Revealed Attention

Yusufcan Masatlioglu Daisuke Nakajima Erkut Y. Ozbay

November 2009

Abstract

The standard revealed preference argument relies on an implicit assumption that a decision maker considers all feasible alternatives. However, the marketing and psychology literatures provide well-established evidence that consumers do not consider all brands in a given market before making a purchase (Limited Attention). In this paper, we illustrate how one can deduce both the decision maker's preference and the alternatives to which she gives attention and inattention from the observed behavior. Further, we provide a choice theoretical foundation for maximizing a single preference relation under limited attention.

JEL Classi cation: D11, D81.

Keywords: Revealed Preferences, Awareness, Attention, Consideration Set

1 Introduction

Revealed preference is one of the most in uential ideas in economics and has been applied to a number of areas of economics, including consumer theory.[1] According to 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 (Samuelson (1938)). Any choice reversal, therefore, observed both empirically and experimentally, is attributed to irrationality since it cannot be expressed as a preference maximization.

The revealed preference argument relies on the implicit assumption that a decision maker (DM) considers all feasible alternatives. Under the full consideration assumption, the standard revealed preference method can be misleading. Although the 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)). For example, while using a search engine, a DM might only pay attention to alternatives appearing on the rst page of the results since it takes too much time to consider all the search results. She then picks the best alternative of those on the rst page, say y. It is possible that her most preferred item, x, does not appear in the rst page. Therefore we, as

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

outside observers, cannot conclude y is better than x even though y is chosen when x is available. Nevertheless, as in the above example, the DM may have a well-defined preference and is maximizing her preference within her bounded understanding of what is available.[2] Therefore, it follows to ask how to elicit the preferences without the full attention assumption. In this paper, we consider a DM who picks her most preferred item from the alternatives she pays attention to, not from the entire feasible set. We then illustrate when and how one can deduce both the DM's preferences and the alternatives to which she does or does not pay attention from her observed choices.

The marketing literature calls the set of alternatives to which a DM pays attention in her choice process as the *consideration set* (Wright and Barbour (1977)). The formation of the consideration set has been extensively studied in marketing and finance literatures (e.g. Hauser and Wernerfelt (1990); Roberts and Lattin (1991)). It has been argued that due to cognitive limitations, DMs 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 limited cognitive capacity (possibly stemming from unawareness as demonstrated in Goeree (2007)[3]), restricts her attention only to a small fraction of the objects present in the associated market (Stigler (1961); Pessemier (1978); Chiang et al. (1998)).[4] In sum, DMs intentionally or unintentionally filter out some alternatives to prevent her cognitive capacity from being overloaded (Broadbent (1958)).

The common property in the formation of consideration sets is that it is unaffected when an alternative she does not pay attention to becomes unavailable. This basic property of the attention filter, which is also documented in the psychology literature (Broadbent (1958)), can be interpreted as the minimal condition. This property is trivially satisfied in the classical choice theory where it is assumed that the DM is able to pay attention to all the available alternatives. Additionally, it is normatively appealing especially when a DM pays attention to all of items she is aware of and is

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

[3] Lavidge and Steiner (1961) presented awareness of an item as a necessary condition to be in the consideration set. 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)).

[4] In addition, 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)), and airport choice (Basar and Bhat (2004)).

unaware that she is unaware of other items. For example, if a PC buyer is not only unaware of a particular PC, but she is also unaware that she overlooks that PC, then, even when that PC becomes unavailable, such a change will not be recognized by her. Therefore, her consideration set will stay the same.

Interestingly, the property is also satisfied when the DM actually chooses the consideration sets by taking the cost of investigation and the expected benefit into account. Suppose the DM excludes x from her consideration. If x becomes unavailable, she has no reason to add or remove any alternative to her consideration set because she could have done so when x was available. Therefore, her consideration set is not affected when x becomes unavailable. Furthermore, this property is also satisfied when the formation is based on many decision heuristics, such as paying attention only to the N-most advertised alternatives or the products appearing in the first page of search results.

In this paper, we refer to the consideration sets satisfying this property as *attention filters*. Under this structure, it is possible to elicit the DM's preference whenever a choice reversal is observed.[5] For example, assume that she chooses x but removing y changes the choice. This can happen only when her consideration set has changed. This would be impossible if she did not pay 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 unchosen alternative, the initially chosen alternative is preferred to the removed one.

Given that our identification strategy relies on the particular choice procedure, where she maximizes her preference within her attention filter, it is natural to ask the falsifiability of our model. We show that our model is fully characterized by weakening the Weak Axiom of the Revealed Preference (WARP). This result renders our model behaviorally testable.

Our method to distinguish between a preference and attention/inattention generates several policy implications. For instance, if a product of a firm is unpopular in the market place, there could be two different explanations: (i) the product has a low evaluation by consumers and/or (ii) it does not the attract attention of consumers. Identifying the right reason will lead to different strategies for the firm to improve the sales.

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

Another implication of our method is a welfare analysis based on the revealed preference. We elicit the revealed preference in a positive approach that is based on a particular choice procedure similar to Green and Hojman (2009).[6] Bernheim and Rangel (2009) criticize the revealed preference approach by arguing that a rationalization of the behavior is not needed for the welfare analysis. Instead, they develop a choice theoretic framework that generates welfare analysis directly from choice data without assuming any particular choice procedure (model-free).[7] They claim that y is strictly welfare improving over x if x is never chosen when y is present. However, this intuitive criteria of welfare analysis is meaningful only if the DM considers all the presented alternatives.[8] In Section 3, we discuss this issue in detail and provide an example where their welfare implication contradicts our revealed preference; that is, y is revealed to be preferred to x even when x is strictly welfare improving over y.

So far we have discussed how one can elicit DM's preference and consideration sets in our model. In doing so, we impose a relatively weak condition on the formation of consideration sets so that our approach is applicable to a wide range of choice data. As a result, although our model is refutable, it provides a simple unifying explanation for several frequently observed behaviors that cannot be captured by the standard choice theory: Attraction E ect, Cyclical Choice, and Choosing Pairwisely Unchosen (see Anomalies section). Our explanations for these choice patterns solely depend on limited attention, hence seemingly irrational behaviors can be explained without introducing changing preference. Nevertheless, depending on the intended application, it is possible to analyze this framework under di erent restrictions on consideration sets (see Concluding Remarks).

There are several related models where the nal choice is made after eliminating several items, which can be interpreted as a choice from limited consideration. The elimination is done by applying a rationale (Manzini and Mariotti (2007); Apesteguia and Ballester (2008))[9], focusing only on alternatives a decision maker can rationalizes to choose by some other criteria (Cherepanov et al. (2008)) or considering only alternatives belonging to undominated categories (Manzini and Mariotti (2008)). Unlike

[6]A similar approach is followed by Ambrus and Rozen (2008); Cherepanov et al. (2008); Manzini and Mariotti (2008); Masatlioglu and Nakajima (2007).

[7]Chambers and Hayashi (2008); Noor (2009) and Rubinstein and Salant (2009) are other interesting papers discussing this normative approach.

[8]Indeed, ? mention that if we know the DM believes that she is choosing from a set that is other than the objective feasible set, we should take it into account for the welfare analysis. (Section III B)

[9]Other papers studying sequential application of rationales are Houy (2007); Houy and Tadenuma (2009).

our model, these models implicitly assume that a DM considers all feasible alternatives at the rst stage and *intentionally* eliminates several alternatives. Therefore, their stories are not applicable to cases where the source of limited consideration is unawareness of some alternatives.

Finally, we would like to compare several other models involving consideration sets in decision theory. Lleras et al. (2009) study a di erent model of choice under limited consideration where a product attracting attention in a crowded supermarket shelf will be noticed when there are fewer products. While this paper is complementary to our paper, their implications are completely di erent (see Concluding Remarks for more discussion). Masatlioglu and Nakajima (2007) propose a model of an iterative search where a decision maker cannot consider all alternatives, which can be because of unawareness like our model. The di erence is that they emphasize that a consideration set depends on the initial starting point and evolves dynamically during the course of search. In the models of Dean and Caplin (2008), a decision maker goes through alternative sequentially and, at any given time, chooses the best one among those she has searched. Unlike our model, their \choice process" data includes not only her choice without time limit, but also what she would choose if she were suddenly forced to quit the search at any given time.

Eliaz and Spiegler (2007) analyze a market where rms would like to manipulate consumers' consideration sets by using costly marketing devices. Eliaz et al. (2009) study a very concrete and reasonable way to construct a consideration set. Indeed, some of consideration sets we shall present as examples are within their models. However, contrary to our model, in their paper, the decision maker's consideration set (called *finalists*) is observed and is directly investigated. In our model the consideration set is an object that must be inferred from the DM's nal choice.

2 Choice with Limited Consideration and Attention Filter

Let X be an arbitrary non-empty nite 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 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.

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:

▷ **Top N:** A decision maker considers only top N alternatives according to some criterion that is different from her preference. For instance:

- Consider only the three cheapest suppliers in the market (Dulleck et al. (2008)).
- Consider the N most advertised products in the market.
- Consider the products that appear in the first page of the websearch and/or sponsored links (Hotchkiss et al. (2004)).

▷ **Top on each criterion:** A decision maker has several criteria and considers only the best alternatives in each criterion. For instance:

- Consider only a job candidate who is the best in each program. Or consider the top two job candidates from all first-tier schools and the top candidate from second-tier schools.
- Consider only the cheapest car, the safest car, and the most fuel efficient car in the market.

▷ **Most popular category:** A decision maker considers alternatives that belong to the most popular "category" in the market. For instance:

- There are several bike shops in the DM's town. The DM first checks online to find the store offering the most varieties of bikes and goes to that one. Therefore, the DM only considers bikes sold in the selected store.[10]

[10] For instance, suppose store A deals with Maker 1 and 2's bikes while store B sells bikes from Maker 2 and 3. Then, the DM compares the number of Maker 1 and 2's bikes with that of Maker 2 and 3's to choose which store to visit.

- Zyman (1999) provides real-world evidence for such behavior. The sale of Sprite is increased dramatically when they are simply repositioned from the category of lemon-limes (less popular category) to soda (more popular category).

▷ **Satisficing** (Simon (1957)): Suppose a consumer searches feasible items according to a particular order (for instance, from cheapest to most expensive). She stops searching when she finds an item that is satisficing (if none of the feasible items are satisficing, she ends up considering all of them). More generally, her stopping rule may arbitrarily depend on what she has seen so far (?). For instance, the DM continues to search until she finds something whose utility value exceeds the doubled average among those she has seen so far.

▷ **Iterative search** (Masatlioglu and Nakajima (2007)): Consider a DM who is buying a house. She starts by considering an exogenously given house. First she compares the house with others in its neighborhood. If it is the best, she buys it. Otherwise she will move to the best one and then compare it with its neighborhood houses again. She continues this process until she reaches the house that is “locally” best. In this example, her consideration set consists of all houses she compares during her search process.

▷ **Searching more when the decision is tough:** Several items are hard to find even if it is feasible to find them. The decision maker first considers alternatives that are feasible and easy to find and if there is an item that dominates all others, she chooses it immediately. Otherwise, she makes an extensive search to find all feasible items. In the former case, the consideration set consists only of easily found (and feasible) alternatives and in the latter case it coincides with the feasible set.

The common property across all of these examples is that an item that does not attract the DM's attention does not change her attention span at all. For instance, if a product that would appear only in the twenty-first page of a search result becomes unavailable, those appearing in the first page stay in the first page. If she is following the sacrificing procedure described above, an alterative to which she does not pay attention comes after some satisficing alternative. Therefore removing that item does not change the timing of stopping so she will search from exactly the same list of alternatives. We call this property of the consideration sets formation an *attention filter*. Formally,

Definition 1. *A consideration set mapping is an* ***attention filter*** *if for any* S,

$(S) = (S \setminus x)$ *whenever* $x \notin (S)$.[11]

As we discuss in the introduction, this property is plausible if the source of limited attention is complete unawareness, that is, a DM is not only unaware of some feasible alternatives but is also unaware that she does not consider these alternatives. If so, when some alternative she is unaware of becomes infeasible, she cannot perceive that change so her attention span does not change.

The other case where this property makes sense is when a DM chooses her consideration set rationally. At the beginning, she is not sure of the values of alternatives but makes some inference.The DM then optimally restricts her attention to a smaller number of alternatives. The inclusion of an alternative in the consideration set is based on the trade-o between the consideration's cost and bene t. For instance, suppose that she considers only a and b when her feasible set is $\{a, b, c, d\}$. Now imagine that d becomes unavailable. She has no reason to add c to her consideration set because she could have done so when d was available. For the same reason, it is not rational to remove b from her consideration set. Therefore, her consideration set should not be a ected when an alterative she does not pay attention to becomes infeasible.

3 Revealed Preference and (In)Attention

In this section, we illustrate how to infer (1) the DM's preference and (2) what the DM pays (and does not pay) attention to from her observed choice that is a CLC with an attention lter. The standard theory concludes that x is preferred to y when x is chosen while y is available. To justify such an inference, one must implicitly assume that she has paid attention to y. Without this hidden assumption, we cannot make any inference because she may prefer y but overlooks it. Therefore, eliciting the DM's preference is no longer trivial because her choice can be attributed to her preference or to her inattention.[12]

This observation suggests that multiple pairs of a preference and an attention lter can generate the same choice behavior. To illustrate this, consider the choice

[11] Throughout the paper, we write $S \setminus x$ instead of $S \setminus \{x\}$ for simplicity.

[12] In the extreme case where choice data satis es the weak axiom of revealed preference, we have no way of knowing whether the decision maker is aware of all alternatives and maximizing a particular preference, or whether she only pays attention to the one she chooses. In the latter, her preference has no signi cant importance. In the Concluding Remarks Section, we discuss the situations where one can pin down the preference even in this extreme case.

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
$z \succ_1 x \succ_1 y$	$_1$	xy	xy	y	xz
$x \succ_2 y \succ_2 z$	$_2$	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 2. *Assume c is a choice by limited attention and there are k different pairs of preference and attention filter representing c, $(\,_1, \succ_1), (\,_2, \succ_2), \ldots, (\,_k, \succ_k)$. In this case,*

- *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 $\,_i(S)$ includes x for all i,*
- *x is revealed **not** to attract attention at S if $\,_i(S)$ excludes x for all i.*

If one wants to know whether x is revealed to be preferred to y, it would appear necessary to check for every $(\,_i, \succ_i)$ whether it represents her choice or not. However, this 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 is an attention filter, it is possible to determine the relative ranking between x and y. To see this, 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, de ne:

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

By the argument analogous to the one above, if xPy then x is revealed to be preferred to y. 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 also be part of her revealed preference. One may wonder whether some revealed preference is overlooked by P_R. The next theorem states that the answer is no: P_R is the revealed preference.

Theorem 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 while y is present. Bernheim and Rangel (2009) de nes that y is strictly welfare improving over x in their model-free approach. However, in our setup, x may be revealed to be preferred to y even though x is never chosen in the presence of y. For example, 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 conduct welfare analysis.[13]

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*).

[13]For a detailed discussion of 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 $x \in$ (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 (S) or (T). Therefore, removing those items from S or T cannot change her consideration set. Hence, we have

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

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

Theorem 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 $xP_Rc(S)$,

(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) $yP_Rc(S)$ for all $y \in S \setminus T$,
$zP_Rc(T)$ for all $z \in T \setminus S$.

4 Characterization

The two preceding theorems characterize revealed preference and revealed (in)attention assuming that observed choice behavior is a CLC with an attention lter. Without this assumption, these theorems are not applicable. Therefore, a question to ponder is: how can we test whether choice data is consistent with CLC with an attention lter? Surprisingly, it turns out that CLC with an attention lter can be simply characterized by only one observable property of choice.

Before we state our property, recall the su cient 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 stating 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 for the possibility of limited attention. To conclude that x^* is chosen from T, we not only need to make sure that the chosen element from T is in S and x^* is available but also that x^* attracts attention. As we have discussed, we can infer this when removing x^* from T changes the DM's 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:

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

WARP with Limited Attention indeed guarantees that the binary relation P de- ned in (1) is acyclic and it fully characterizes the class of choice functions generated by an attention lter. 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 3. *(Characterization) c satisfies WARP with Limited Attention if and only if c is a CLC with an attention filter.*

Theorem 3 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 *à* la Samuelson and to derive the DM's preferences and attention filter based on Theorem 1 and 2 from the observed choice data.

5 Anomalies

Our limited attention model is capable of accommodating several frequently observed behaviors: Attraction Effect, Cyclical Choice, and Choosing Pairwisely Unchosen. We will overview them and illustrate how our model accommodates them. In addition to that, we elicit the DM's preference, attention and inattention from choice data.

Attraction Effect

The attraction effect refers to a phenomenon where adding an inferior alternative to the choice set affects the choice.[14] It is observed that consumers are more likely to choose a product when it dominates another product in the set. Lehmann and Pan (1994) experimentally show that introducing new products causes an attraction effect particularly by affecting the composition of consideration sets. Consumers include products in the consideration set that are clearly superior to some other product. Our explanation for the attraction effect is in line with their findings.

For illustration purposes, suppose that there are two products, x and y. If the DM considers both alternatives, she will find out that x is better than y although there is a trade off. Now imagine that the DM faces a budget set which includes a third inferior product d_y, that is dominated by y in every aspect but not by x. The presence of an inferior element, d_y, attracts the DM's attention towards y while ignoring x when

[14] This phenomenon is well-documented and robust in behavioral research on marketing (Huber et al. (1982); Tversky and Simonson (1993)), including in choice among monetary gambles, political candidates, job candidates, environmental issues, and medical decision making. Advertising inferior alternatives is commonly used as a marketing strategy to invoke the attraction effect on the customers.

$\{x, y, d_y\}$ is presented. Thus, she considers y and d_y (and, more importantly, she will not consider x) and chooses y from $\{x, y, d_y\}$, which will generate a typical attraction e ect choice pattern with three alternatives:

$$c(xyd_y) = y, \quad c(xy) = x, \quad c(yd_y) = y, \quad c(xd_y) = x.$$

While most of the theoretical literature on attraction e ects can only accommodate choice sets with three alternatives, our model can generate attraction e ect choice patterns when there are more than one inferior yet attracting alternatives.[15] For example, assume that there is a fourth product d_x which is clearly dominated by x but not by y.Her choice will be:

$$c(xyd_xd_y) = x, \quad c(xyd_y) = y, \quad c(xy) = x.$$

Here, as in the previous example, the presence of d_y attracts the DM to pay attention to y while ignoring x when $\{x, y, d_y\}$ is presented. However, adding d_x (which is dominated by x but not by y) to the alternative set attracts the attention of the DM to x again. Therefore, when $\{x, y, d_x, d_y\}$ is presented, the DM pays attention to x and chooses x as she would when choosing x from x, y.[16]

We illustrated that the notion of attraction e ect is consistent with our model. We now consider another implication: one can elicit the preference of a decision maker whose choice behavior follows the same pattern above without any knowledge about the DM's preference and consideration sets. Given the discussion above, it is true that this choice pattern is within our model. Therefore, Theorem 1 and 2 are applicable. First, we can conclude that she pays attention to d_x at $\{x, y, d_x, d_y\}$ because otherwise her attention span would be the same between $\{x, y, d_x, d_y\}$ and $\{x, y, d_y\}$, resulting in the same item being chosen. Therefore, x must be preferred to d_x. Similarly, the DM pays attention to d_y at $\{x, y, d_x\}$ and she prefers y over d_y. Therefore, our approach will elicit the clear dominance as a revealed preference: $x \succ d_x$, $y \succ d_y$.

Cyclical Choice

May (1954) provides the rst experiment where cyclical choice patterns are observed and these results have been replicated in many di erent choice environments

[15] There are two exceptions: Ok et al. (2008) and de Clippel and Eliaz (2009). While Ok et al. (2008) propose a reference-dependent choice model that accounts for the attraction e ect, de Clippel and Eliaz (2009) provide a characterization for both the attraction and compromise e ects which are derived from an intra-personal bargaining problem among di erent selves of an individual. However, these models cannot accommodate both Cyclical and Choosing Pairwisely choice patterns.

[16] The generalization of the attraction e ect lies outside of recent models provided in Cherepanov et al. (2008), Manzini and Mariotti (2007), and Lleras et al. (2009) since it does not satisfy the axiom called Weak WARP.

(e.g. Tversky (1969); Loomes et al. (1991); Manzini and Mariotti (2009b); Mandler et al. (2008)). In a cyclical choice pattern:

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

We have already illustrated that this choice patter can be captured by our model at the beginning of Section 3. Now let us elicit her preference. Since she exhibits a choice reversal when y is removed from $\{x, y, z\}$, we can identify that y attracts her attention when these three elements are present. Therefore, we can conclude that she prefers x over y. However, as illustrated before, we cannot determine the ranking of z.

Choosing Pairwisely Unchosen

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

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

Since removing x or y from $\{x, y, z\}$ changes her choice, it is revealed that z is better than x and y but we cannot determine her preference between x and y. Since her revealed preference has no cycle, her behavior is captured by our model through Lemma 1 and Theorem 3.

Note that the best element, z, is not chosen in any binary choice so we can conclude that she pays attention to z only when x and y are present. Applying Theorem 2, we pin uniquely down her consideration set except for $\{x, y\}$:

Revealed Preference	zP_Rx and zP_Ry			
	xyz	xy	yz	xz
Revealed Attention	xyz	x	y	x
Revealed Inattention	-	-	z	z

Table 2: Choosing Pairwisely Unchosen

Here, the DM pays attention to z only when both x are y are available. One possible story that generates such an attention pattern is \searching more when the decision is tough" (as discussed in Section 2). Suppose her true preference is $z \succ x \succ y$ where the decision between x and y is very tough and z is hard to nd. She makes an extensive search to nd z only if she sees both x and y. If either x or y is missing, she does not bother to search, and therefore overlooks z.

6 Concluding Remarks

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)). For example, Goeree (2007) shows that relaxing the full attention assumption by allowing customers to be unaware of some computers in the market is enough to explain the high markups in the PC industry.

In this paper, we study the implications of limited attention on revealed preference. We illustrate when and how one can deduce both the preference and consideration sets of a DM who follows a CLC with an attention filter. The distinction between revealed preference and revealed attention is crucial. For instance, if a product is not popular in a market, it is very important for a firm to know the reason, which can either be that it is not liked by consumers or that it does not attract the attentions of consumers. Our model provides a theoretical framework to distinguish these two possibilities. Similarly, a social planner can find a proper strategy to make sure that people choose the right option in 401(K) plans and health insurance. Hence, in a welfare analysis it is important to understand the underlying model of the DM.

Since revealed preference and (in)attention are the main focus of the paper, we impose a rather weak restriction on consideration sets. Such a weak condition allows us to apply our revealed preference and (in)attention theorems to seemingly irrational choice patterns (i.e. Attraction Effect, Cyclical Choice, and Choosing Pairwisely Unchosen). Nevertheless, depending on the intended application, it is possible to analyze the CLC framework under different restrictions.

◇ In many real-world markets, products compete 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 exist many others, then it is easier to be considered when some of other alternatives 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, i.e., $x \in \ (T)$ implies $x \in \ (S)$ whenever $x \in S \subset T$. Lleras et al. (2009) extensively study consideration sets which satisfy this property. Our model and their model are independent; that is, there exists a choice data consistent with one of them but not the other. They also consider the cases where both conditions are satisfied.

◇ Lleras et al. (2009) also consider another 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 job market to consider

Fortunately, one can extend our revealed attention and inattention results (Theorem 1-2) to this special case simply by replacing the revealed preference P_R with the completely identi ed preference.

In a nutshell, without the full consideration assumption, the standard revealed preference method and hence the welfare analysis based on it can be misleading. This paper enriches the standard model of rational choice by relaxing the full attention assumption and illustrates how to deduce both DM's preference and the alternatives to which the DM pays attention or ignores from observed choices. This setup serves as a theoretical foundation for the aforementioned marketing and nance literatures on consideration sets and also a ects our understanding of the associated welfare analysis.

References

Ambrus, A. and K. Rozen, \Revealed Con icting Preferences: Rationalizing Choice with Multi-Self Models," 2008. Cowles Foundation Discussion Paper No. 1670.

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, Gozen 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.

Chambers, C. P. and T. Hayashi, \Choice and Individual Welfare," 2008. HSS California Institute of Techology WP 1286.

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 E ects," Working Papers, Brown University 2009.

Dean, Mark and Andrew Caplin, \Search, Choice, and Revealed Preference," 2008. Mimeo.

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, \Costly 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.

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.

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, 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.

Lavidge, Robert J. and Gary A. Steiner, \A model for predictive measurements of advertising e ectiveness.," *Journal of Marketing*, 1961, *25*, 59{62.

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

Lleras, J., Yusufcan Masatlioglu, Daisuke. 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, \Choice by Iterative Search," 2007. Mimeo, Department of Economics, University of Michigan.

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

Noor, Jawwad, \Temptation, Welfare and Revealed Preference," 2009. Mimeo.

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.

Rubinstein, Ariel and Yuval Salant, \Eliciting Welfare Preferences from Behavioral Datasets," 2009. mimeo.

Samuelson, Paul A., \A note on the pure theory of consumers' behavior," *Econometrica*, 1938, pp. 61{71.

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.

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.

7 Appendix

The Proofs of Theorem 1, Theorem 2 and Theorem 3

Notice that the if-parts of Theorem 1 and Theorem 2 have been already shown in the main text. The following proofs use these results.

Proof of Theorem 3

Suppose c is a CLC with an attention filter represented by ($\succ$,). Then Theorem 1(if part) implies that $\succ$ must include P so P must be acyclic. Therefore, by Lemma 1, c must satisfy WARP(LA).

Now suppose that c satisfies WARP(LA). By Lemma 1, P is acyclic so there is a preference $\succ$ that includes P. Pick any such preference arbitrarily and define

$$(S) = \{x \in S \ : \ c(S) \succ x\} \cup \{c(S)\} \tag{2}$$

Then, it is clear that $c(S)$ is the unique $\succ$-best element in (S) so all we need to show is that is an attention filter. Suppose $x \in S$ but $x \notin (S)$ (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 $(S) = (S \setminus x)$. □

Proof of Theorem 1 (The Only-If part)

Suppose xP_Ry does not hold. Then there exists a preference that includes P_R and ranks y better than x. The proof of Theorem 3 shows that c can be represented by such a preference so x is not revealed to be preferred to y. □

Proof of Theorem 2 (The Only-If parts)

(Revealed Inattention) Suppose x is not revealed to be preferred to $c(S)$. Then pick a preference that includes P_R and puts $c(S)$ above x. The proof of Theorem 3 shows that c can be represented by such a preference and an attention filter with $x \in (S)$.

(Revealed Attention) Suppose there exists no T which 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 with $x \notin (S)$. If $c(S)P_Rx$ does not hold, we have already shown that c can be represented with $x \succ c(S)$ and $x \notin (S)$ 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, $\bar{P}$, where $a\bar{P}b$ if and only if \$aP_Rb$" or \$a = c(S)$ and not $bP_Rc(S)$." That is, $\bar{P}$ 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 $\bar{P}$ is also acyclic. Given this, take any preference relation $\succ$ that includes $\bar{P}$, which includes P_R as well. We have already shown that $\tilde{\ }(S) \equiv \{z \in S : c(S) \succ z\} \cup \{c(S)\}$ is an attention filter and $(\tilde{\ }, \succ)$ represents c. Now define as follows :

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

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

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

That is, is obtained from $\tilde{}$ by removing from x 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 this true, the condition of the statement is satis ed for $T = S$. Hence, $(S') \subset \tilde{}(S')$ always includes $c(S')$. Furthermore the proof of Theorem 3 shows that $(\tilde{}, \succ)$ represents c. Therefore, $(\ , \succ)$ also represents c so we only need to show that is an attention lter.

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

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

Case I: $y = x$

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

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

Since $y \notin (T)$ is equivalent to $y \notin \tilde{}(T)$, we have $\tilde{}(T) = \tilde{}(T \setminus y)$. Therefore, $c(T \setminus y) = c(T) = c(S)$. By construction of and $\tilde{}$, 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, $(T) = \tilde{}(T) \setminus x = \tilde{}(T \setminus y) \setminus x = (T \setminus y)$.

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 $(T) = \tilde{}(T) = \tilde{}(T \setminus y) = (T \setminus y)$. □