c(S)$, then $c(S) \succ y$.*

Proof. Since $y \in \Gamma(S)$, there exists $T \supset S$ such that $y \in \Gamma(T)$. Furthermore, T has a minimum block B where y is the worst element and none of elements in B is in $\Gamma(T')$ for any $T' \supsetneq T$. There are three easy cases: (i) if $c(S) = c(T)$ then by Lemma 3 (ii) we have $c(S) = c(T) \succ y$, (ii) if $y = c(T)$ then we have $c(S)P''y$ so it must be $c(S) \succ y$, and finally (iii) if $c(S) \in B$, then $c(S) \succ y$ by the construction. Therefore, we only need investigate the case when $y \neq c(T) \neq c(S)$ and $c(S) \notin B$. Note that $c(T) \succ y$ in this case by Lemma 3 (ii).

Now let $S' = S \setminus B$. Since $y \in B$, S' is a proper subset of S.

Case I: $c(S'') \neq c(S)$ for some S'' where $S' \subset S'' \subset S$.

By Lemma 3 (iii), S has a minimum block B' that is a subset of $S \setminus S'' \subset B$. Since $c(S) \notin B'(\subset B)$, every element in B' is worse than $c(S)$ by Lemma 3 (ii). Since y is the worst element in B that is a superset of B', we conclude $c(S) \succ y$.

Case II: $c(S'') = c(S)$ for all S'' where $S' \subset S'' \subset S$.

Since $y \neq c(T) = c(T \setminus (B \setminus y)) \neq c(T \setminus B)$, and $c(S \setminus (B \setminus y)) \in T \setminus (B \setminus y)$, we have $c(S \setminus (B \setminus y))\ P''\ y$. Therefore, $c(S) \succ y$ because of $c(S \setminus (B \setminus y)) = c(S)$. □

The Proof of Theorem 3

The if-part is demonstrated in the main body. For the only-if part, notice that A3 and A4 imply that P is acyclic so the proof of Theorem 1's only-if part is applicable by setting $\succ$ to be equal to P^*. Given this, it is easy to see that Γ constructed in the proof has the property $\Gamma(S) = S$ whenever $|S| = 2$.

The Proof of Proposition 4

The if-parts of both revealed attention and inattention can be shown exactly in the same way as those in Proposition 2. The only-if part of the revealed inattention can be obtained by the proof of Theorem 3, as we show that $\Gamma(S) = \{x \in S : c(S)\, P^* x\} \cup \{c(S)\}$ represents c along with $\succ = P^*$.

As for the only-if part of the revealed attention, we argue that the proof of the only-if part (revealed attention) of Proposition 2 is applicable by setting $\succ$ to be equal to P^*. The only issue is that Γ constructed in that proof satisfies the property $\Gamma(S) = S$ whenever $|S| = 2$. To see this, notice that $\widetilde{\Gamma}$ has this property so it is enough to verify $\{x, a\} \notin \mathcal{D}$ for any y whenever there is no T satisfying the properties stated in the Proposition 2. Be reminded that this implies $c(S) = c(S \setminus x)$ particularly. Suppose the contrary that $\{x, a\} \in \mathcal{D}$. Then $c(S) = c(xa)$, where $c(S)$ cannot be x. so $c(S \setminus x) = c(S) = c(xa) = a$. Hence, we have $a \in S$ and aP^*x. Furthermore, every element in S other than x or a must be P^*-better than a. Therefore, a is the P^*-worst element in $S \setminus x$.

However, a decision maker who follows CLC with an attention filter can never choose the worst element if she has full attention at all binary sets. If so, she must not pay attention to any other alternative in that decision problem. Then, removing items other than the worst one to make a doubleton does not change her attention so she must necessarily have inattention at some binary problem. Therefore, we have a contradiction and conclude that it cannot be $\{x, a\} \in \mathcal{D}$.

The Proof of Theorem 4

We have already shown the if-part of the statement in the main text so we shall show the only-if part. Take any completion of P'', denoted by $\succ$. (P'' is defined in the proof of Theorem 2. Such $\succ$ exists because Strong Consistency guarantees that P'' is acyclic.) Then define $x \rhd' y$ if and only if

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

Since $x \rhd' y$ only if $y \succ x$ and $\succ$ is a preference, $\rhd'$ is acyclic. Let $\rhd$ be the transitive closure of $\rhd'$. It is acyclic as well. Therefore, all we need to show is that $(\Gamma_{\rhd}, \succ)$ represents c.

Lemma 4. *If $y \succ c(S)$, then $y \notin \Gamma_{\rhd}(S)$*

Proof. Let $x = c(S)$. The statement is trivial when $y \notin S$ so assume $y \in S$. Since $\rhd$ is the transitive closure of $\rhd'$, it is $\Gamma_{\rhd}(S) \subset \Gamma_{\rhd'}(S)$ so it is enough to show $y \notin C_{\rhd'}(S)$.

The statement is true when $|S| = 2$ by the definition of $\rhd'$. Suppose there exists S and $y \in S$ such that $y \succ x$ and $y \in \Gamma_{\rhd}(S)$. Without loss of generality, assume that S has the smallest cardinality among such sets. We shall lead a contradiction by showing several claims.

Claim 3. *For any $S' \subsetneq S$, $x \neq c\,(S')$ whenever $y \in S'$.*

Proof. Since $y \in \Gamma_{\rhd'}(S)$, it must be $y \in \Gamma_{\rhd'}(S')$ as well. If $x = c(S')$ then this violates the assumption that S has the smallest cardinality at which the statement of Lemma 4 is violated. □

Now consider all budget sets that can be obtained by removing one element from S. Notice that there are $|S|$ such decision problems and only elements in S may be chosen from those sets.

Claim 4. *For any $z \in S$, there exists $z' \in S \setminus \{z\}$ such that $z = c\,(S \setminus z')$.*

Proof. Suppose not. By the pigeonhole principle, there must exist $\alpha \in S$ such that

$$\alpha = c\,(S \setminus \beta) = c\,(S \setminus \gamma)$$

for some distinct $\beta, \gamma \in S$. By Expansion, $c\,(S) = \alpha$ so it must be $\alpha = x$. Since y must be included either in $S \setminus \alpha$ or $S \setminus \beta$, Claim 3 implies that $x \neq c(S \setminus \alpha)$ or $x \neq c(S \setminus \beta)$. This is a contradiction. □

Claim 5. *$x = c\,(S \setminus y)$ and $y = c\,(S \setminus x)$.*

Proof. The combination of Claim 3 and 4 immediately implies $x = c\,(S \setminus y)$. By Claim 4, y must be chosen from $S \setminus z$ for some $z \in S$. If $z \neq x$, then then we have $yP'x$ so it must be $y \succ x$. Therefore, it cannot be $y \rhd' x$. Hence, $y = c\,(S \setminus x)$. □

Now take any $z \in S \setminus \{x, y\}$. (Notice that $|S| \geq 3$). If $|S| = 3$, Claim 4 requires $z = c(S \setminus x)$ or $z = c(S \setminus y)$ but both possibilities are excluded by Claim 5. Suppose $|S| \geq 4$. Let $\alpha = c(S \setminus z)$. By Claim 4 and Claim 5, $\alpha \in S \setminus \{x, y, z\}$. Hence, we have $\alpha P' x$ so it must be $\alpha \succ x$. Now consider $c(S \setminus \alpha)$, which must be something other than x. Hence, we have $xP'\alpha$ so is $x \succ \alpha$. This is a contradiction. Therefore, there is no S such that $x = c(S)$ but $y \in S$ so Lemma 4 is proven.

Lemma 5. $c(S) \in \Gamma_{\triangleright}(S)$.

Proof. Let $x = c(S)$ but there exists $y \in S$ such that $y \triangleright x$. If $y \triangleright' x$ (i.e. before taking the transitive closure), then it must be $x \succ y$ and $c(xy) = y$. $c(xy) = y$ and $c(S) = x$ imply $yP'x$ so we cannot have $x \succ y$, which is a contradiction. Therefore, it cannot be $y \triangleright' x$. so there must exist $z_1, \ldots, z_k$ such that

$$y \triangleright' z_1 \triangleright' z_2 \triangleright' \cdots \triangleright' z_k \triangleright' x.$$

By definition of $\triangleright'$ it must be

$$x \succ z_k \succ \cdots \succ z_2 \succ z_1 \succ y$$

and

$$c(yz_1) = y,\ c(z_1 z_2) = z_1, \ldots, c(z_{k-1} z_k) = z_{k-1},\ c(z_k x) = z_k$$

Since $x \succ y$, we cannot have $y = c(xy)$ (if so, it would be $y \triangleright' x$ and we have shown that it would lead a contradiction.) so it must be $x = c(xy)$.

Now consider $c(xyz_1 \ldots z_k)$. It cannot be x because, if so, it must be $z_k P' x$ so is $z_k \succ x$. It cannot be z_i (because if so $z_{i-1} \succ z_i$ or $y \succ z_1$). Therefore, it must be $y = c(xyz_1, \ldots, z_k)$.

Since $x = c(xy)$, there must exist i such that:

$$y = c(xyz_i z_{i+1} \ldots z_k) \neq c(xyz_{i+1} \ldots z_k)$$

which implies $yP'z_i$ so $y \succ z_i$. This is a contradiction. Therefore, we conclude that there is no $y \in S$ such that $y \triangleright x$. □

These two lemmas prove that c is represented by $(\Gamma_{\triangleright}, \succ)$. □

The Proof of Proposition 5

The if-parts of both the revealed preference and the revealed order are shown in the mainbody. To prove the only-if parts, the proof of the only-if part of Theorem 4 is applicable. If not $xP'_R y$, it has been shown that a preference with $y \succ x$ can represent c. If $yP'_R x$ but $x = c(xy)$, then we indeed define $x \triangleright y$ to represent c.