二松学舎大学国際政治経済学部

Discussion Paper Series

Generalized Revealed Attention

Yukinori Iwata

November 2013

Discussion Paper (Econ) No.4

FACULTY OF INTERNATIONAL POLITICS AND ECONOMICS

NISHOGAKUSHA UNIVERSITY

Generalized Revealed Attention

Yukinori Iwata*

November 22, 2013

Abstract

This paper proposes a natural extension of the limited attention model introduced by Masatlioglu et al. (American Economic Review 102(5): 2183-2205, 2012). We assume that attributes or frames that affect decision making are attached to each alternative. A decision maker does not always pay attention to all feasible alternatives, and moreover, her attention varies according to the salience of attributes for alternatives (Generalized Limited Attention). We provide characterizations of revealed preference, revealed attention/inattention, and a choice behavior that maximizes a single preference under generalized limited attention. Our model explains more choice behaviors than the limited attention model by Masatlioglu et al. (2012) and improves the power of inferring revealed preference and revealed attention/inattention.

JEL classifications: D01, D11

Keywords: revealed preference, revealed attention, consideration set, generalized limited attention, frames

1 Introduction

Many methods that explain choice behaviors that seem to be irrational have been recently proposed. One of the most influential methods is a two-stage choice procedure. That is, a decision maker (DM) makes a choice from a subset of feasible alternatives after intentionally or unconsciously eliminating some alternatives from the set of all feasible alternatives. See Cherepanov et al. (2013), Lleras et al. (2011), and Manzini and Mariotti (2007) for typical examples.

In a recent paper, Masatlioglu, Nakajima, and Ozbay (2012), and henceforth MNO, propose a revealed preference model with limited attention, where a DM chooses her most preferred alternative from the set of alternatives to which she pays attention, not from the set of all feasible alternatives. The set of alternatives that she considers, which is called the *consideration set*, is specified by her *consideration set mapping* that assigns a nonempty subset of all alternatives to

*Faculty of International Politics and Economics, Nishogakusha University, 6-16 Sanbancho, Chiyoda-ku, Tokyo 102-8336, Japan; E-mail: y-iwata@nishogakusha-u.ac.jp

each subset of all alternatives. MNO assume that a consideration set mapping is an *attention filter* that requires that a consideration set does not change when an alternative to which a DM does not pay attention becomes unavailable. A DM follows a *choice with limited attention* (CLA) if there exist a single preference over alternatives and an attention filter such that the chosen alternative is the best element in the consideration set according to the preference.

MNO develop a method of identifying both DM's revealed preference and revealed attention/inattention under the CLA assumption. In the MNO model, a DM's attention is elicited as follows: She chooses x from S, while a choice reversal is caused by removing y from S. Then, it is possible to conclude that y is revealed to attract attention at S because this would be impossible if she did not pay attention to y according to the CLA assumption (Revealed Attention). In this situation, it is also possible to conclude that she prefers x to y because x is chosen while y attracts her attention (Revealed Preference). Furthermore, MNO characterize a choice function that satisfies a CLA by the weakening of the Weak Axiom of Revealed Preference (WARP).

In the limited attention model proposed by MNO, a consideration set mapping is defined only on a collection of subsets of all alternatives. This setting implicitly implies that a DM's consideration set remains the same while the set of feasible alternatives remains the same. This paper introduces any other factors that influence a DM's consideration set even if the set of feasible alternatives remains the same. We assume that some *attributes* that affect a DM's attention are attached to each alternative. Each attribute is characterized by *salience* and is represented by a vector consisting of real numbers. The profile of attributes is a matrix consisting of vectors, each of which represents an attribute for alternatives. We assume that a positively (resp. negatively) salient attribute for an alternative attracts (resp. diverts) a DM's attention to the alternative. For instance, in the literature on marketing, a consumer changes her attention according to the salience of attributes such as advertisements, people's choices and evaluation, positive or negative campaigns, and good or bad news and rumors.

We assume that a DM's consideration set mapping associates her consideration set to each pair of the set of feasible alternatives and the profile of attributes. It satisfies a *generalized attention filter* if her consideration set remains the same when an alternative to which she does not pay attention becomes unavailable or the salience of an attribute for the alternative decreases. A DM maximizes her preference over alternatives within her generalized attention filter. By a similar way to the MNO limited attention model, we can elicit both DM's revealed preference and revealed attention/inattention whenever a choice reversal is observed as a consequence of removing an unchosen alternative or decreasing the salience of an attribute for the alternative. We also show that a choice behavior in our model is fully characterized by a generalization of the weakening of WARP proposed by MNO.

The MNO model of a choice with limited attention can explain some anomalies such as attraction effects, cyclical choices, and choosing pairwisely unchosen, while they also show that revealed preference inferred in the MNO model could be incomplete. We show that our model can elicit further revealed preference that is not inferred in the MNO model. As MNO pointed out, a policymaker may be *forced* to make a welfare judgment even when a DM's revealed preference is extremely incomplete. In addition, the policymaker may *nudge* people to help them better choices as discussed in the literature on libertarian paternalism (See Thaler and Sunstein 2008). Our model will be able to improve the policymaker's welfare analysis by inferring more revealed preferences than the MNO model and advise the policymaker which attribute for alternatives attracts or diverts people's attention.

The remaining sections are organized as follows. Section 2 discusses reasonable examples in this study and the related literatures. Section 3 introduces the basic model. Section 4 provides the main results in this study. Section 5 concludes the paper. The proofs of our results are presented in the appendix.

2 Examples and the related literatures

In this section, we provide some examples, each of which captures a property of our model. The key point of our model is that the salience of attributes attached to each alternative varies, while a DM's attention also changes. Some of the examples are analyzed in a probabilistic model, but we describe them in a non-probabilistic model. Moreover, these choice behaviors are often explained by changing a DM's underlying preference, but we will explain them by changing her attention, while preserving her preference. Note that these examples contain a DM's behavior that the MNO limited attention model cannot explain because her attention would change even if the set of feasible alternatives does not change at all.

Additional information about alternatives may influence a DM's choice, which is well-known as a framing effect.[1] For example, the way of expressing a question or a statement, e.g., whether it includes a positive phrase or negative one, influences the DM's attention and choice (See Tversky and Kahneman 1981).

Example 1 (Framing)**.** A surgeon explains a surgery to a patient in one of two ways: "Of 100 people having surgery 90 live through the post-operative period." Or "Of 100 people having surgery 10 die during surgery or the post-operative period." Then, the patient will undergo the surgery in the former, but she will not in the latter.

In the literature on marketing, any information about products (e.g., advertisements, characteristics, and popularity etc.) may influence a consumer's attention and choice. For example, a consumer considers the N-most advertised product in the market. Such information varies, while the consumer's consideration also changes.

[1] Salant and Rubinstein (2008) explain a frame as observable information, other than the set of feasible alternatives, which is irrelevant in the rational assessment of the alternatives but nonetheless affects behavior.

Example 2 (Markting)**.** A person looks for an apartment appearing on the web. As a result of using a search engine, she will choose the best apartment among ones that appear on the first page of the searching results according to her preference. Then, her choice will change according to which apartments appear on the first page.

A DM's behavior may be influenced by what other people are doing, which is well-known as herd behavior (See Banerjee 1992). A DM will make a choice according to information available from observing other people's behavior rather than information she has.

Example 3 (Following the Herd)**.** Restaurants A and B are next to each other. Many people are waiting in front of A, but nobody is waiting in front of B. Then, a person who arrives at the restaurants will go to A even if she would go to B when nobody is waiting in front of both A and B.

A DM's behavior may be influenced by what other people think and say. What we should do or should not do depends on other people's views and opinions about it. In this situation, a DM will change her choice according to what other people expect her to do. We often follow such a norm intentionally or unconsciously.[2]

Example 4 (Following the Norm)**.** A person who is smoking a cigarette in a hospital waiting room puts out it because another person next to her asks her not to smoke. Furthermore, in the long run, the rate of women's smoking in the USA depends on people's views about women's smoking throughout the 20th century (See Akerlof and Kranton 2010).

We now make a few comments on the related literatures. Salant and Rubinstein (2008) study a choice model with frames. In their model, a set of frames is abstract, and there exists an ordering over a feasible set with respect to each frame and an ordering changes while a frame changes. On the other hand, in our model, a preference remains the same but a consideration set changes while the salience of attributes for alternatives changes. Salant and Rubinstein (2008) analyze relationship between a choice with frames and the standard rational choice theory. They also discuss the limited attention model, but they assume that the number of alternatives to which a DM pays attention is observable while a consideration set in our model is not observable.

In the literature on social choice theory, Iwata (2011, 2012) proposes a similar two-stage social choice procedure in which each attribute is interpreted as an agenda setter in society and the salience of an attribute as an agenda setter's opinion. In this social choice framework, the set of feasible alternatives is endogenously determined by aggregating agenda setters' opinions, and a consideration set mapping is called a qualification function that specifies the set of

[2]For example, Akerlof and Kranton (2010) define norms as "the social rules regarding how people should behave in different situations (p.4)" and state "these rules are sometimes explicit, sometimes implicit, largely internalized, and often deeply held. (p.4)"

qualified alternatives for collective decision making. The difference of qualification functions from consideration set mappings is that a qualification function is observable and each agenda setter's opinion is ternary, that is, it consists of 1, 0, or -1.

In the literature on consumer research, decision making procedures in which a consumer's preference plays no role are proposed. See Bettman et al. (1998) for a comprehensive discussion on the literature. For example, Elimination-by-aspects (EBA) is a decision making procedure as follows: some alternatives that do not meet a threshold for the most important attribute are eliminated first. If more than one alternative remain, then this elimination process is applied for the second most important attribute, with the process repeated until a single alternative remains. The decision making procedures like EBA determine a consumer's final choice and therefore her preference plays no role in making a choice. On the other hand, we interpret these decision making procedures as the first stage of decision making in which some alternatives are eliminated. If more than one alternative remain at this stage, then a DM makes a choice from the remaining alternatives according to her preference.

3 The model

Let $X = \{x_1, x_2, \ldots, x_m\}$ be a finite set of alternatives. Let $\mathscr{X}$ denote the set of all nonempty subsets of X, whose interpretation is the collection of all the objectively observable alternatives a DM could be facing. Let $S, T \in \mathscr{X}$ denote a nonempty subset of X. They are interpreted as a set of feasible alternatives the DM is actually facing.

We assume that some *attributes* that influence DM's attention to alternatives may or may not be attached to every alternative. Let $\mathcal{A} = \{1, 2, \ldots, n\}$ be a finite set of attributes for alternatives. Each attribute for alternatives is characterized by *salience*. Let A_j denote the salience of n attributes for alternative x_j, represented by an $n \times 1$ column vector consisting of real numbers. Let a_{ij} be the ith component of A_j with $a_{ij} \in I_i$, where I_i is a finite set of real numbers or a closed interval with $[\underline{\alpha}_i, \overline{\alpha}_i]$ and represents a range of the salience of attribute i. Let us now explain the interpretation of a_{ij}. We assume $a_{ij} \neq 0$ if attribute i is attached to alternative x_j and $a_{ij} = 0$ otherwise. We assume that every a_{ij} is observable. The salience of each attribute attracts or diverts a DM's attention to alternatives. We assume that if $a_{ij} > 0$ (resp. $a_{ij} < 0$), attribute i attracts (resp. diverts) a DM's attention to alternative x_j. If $a_{ij} > a'_{ij} > 0$ (resp. $0 > a'_{ij} > a_{ij}$), then attribute i for alternative x_j is more positively (resp. negatively) salient at A_j than at A'_j.[3]

An *attribute profile* A is an $n \times m$ matrix consisting of m column vectors $A_1, \ldots, A_m$. Let $\mathscr{A}$ be the set of all attribute profiles. (S, A) is a pair of a subset S of X and an attribute profile A and is called a *profile*. Let $(S \backslash x_j, A)$

[3] We should deal with attributes for alternatives carefully. For example, a waiting line in front of a restaurant is positively salient for a person who has much time, but negatively salient for another person who has little time.

be the profile obtained by removing x_j from S at (S, A).[4] Let $(S, A \backslash a'_{ij})$ be the profile obtained by replacing a_{ij} of A with a'_{ij} at (S, A). A *consideration set mapping*, $f : \mathscr{X} \times \mathscr{A} \to \mathscr{X}$, is a function that assigns a nonempty subset of S to each profile (S, A). A consideration set mapping specifies which alternatives a DM pays attention to for any profile (S, A). For all (S, A), $f(S, A)$ is called the *consideration set* at (S, A). Since a consideration set mapping in the MNO model is a function defined on $\mathscr{X}$, it is possible to interpret that their model is the special case of our model, where the unique A exists in $\mathscr{A}$.

We now propose a property of a consideration set mapping. A *generalized attention filter* requires that if the DM does not pay attention to alternative x_j at (S, A), then removing x_j from S or decreasing the salience of an attribute i for x_j does not change her consideration set.

Definition 1. *A consideration set mapping f is a generalized attention filter if for any (S, A) and any a'_{ij} with $a_{ij} \geq a'_{ij}$, $f(S \backslash x_j, A) = f(S, A) = f(S, A \backslash a'_{ij})$ whenever $x_j \notin f(S, A)$.*

If A uniquely exists in $\mathscr{A}$, then a generalized attention filter is equivalent to an attention filter proposed by MNO. An *extended choice function*, $c : \mathscr{X} \times \mathscr{A} \to X$, assigns an alternative in S to each profile (S, A). Let $c(S, A)$ be an alternative chosen from S at (S, A). Let $\succ$ be a DM's *preference* over X, that is, a complete, transitive, and antisymetric binary relation over X. We assume that a DM obeys a choice with generalized limited attention but her preference and consideration set mapping are not observable.

Definition 2. *An extended choice function c is a choice with generalized limited attention (CGLA) if there exists a preference $\succ$ and a generalized attention filter f such that $c(S, A)$ is the $\succ$-best alternative in $f(S, A)$, that is, $c(S, A) = \max_{\succ} f(S, A)$.*

If we additionally assume that a consideration set always contains just one alternative, that is, $\# f(S, A) = 1$ for all (S, A), then we have $c(S, A) = f(S, A)$.[5] This observation implies that a consideration set mapping determines an alternative chosen from a feasible set and a DM's preference plays no role, which is a similar decision-making procedure proposed in the literature on consumer research. However, this paper is not always the case because a consideration set is not singleton.

4 Characterization results

This section provides main results in this study. As in the MNO model, we first illustrate the way of inferring (i) the DM's preference and (ii) the DM's attention/inattention from her choice data that satisfies a CGLA. MNO conclude that a DM prefers x to y when choosing x from S is changed as a result of removing

[4] That is, $S \backslash x_j$ is equivalent to $S \backslash \{x_j\}$.

[5] $\#$ represents the number of alternatives in a set.

y from S. In our model, a choice reversal causes not only when y is removed but also when the salience of an attribute for y is decreased. This observation is useful to understand revealed preference in our model. For any distinct x_j and x_k, define $x_j P x_k$ if there exists (S, A) such that $c(S, A) = x_j \neq c(S \backslash x_k, A)$ or $c(S, A) = x_j \neq c(S, A \backslash a'_{ik})$ with $x_k \in S$ and $a_{ik} \geq a'_{ik}$. Thus, when a choice reversal happens by removing an alternative or decreasing the salience of an attribute for the alternative, the initially chosen alternative is preferred to the operated alternative. Let P_R be the transitive closure of P.

The next theorem shows that P_R is the revealed preference in our model.

Theorem 1 (Revealed Preference)**.** *Suppose c is a CGLA. Then, x_j is revealed to be preferred to x_k if and only if $x_j P_R x_k$.*

As shown in MNO, we conclude that x_j is revealed not to attract attention whenever x_j is revealed to be preferred to x_k but x_k is chosen from S at (S, A) with $x_j \in S$. Furthermore, MNO infer that x is revealed to attract attention at S when x is chosen from S or removing x from S causes a choice reversal. They also find the indirect way of inferring that x is revealed to attract attention at S. Suppose that removing x from T, not S causes a choice reversal. Then, a DM pays attention to x at T. Suppose that all alternatives that belong to S but not to T is revealed to be preferred to the chosen alternative from S and all alternatives that belong to T but not to S is revealed to be preferred to the chosen alternative from T. Then, a DM does not attract attention to those alternatives at S and T, respectively. Therefore, removing them from S or T cannot change her consideration set, which implies that x is paid attention to at S.

We apply this idea to identify revealed attention/inattention in our model. When x is revealed to attract attention at (T, B), we need the following equation to infer that x is revealed to attract attention at (S, A) indirectly:

$$f(S, A) = f(S \cap T, C) = f(T, B).$$

That is, removing all alternatives that belong to either S or T but not both reaches $S \cap T$ by a similar way to MNO. But, how to reach from both A and B to C? Decreasing the salience of an attribute for an alternative does not change a DM's consideration set only when she does not pay attention to the alternative. As we mentioned, an alternative that is revealed not to attract attention at (S, A) and (T, B) when it is revealed to be preferred to the chosen alternative at (S, A) and (T, B). Otherwise, A_j must be equivalent to B_j because we cannot decrease the salience of any attribute for alternative x_j.[6]

The following theorem captures the idea above and characterizes revealed attention and inattention in our model.

Theorem 2 (Revealed (In)Attention)**.** *Suppose c is a CGLA.*

1. *x_j is revealed not to attract attention at (S, A) if and only if $x_j P_R c(S, A)$,*

[6]For all $x_j \notin S \cup T$, we can constract C_j by setting $c_{ij} = \min(a_{ij}, b_{ij})$.

2. *x_j is revealed to attract attention at (S, A) if and only if there exists (T, B) (possibly equal to (S, A)) such that:*

   (a) *$c(T, B) \neq c(T \backslash x_j, B)$ or $c(T, B) \neq c(T, B \backslash b'_{ij})$ with $x_j \in T$ and $b_{ij} \geq b'_{ij}$,*

   (b) *$x_k P_R c(S, A)$ for all $x_k \in S \backslash T$,*
   $x_l P_R c(T, B)$ for all $x_l \in T \backslash S$,

   (c) *for all $x_k \in S \cup T$,*
   $A_k \geq B_k$ if $x_k P_R c(S, A)$ and $x_k P_R c(T, B)$,
   $A_k = B_k$ otherwise.

MNO propose a choice behavioral hypothesis, that is, a weakening of the Weak Axiom of Revealed Preference (WARP) to characterize the MNO model of a choice with limited attention. WARP with Limited Attention (WARP(LA)) requires that there exists the "best" alternative x^* in any set S in the sense that x^* must be chosen from any set T including x^* whenever the choice from T is still in S and removing x^* from T changes the DM's choice. In our model, the DM's choice changes not only when x^* is removing from T but also when the salience of an attribute for x^* is decreased at (T, B). Thus, we generalize WARP(LA) as follows:

WARP with Generalized Limited Attention (WARP(GLA)): For any (S, A), there exists $x_j \in S$ such that, for any (T, B) with $x_j \in T$, if c$(T, B) \in S$, and $c(T, B) \neq c(T \backslash x_j, B)$ or $c(T, B) \neq c(T, B \backslash b'_{ij})$ with $b_{ij} \geq b'_{ij}$, then $c(T, B) = x_j$.

If the unique A is in $\mathscr{A}$, then WARP(GLA) is equivalent to WARP(LA). WARP(GLA) characterizes the class of extended choice functions that satisfy a generalized attention filters.

Theorem 3 (Characterization)**.** *c satisfies WARP(GLA) if and only if c is a CGLA.*

As in MNO, Theorem 3 shows that a CGLA is captured by a single behavioral postulate. Thus, it is possible to test our model by using a method found in the literature on the revealed preference theory and to elicit the DM's preferences and generalized attention filter by following Theorems 1 and 2 from observed choice data.

We now highlight the difference between the MNO model and our model. MNO suggest that their limited attention model can explain some anomalies, including attraction effects, cyclical choice, and choosing pairwisely unchosen. At the same time, they show that revealed preferences are incomplete in those examples. We will show that our model is able to explain further revealed preference that MNO cannot infer in their examples. In what follows, we assume $X = \{x_1, x_2, x_3\}$.

The attraction effect is a phenomenon where adding an irrelevant alternative to a feasible set affects the choice.

Example 5 (Attraction Effect). Suppose that there exists one attribute for alternatives. Let $I_1 = \{-1, 0, 1\}$. Suppose that attribute 1 determines which alternative is a decoy of which alternatives. We assume that for all (S, A) with $x_j \in S$, if $a_{1j} = -1$, then $x_j \in f(S, A)$ if and only if there exists $x_k \in S$, $k \neq j$, with $a_{1k} = 1$, which interprets that x_k is a decoy of x_j, or $a_{1k} = -1$ for all $x_k \in S$. We assume that she considers everything in any other cases. It is clear that the DM's consideration set mapping is a generalized attention filter.

Let $A = (0, -1, 1)$. Then, a typical attraction effect choice pattern is observed as follows:

$$c(X, A) = x_2,\ c(X \backslash x_1, A) = x_2,\ c(X \backslash x_2, A) = x_1,\ c(X \backslash x_3, A) = x_1.$$

In this situation, x_3 is a decoy of x_2 and therefore an attraction effect is caused by removing x_3 from X, that is, $c(X, A) = x_2$ but $c(X \backslash x_3, A) = x_1$. Suppose that the DM's preference is $x_2 \succ x_1 \succ x_3$. As in MNO, it is possible to show that x_2 is revealed to be preferred to x_3, but we cannot determine the ranking of x_1 in this situation.

We now show that our model is able to determine the ranking of x_1. Let $B = (1, -1, 0)$. Then, x_1 is a decoy of x_2 and the choice pattern changes as follows:

$$c(X, B) = x_2,\ c(X \backslash x_1, B) = x_3,\ c(X \backslash x_2, B) = x_1,\ c(X \backslash x_3, B) = x_2.$$

By the same reasoning above, we can infer that x_2 is revealed to be preferred to x_1.

Furthermore, let $C = (-1, 0, 1)$ and $D = (-1, 0, 0)$. Then, x_3 is a decoy of x_1 at C but not at D. Again, the choice pattern changes as follows:

$$c(X, C) = x_2,\ c(X \backslash x_1, C) = x_2,\ c(X \backslash x_2, C) = x_1,\ c(X \backslash x_3, C) = x_2;$$
$$c(X, D) = x_2,\ c(X \backslash x_1, D) = x_2,\ c(X \backslash x_2, D) = x_3,\ c(X \backslash x_3, D) = x_2.$$

Thus, we can infer x_1 is revealed to be preferred to x_3 because we have $c(X \backslash x_2, C) = x_1$ and $c(X \backslash x_2, D) = x_3$.

The cyclical choice pattern is related to a DM's intransitive preference.

Example 6 (Cyclical Choice). Suppose that there exists one attribute for alternatives. Let $I_1 = \{-1, 0, 1\}$. We assume that a DM considers alternatives in S in the following way. She first pays attention to alternatives for which the salience of attribute 1 is the largest in alternatives in S, that is $\max_{j: x_j \in S} a_{1j}$. Then, she successively considers alternatives for which the salience of attribute 1 is "close" to that of alternatives that have already considered. We define that x_j is close to x_k if and only if $|a_{1j} - a_{ik}| \leq 1$. Therefore, $x_l \in f(S, A)$ if and only if there exists a sequence $x_j, x_k, x_l \in S$ such that $x_j \in f(S, A)$, $|a_{1j} - a_{1k}| \leq 1$, and $|a_{1k} - a_{1l}| \leq 1$. It is easily possible to check that her consideration set mapping is a generalized attention filter.

Suppose that her true preference is $x_1 \succ x_2 \succ x_3$ and let $A = (-1, 0, 1)$. Then, a typical cyclical choice pattern is observed as follows:

$$c(X, A) = x_1,\ c(X \backslash x_1, A) = x_2,\ c(X \backslash x_2, A) = x_3,\ c(X \backslash x_3, A) = x_1.$$

As in MNO, we conclude, from this choice data, that x_2 attracts her attention at (X, A) and x_1 is revealed to be preferred to x_2 because $x_1 = c(X, A) \neq c(X \backslash x_2, A)$. However, as shown in MNO, it is impossible identify the ranking of x_3 in this situation.

We now identify the ranking of x_3. Let $B = (-1, 1, 1)$ and $C = (-1, 1, 0)$. The choice pattern changes as follows:

$$c(X, B) = x_2,\ c(X \backslash x_1, B) = x_2,\ c(X \backslash x_2, B) = x_3,\ c(X \backslash x_3, B) = x_2;$$
$$c(X, C) = x_1,\ c(X \backslash x_1, C) = x_2,\ c(X \backslash x_2, C) = x_1,\ c(X \backslash x_3, C) = x_2.$$

By $x_2 = c(X, B) \neq c(X, C)$, we infer that the DM prefers x_2 over x_3. By the transitivity of her revealed preference, x_1 is revealed to be preferred to x_3.

In the example of choosing pairwisely unchosen, a DM never choose an alternative in any binary choice.

Example 7 (Choosing Pairwisely Unchosen)**.** Suppose that there exist two attributes for alternatives. Let $I_1 = [-1, 1]$ and $I_2 = \{-1, 0, 1\}$. In this example, attribute 1 represents how easy to find alternatives. If the salience of attribute 1 for an alternative is positive, then the DM can easily find the alternative. However, if the salience of attibute 1 for an alternative is not positive even if it is feasible, then the DM cannot find it unless she makes an extensive search. She makes an extensive search only when there exists no alternative that "dominates" all easily found alternatives or the salience of attribute 1 for every feasible alternative is not positive. We define that alternative x_j dominates alternative x_k if and only if (i) the DM prefers x_j to x_k and (ii) the salience of attribute 2 for x_j is at least as large as that of x_k.

As a consequence of an extensive search, she finds alternatives in a decreasing order of the salience of attribute 1 for alternatives in S unless the sum of the nonpositive salience of attribute 1 exceeds -1. In any other cases, she finds alternatives that have the largest nonpositive salience of attribute 1. For instance, let A be such that $a_{11} = a_{12} = -1/4$ and $a_{13} = -2/3$. Then, the DM must make an extensive search to find alternatives. In this case, we have $f(S, A) = f(S \backslash x_3, A) = \{x_1, x_2\}$, $f(S \backslash x_1, A) = \{x_2, x_3\}$, and $f(S \backslash x_2, A) = \{x_1, x_3\}$ because we have $a_{11} + a_{12} \geq -1$ but $a_{11} + a_{12} + a_{13} < -1$, $a_{11} + a_{13} \geq -1$, and $a_{12} + a_{13} \geq -1$.

Suppose that her true preference is $x_3 \succ x_1 \succ x_2$ and let $A = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \end{pmatrix}$. Then, as in MNO, a typical choice pattern of choosing pairwisely unchosen appears:

$$c(X, A) = x_3,\ c(X \backslash x_1, A) = x_2,\ c(X \backslash x_2, A) = x_1,\ c(X \backslash x_3, A) = x_1.$$

In this situation, x_3 is hard to find because $a_{13} \leq 0$, and x_1 does not dominate x_2 and is not dominated by x_2 because $x_1 \succ x_2$ but $a_{21} < a_{22}$. Then, the DM makes an extensive search to find x_3 only when x_1 and x_2 are present. Since removing x_1 or x_2 from X changes her choice, we conclude that x_3 is revealed to be preferred to x_1 and x_2 but, as in MNO, it is impossible to determine the ranking between x_1 and x_2 in this situation.

Let $B = \begin{pmatrix} 0 & -1/2 & -1 \\ -1 & 1 & 0 \end{pmatrix}$. Then, the choice pattern changes as follows:

$$c(X, B) = x_1,\ c(X \backslash x_1, B) = x_2,\ c(X \backslash x_2, B) = x_3,\ c(X \backslash x_3, B) = x_1.$$

In this situation, the DM always makes an extensive search to find alternatives. She overlooks x_3 when x_1 and x_2 are present because we have $a_{11} \geq a_{12} \geq a_{13}$ and $a_{11} + a_{12} \geq -1$ but $a_{11} + a_{12} + a_{13} < -1$. Since $x_1 = c(X, B) \neq c(X \backslash x_2, B)$, we correctly infer that x_1 is revealed to be preferred to x_2.

5 Concluding remarks

This paper proposed a natural extension of the limited attention model introduced by MNO. A DM does not always pay attention to all feasible alternatives, and moreover, her attention varies according to the salience of attributes for the alternatives (GLA). We showed the way of inferring both the preference and consideration sets of a DM who follows a CGLA. We provided a characterization of a choice behavior that obeys CGLA. We showed that our model explains more choice behaviors than the limited attention model by MNO and improves the power of inferring revealed preference and revealed attention/inattention.

As MNO emphasize, it is crucial to distinguish between revealed preference and revealed (in)attention. For instance, it is important for a firm to know the reason why a product is not popular, that is, whether consumers do not prefer the product or it does not attract consumers' attention. In addition to MNO's observation, we further argue that the salience of advertisements for the product may influence consumers' attention. Thus, it is also important to know which advertisement attracts consumers' attention.

Acknowledgment The author is grateful to Naoki Yoshihara for his valuable suggestions.

Appendix

Lemma 1. *P is acyclic if and only if c satisfies WARP(GLA).*

Proof. (The if-part): Suppose P has a cycle: $x^1 P x^2 P \cdots P x^t P x^1$. Without loss of generality, we can rearrange the order of alternatives such as $x_1 P x_2 P \cdots P x_t P x_1$. Then for all $j \in \{1, \ldots, t-1\}$, there exists (T^j, B^j) such that $x_j = c(T^j, B^j) \neq c(T^j \backslash x_{j+1}, B^j)$ or $x_j = c(T^j, B^j) \neq c(T^j, B^j \backslash b'_{i(j+1)})$ with $x_{j+1} \in T^j$ and

$b^j_{i(j+1)} \geq b'_{i(j+1)}$, and $x_t = c(T^t, B^t) \neq c(T^t \backslash x_1, B^t)$ or $x_t = c(T^t, B^t) \neq c(T^t, B^t \backslash b'_{i1})$ with $x_1 \in T^t$ and $b^t_{i1} \geq b'_{i1}$. Consider the profile (S, A) with $\{x_1, \ldots, x_t\} \equiv S$. Then, for all $x_j \in S$, there exists (T, B) such that $c(T, B) \in S$, and $c(T, B) \neq c(T \backslash x_j, B)$ or $c(T, B) \neq c(T, B \backslash b'_{ik})$ with $x_k \in T$ and $b_{ik} \geq b'_{ik}$ but $x_j \neq c(T, B)$, which implies that c violates WARP(GLA).

(The only-if part): Suppose P is acyclic. Then for all (S, A), there exists at least one alternative $x_j \in S$ such that there is no $x_k \in S$ with $x_k P x_j$, which implies that there is no $x_k \in S$ such that $x_k = c(T, B) \neq c(T \backslash x_j, B)$ or $x_k = c(T, B) \neq c(T, B \backslash b'_{ij})$ with $x_j \in T$ and $b_{ij} \geq b'_{ij}$. Therefore, whenever $c(T, B) \in S$ and $c(T, B) \neq c(T \backslash x_j, B)$ or $c(T, B) \neq c(T, B \backslash b'_{ij})$ with $x_j \in T$ and $b_{ij} \geq b'_{ij}$, we have $x_j = c(T, B)$, which implies that c satisfies WARP(GLA). □

Proof of Theorem 3. Suppose c is a CGLA represented by $(\succ, f)$. Then the if-part of Theorem 1 implies that $\succ$ includes P. Hence, P must be acyclic. It follows from Lemma 1 that c satisfies WARP(GLA).

We now suppose c satisfies WARP(GLA). Since P is acyclic by Lemma 1, we can find a preference $\succ$ that includes P. Consider any preference $\succ$ that includes P and define the following consideration set mapping:

$$f(S, A) = \{x_j \in S : c(S, A) \succ x_j\} \cup \{c(S, A)\}.$$

Then, it is clear that $c(S, A)$ is the unique $\succ$-best element in $f(S, A)$. We now show that f is a generalized attention filter. We shall show that $f(S, A) = f(S \backslash x_j, A)$ and $f(S, A) = f(S, A \backslash a'_{ij})$ with $a_{ij} \geq a'_{ij}$. Suppose $x_j \in S$ but $x_j \notin f(S, A)$. Then, we have $x_j \neq c(S, A)$. By the construction of f, we have $x_j \succ c(S, A)$, and therefore $c(S, A) P x_j$ does not hold, which implies that we have $c(S, A) = c(S \backslash x_j, A)$ and $c(S, A) = c(S, A \backslash a'_{ij})$ with $x_j \in S$ and $a_{ij} \geq a'_{ij}$. By the construction of f, we have $f(S, A) = f(S \backslash x_j, A)$ and $f(S, A) = f(S, A \backslash a'_{ij})$ with $a_{ij} \geq a'_{ij}$. □

Proof of Theorem 1. (The only-if part): Suppose $x_j P_R x_k$ does not hold. Then, there exists a preference $\succ$ that includes P_R and ranks x_k higher than x_j. The proof of Theorem 3 shows that c can be represented by the preference $\succ$. Therefore, x_j is not revealed to be preferred to x_k. □

Proof of Theorem 2. (The only-if part): (Revealed Inattention) Suppose x_j is not revealed to be preferred to $c(S, A)$. Then, consider a preference $\succ$ that includes P_R and ranks $c(S, A)$ higher than x_i. The proof of Theorem 3 shows that c can be represented by the preference $\succ$ and a generalized attention filter f with $x_j \in f(S, A)$.

(Revealed Attention) Suppose there exists no (T, B) that satisfies the condition. We shall prove that if c is a CGLA, then it can be represented by some generalized attention filter f with $x_j \notin f(S, A)$. If $c(S, A) P_R x_j$ does not hold, we have already shown that c can be represented by $(\succ, f)$ with $x_j \succ c(S, A)$ and $x_j \notin f(S, A)$. Thus, x_j is not revealed to attract attention at (S, A), and therefore, we consider the case where $c(S, A) P_R x_j$.

Now construct a binary relation $\tilde{P}$ as follows: $a\tilde{P}b$ if and only if "aP_Rb" or "$a = c(S,A)$ and not $bP_Rc(S,A)$." That is, $\tilde{P}$ ranks $c(S,A)$ as high as possible unless it contradicts P_R. Since P_R is acyclic and c is represented by a generalized attention filter, it is possible to show that $\tilde{P}$ is also acyclic. Given this observation, consider any preference $\succ$ that includes $\tilde{P}$ as well as P_R. We have already shown that $\tilde{f}(S,A) \equiv \{x_j \in S : c(S,A) \succ x_j\} \cup \{c(S,A)\}$ is a generalized attention filter and c is represented by $(\succ, \tilde{f})$. We now define f as follows:

$$f(S',A') = \begin{cases} \tilde{f}(S',A') & \text{for } (S',A') \notin \Omega \\ \tilde{f}(S',A')\backslash x_j & \text{for } (S',A') \in \Omega \end{cases},$$

where Ω is a collection of profiles such that

$$\Omega = \left\{ (S',A') \in \mathscr{X} \times \mathscr{A} : \begin{array}{c} c(S',A') = c(S,A), \\ x_kP_Rc(S,A) \text{ for all } x_k \in (S\backslash S') \cup (S'\backslash S), \text{ and} \\ \text{for all } x_k \in S \cup S', \\ A_k \geq A'_k \text{ if } x_kP_Rc(S,A), \\ A_k = A'_k \text{ otherwise} \end{array} \right\}.$$

Thus, f is obtained from $\tilde{f}$ by removing x_j from S' of any profile (S',A') where $c(S',A') = c(S,A)$, any alternative that belongs to S and S' but not to both is revealed to be preferred to $c(S,A)$, and for any alternative in S or S', $A_j \geq A'_j$ if it is revealed to be preferred to $c(S,A)$ and $A_j = A'_j$ otherwise. Note that $x_j \neq c(S,A)$ by $c(S,A)P_Rx_j$. Therefore, $f(S',A') \subseteq \tilde{f}(S',A')$ always contains $c(S',A')$. Furthermore, since $(\succ, \tilde{f})$ represents c, $(\succ, f)$ also represents c. Thus, we only need to show that f is a generalized attention filter.

Note that $\tilde{f}$ is a generalized attention filter and $c(S',A') = c(S'',A'')$ whenever $\tilde{f}(S',A') = \tilde{f}(S'',A'')$ because $(\succ, \tilde{f})$ represents c.

Suppose $x_k \notin f(T,B)$. We shall prove $f(T,B) = f(T\backslash x_k, B)$ and $f(T,B) = f(T, B\backslash b'_{ik})$ with $b_{ik} \geq b'_{ik}$. We first show that $f(T,B) = f(T\backslash x_k, B)$. Note that since B is fixed, condition (c) of the statement is always satisfied. Therefore, it is possible to show that f is a generalized attention filter by a similar argument to the proof of Theorem 2 in MNO.

We now show that $f(T,B) = f(T, B\backslash b'_{ik})$ with $b_{ik} \geq b'_{ik}$. We distinguish three possible cases: (i) $x_k = x_j$, (ii) $(T,B) \in \Omega$ and $x_k \neq x_j$, and (iii) $(T,B) \notin \Omega$ and $x_k \neq x_j$.

Case (i): If $(T,B) \notin \Omega$, then we have $f(T,B) = \tilde{f}(T,B) = \tilde{f}(T, B\backslash b'_{ik}) = f(T, B\backslash b'_{ik})$. If $(T,B) \in \Omega$, then $c(T,B) = c(T, B\backslash b'_{ik})$ must hold (otherwise, the condition of the statement is satisfied). It follows from construction of $\tilde{f}$ and f that $f(T,B) = \tilde{f}(T,B)\backslash x_j = \tilde{f}(T, B\backslash b'_{ik}) = f(T, B\backslash b'_{ik})$.

Case (ii): Since $x_k \notin f(T,B)$ is equivalent to $x_k \notin \tilde{f}(T,B)$, we have $\tilde{f}(T,B) = \tilde{f}(T, B\backslash b'_{ik})$ with $b_{ik} \geq b'_{ik}$. Therefore, we have $c(T, B\backslash b'_{ik}) = c(T,B) = c(S,A)$. By construction of f and $\tilde{f}$, $x_k \succ c(S,A)$ must hold, which implies $x_kP_Rc(S,A)$. It follows from $(T,B) \in \Omega$ that $(T, B\backslash b'_{ik}) \in \Omega$. Therefore, we have $f(T,B) = \tilde{f}(T,B)\backslash x_j = \tilde{f}(T, B\backslash b'_{ik})\backslash x_j = f(T, B\backslash b'_{ik})$.

Case (iii): If $(T, B \backslash b'_{ik}) \in \Omega$, we have $c(T,B) = c(T, B\backslash b'_{ik}) = c(S,A)$ and $x_k P_R c(S,A)$ by a similar argument above. Therefore, $(T,B) \in \Omega$ must hold, which is a contradiction. Thus, we have $(T, B\backslash b'_{ik}) \notin \Omega$, which implies that $f(T,B) = \tilde{f}(T,B) = \tilde{f}(T, B\backslash b'_{ik}) = f(T, B\backslash b'_{ik})$. $\square$

References

[1] Akerlof, G.A., Kranton, R.E. (2010) "Identity economics: How our identities shape our work, wages, and well-being," Princeton, Princeton University Press.

[2] Banerjee, A.V. (1992) "A simple model of herd behavior," *Quarterly Journal of Economics* 107, 797-817.

[3] Bettman, J.R., Luce, M.F., Payne, J.W. (1998) "Constructive consumer choice processes," *Journal of Consumer Research* 25, 187-217.

[4] Cherepanov, V., Feddersen, T., Sandroni, A. (2013) "Rationalization," *Theoretical Economics*, forthcoming.

[5] Iwata, Y. (2011) "Social choice theory with collectively qualified agendas," presented at the Autumn Meeting of the Japanese Economic Association.

[6] Iwata, Y. (2012) "Social choice theory with collectively qualified agendas: The case where there exist voting agenda setters," presented at the 11th International Meeting of the Society for Social Choice and Welfare.

[7] Lleras, J.S., Masatlioglu, Y., Nakajima, D., Ozbay, E. (2011) "When more is less: Limited consideration," unpublished.

[8] Manzini, P., Mariotti, M. (2007) "Sequentially rationalizable choice," *American Economic Review* 97(5), 1824-1839.

[9] Masatlioglu, Y., Nakajima, D., Ozbay, E.Y. (2012) "Revealed attention," *American Economic Review* 102(5), 2183-2205.

[10] Salant, Y., Rubinstein, A. (2008) "(A, f): Choice with frames," *Review of Economic Studies* 75, 1287-1296.

[11] Thaler, R.H., Sunstein, C.R. (2008) "Nudge: Improving desicions about health, wealth, and happiness," New Haven, Yale University Press.

[12] Tversky, A., Kahneman, D. (1981) "The framing of decisions and the psychology of choice," *Science* 211, 453-458.