COMPLETING INCOMPLETE REVEALED PREFERENCE UNDER LIMITED ATTENTION

By YUSUFCAN MASATLIOGLU† and DAISUKE NAKAJIMA‡

†University of Michigan ‡Otaru University of Commerce

This paper discusses how to improve the identification of the preference of a decision-maker (DM) with limited attention proposed by Masatlioglu, Nakajima and Ozbay (2012). in "Revealed Attention". Their identification method relies on choice reversals so the obtained revealed preference is often incomplete. We propose three approaches to address this problem. The first one is accommodating a model-free approach, which respects the DM's choice in making a welfare analysis, as long as it does not contradict the revealed preference of Masatlioglu *et al*. The second approach incorporates the DM's exogenously obtained attention/inattention information into the model of Masatlioglu *et al*. The third approach is to take framings that influence the DM's attention into effect for the identification.
JEL Classification Numbers: D11, D60, D81.

1. Introduction

Revealed preference is one of the most important concepts in economics and has been applied to virtually all areas of economics. The basic idea is simple: whenever we observe a decision-maker (DM) choosing x while y is available, we infer that she prefers x over y. (Samuelson, 1938).

However, ample evidence suggests that people do not always consider all alternatives in making a decision. For instance, Lapersonnea *et al*. (1995) report that 22% of new car buyers consider only one brand.

This immediately raises a serious question against the revealed preference theory: if a DM does not consider all available alternatives, how can we infer her preference from her choice? She may prefer y over x but choose x simply because she overlooks y for some reason. For instance, imagine she uses a web search engine to look for alternatives. The search engine provides many results so the DM looks only at the first page of the search results, but y appears only on the second page.[1]

The pioneering work by Masatlioglu *et al*. (2012) extends the revealed preference theory to deal with a DM with limited attention. Their key idea is as follows: to infer a DM's preference, it is not enough to observe her choice in a single occasion. Rather, we need to observe how her choice *changes* when some alternatives become unavailable.

Masatlioglu *et al*. (2012) introduce the idea of consideration sets[2] into the revealed preference theory. Given a set of available alternatives, a DM only considers its subset (the consideration set) and picks from it only her most preferred alternatives.

---

[1] Aumann (2005) argues that such behaviour is still considered rational (at least boundedly rational) because she is choosing the best alternatives under her limited information about what is available. Thus, it is still meaningful to elicit her preference.

[2] Marketing studies extensively analyse consideration sets (see Wright and Barbour, 1977 and Hauser and Wernerfelt, 1990).

Specifically, Masatlioglu *et al*. impose a restriction on how her consideration set can be influenced when her feasible set changes, called an *attention filter*. The idea is simple: her consideration set should *not* be influenced when an alternative to which she does *not* pay attention is removed.[3] A choice function that is generated by preference maximization and an attention filter is called "choice with limited attention" (CLA).

Masatlioglu *et al*. demonstrate how to identify a preference of a DM following the CLA model. Imagine a DM picks x from $\{x, y, z\}$, but her choice shifts to z when y becomes unavailable. Masatlioglu *et al*. utilize such a seemingly irrational choice reversal to identify her preference. Suppose that she does not pay attention to y when it is available. Then, the property of an attention filter requires that her consideration set would be uninfluenced when y is removed so her choice could not shift from x to z. Thus, we conclude that she pays attention to y at that time. We now know that she picks x while considering y. This makes it possible to infer her preference: x over y.

Masatlioglu *et al*. (2012) provide the necessary and sufficient condition to identify: (i) the DM's preference; (ii) what the DM considers (revealed attention); and (iii) what the DM does not consider (revealed inattention). In addition, it provides the condition for a choice function to be compatible with the CLA model.

While the use of choice reversals for the identification is a clever idea, it has one serious drawback. When only a few choice reversals are observed, the identified preference is incomplete despite that the DM's true preference is complete. In the most extreme case where the DM's choice is perfectly compatible with the standard preference maximization, the CLA model cannot identify her preference at all. This is because we can attribute her choice entirely to her preference (maximizing her preference always considering all options) or to her inattention (considering only what she chooses).

The present paper discusses how to overcome such incompleteness of the revealed preference. Suppose the model of Masatlioglu *et al*. identifies that the DM prefers x over y but cannot find the ranking of z. We would like to rank z relative to x and y while respecting the already identified preference $x > y$. We discuss three approaches. The first approach aims to make a welfare ranking without identifying the DM's true preference regarding z, while the other two approaches still attempt identification through some other means.

The first approach is a compromise to model-free approaches, such as those proposed by Bernheim and Rangel (2009), Chambers and Hayashi (2012) and Apesteguia and Ballester (2014). These model-free approaches attempt a welfare analysis without modelling how the DM actually makes a decision. Masatlioglu *et al*. (2012) point out that model-free approaches may contradict the revealed preference elicited through the CLA model (Example 1 in Masatlioglu *et al*.). Nevertheless, we explain that the method of Apesteguia and Ballester (2014), if used only when Masatlioglu *et al*. cannot identify the DM's preference, can provide the secondary criterion of a welfare judgment, while the methods of the other two papers cannot.

The second approach is to use information about the DM's attention and inattention obtained exogenously through some other method, such as surveys and tracking systems used by Internet service providers. We illustrate how incorporating such information into the model of Masatlioglu *et al*. creates extra revealed preference, attention and inattention.

---

[3] Masatlioglu *et al*. (2012) illustrate that many heuristics often used to form a consideration set indeed generate an attention filter. They also explain that the DM's consideration set must be an attention filter if: (i) she forms a consideration set taking the cost of consideration into account; or (ii) she wrongly believes that her consideration set is her feasible set and is unaware that she is overlooking something (see pp. 2187–89 of Masatlioglu *et al*.).

The third approach is to utilize any observable factor that affects the DM's consideration set. Such a factor is called a framing. For instance, depending on which search engine the DM uses to find/narrow down alternatives, she may consider different sets of alternatives and so may choose different alternatives. We can identify her preference by utilizing a richer data set that includes not only her feasible sets and choices but also framings.

The organization of the paper is as follows. In the next section, we briefly review the model of Masatlioglu *et al.* and their identification method (revealed preference). Section 3 illustrates the incompleteness of the revealed preference of Masatlioglu *et al.* Sections 4–6 discuss the methods used to overcome this incompleteness and Section 7 concludes.

2. Review of model of Masatlioglu *et al.* (2012)

This section formally reviews the model of Masatlioglu *et al.* (2012) and illustrates its power and limitation. Throughout the paper, let X be a finite set of alternatives that may be available for a DM to choose. $\mathcal{X}$ stands for the collection of all non-empty subsets of X, which are the set of decision problems the DM may face.

2.1 Attention filter

The DM, when facing a decision problem $S \in \mathcal{X}$, does not always consider all alternatives in S but only those in its subset, denoted by $\Gamma(S)$, which is called her consideration set (under S). Formally, $\Gamma(\cdot)$ is a mapping from $\mathcal{X}$ to $\mathcal{X}$, where $\Gamma(S) \subset S$ for all $S \in \mathcal{X}$. Furthermore, Masatlioglu *et al.* restrict their attention to a particular type of consideration set as follows:

Definition 1 (Attention Filter): *A consideration set mapping Γ is an **attention filter** if* $\Gamma(S \setminus x) = \Gamma(S)$ *whenever* $x \notin \Gamma(S)$.[4]

Definition 1 says that if an alternative does not attract the attention of a DM, her consideration set does not change when it becomes unavailable. Masatlioglu *et al.* argue that this property is plausible: (i) when the limited consideration is entirely due to unawareness; and (ii) when the DM rationally restricts her attention, taking the cost of contemplating alternatives and her prior belief about the values of alternatives into account. Masatlioglu *et al.* also provide several heuristics that are often used to filter alternatives before serious comparisons, such as: (i) Top N (considering only top N options according to some criterion other than her preference); (ii) Top on Each Criterion (using several criteria and considering only those that are best/top N options in each criterion); and (iii) Most Popular Category (looking for the most popular category (including the largest number of available alternatives) and considering only those belonging to it). It is easy to verify that those heuristics generate an attention filter.[5]

2.2 Choice with limited attention

The DM, facing a decision problem S, chooses her most preferred alternative in $\Gamma(S)$. Masatlioglu *et al.* assume her preference is asymmetric and transitive, denoted by $\succ$. That

---

[4] Throughout the paper, unless it leads to confusion, we abuse the notation by omitting set delimiters: for example, writing $c(xy)$ instead of $c(\{xy\})$ or $S \setminus x$ instead of $S \setminus \{x\}$.

[5] See section IIA of Masatlioglu *et al.* for more detailed discussion.

is, $>$ is a standard preference over X without indifferences. The pair of an attention filter and a preference generates a choice function $c(\cdot)$, which is a mapping from $\mathcal{X}$ to X with $c(S) \in S$ for all $S \in \mathcal{X}$. A choice function generated by an attention filter and a preference is called a choice with limited attention (CLA). Formally:

Definition 2 (Choice with Limited Attention): *A choice function c is a **choice with limited attention (CLA)** if there exists an asymmetric and transitive preference $>$ over X and an attention filter Γ such that $c(S)$ is the $>$-best element in $\Gamma(S)$ for all $S \in \mathcal{X}$.*

2.3 Revealed preference and revealed attention/inattention

Imagine that we observe DM's choices (her choice function $c(\cdot)$) without knowing her preference $>$ and attention filter $\Gamma(\cdot)$. The main question of Masatlioglu *et al.* is how to identify them through observed data.[6]

For instance, imagine we observe the following choice data:

Example 1

S	xyz	xy	yz	xz
$c(S)$	x	x	y	z

The difficulty of the identification is that there are multiple CLA models that can explain the above choice.

	S	xyz	xy	yz	xz	Preference
CLA Model 1	$\Gamma_1(S)$	xy	xy	y	xz	$z >_1 x >_1 y$
CLA Model 2	$\Gamma_2(S)$	xyz	xy	yz	z	$x >_2 y >_2 z$

One can easily verify that both CLA Models 1 and 2 represent c, and Γ_1 and Γ_2 are both attention filters. There is no way to determine whether she prefers x or z.

Nevertheless, if these two CLA models were the only ones that represent c, we could at least conclude that the DM prefers x to y, and pays attention to y at xyz because both models agree on them.

Motivated by this example, Masatlioglu *et al.* extend the definitions of the revealed preference to the CLA model and also introduce the concepts of the revealed attention and the revealed inattention as follows:

Definition 3: *Assume c is a CLA and there are k different pairs of preference and attention filter which represent c: $(\Gamma_1, >_1), (\Gamma_2, >_2), \ldots, (\Gamma_k, >_k)$. In this case,*

- *x is revealed to be preferred to y if $x >_i y$ for all i,*
- *x is revealed to attract attention at S if $x \in \Gamma_i(S)$ for all i,*
- *x is revealed **not** to attract attention at S if $x \notin \Gamma_i(S)$ for all i.*

---

[6] de Clippel and Rozen (2014) discuss the case where the economist has a limited database where not all decision problems (and her choices) are observed.

Logically, one can verify whether x is revealed preferred to y by looking at all CLA models representing c. However, it is not practical when there are many alternatives. Indeed, Example 1, involving just three elements, has 13 distinct CLA representations.

Masatlioglu *et al*. provide a handy way to elicit the DM's preference, attention and inattention. Consider the following example:

S	$xyza$	xyz	xz
$c(S)$	y	x	z

When z is removed from xyz, the DM's choice shifts from x to z. This is not possible unless her consideration set changes before and after. Given the property of an attention filter, y must have attracted her attention at xyz (revealed attention). Thus, we are sure that she chooses x from xyz while considering y, so we conclude that she prefers x to y (revealed preference). Now we also find that she picks y from $xyza$, although she prefers x to y so she must have overlooked x at $xyza$ (revealed inattention).

In general, whenever we observe a choice reversal, $x = c(S) \neq c(S \setminus y)$ for some $S \in \mathcal{X}$, we can infer that the DM pays attention to y at decision problem S (otherwise, removing y could not affect her consideration so could not affect her choice). This means that she chooses x from S while considering y. Thus, we conclude that she prefers x to y. In such a case, let us write xPy.

xPy is not a necessary condition for the identification. Even if we never find such a choice reversal, we may find a chain of choice reversals. For instance, observing $x = c(S) \neq c(S \setminus z)$ and $z = c(T) \neq c(T \setminus y)$, we infer that she prefers x to z and z to y. Thus, we can indirectly identify her preference of x over y. Such indirect identifications are obtained by taking the transitive closure of P, denoted by P_R.[7] The following theorem states that xP_Ry is the necessary and sufficient condition for us to identify her preference of x over y. Thus, P_R is the revealed preference of the CLA model.

Theorem 1 (Revealed Preference): *(Revealed Preference) Suppose c is a CLA. Then, x is revealed prefered to y if and only if xP_Ry.*

Using the revealed preference, Masatlioglu *et al*. also provide the necessary and sufficient condition for the revealed attention and inattention as follows.

Theorem 2 (Revealed Attention/Inattention): *(Revealed (In)Attention) Suppose c is a CLA. 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 $z \in T \setminus S$.*

---

[7] It is equivalent to say that xP_Ry if and only if there exists a sequence of $a_0, a_1, \ldots, a_n$ with $a_0 = x$ and $a_n = y$ such that $x_0Px_1, x_1Px_2, \ldots, x_{n-1}Px_n$.

Masatlioglu *et al*. also provide the necessary and sufficient condition for a choice data to be compatible with the CLA model, called the "WARP with limited attention". It is a relaxation of the weak axiom of revealed preference (precisely, Sen's alpha condition). It is, indeed, equivalent that the revealed preference characterized by Theorem 1 has no contradiction (cycle), such as "xP_Ry and yP_Rx" or "xP_Ry, yP_Rz and zP_Rx".

3. Incompleteness of revealed preference of Masatlioglu *et al*. (2012)

Although Masatlioglu *et al*. successfully extend the idea of the revealed preference to decision-making with limited attention, their revealed preference is often incomplete despite that the true preference is complete. That is, we cannot always compare x <u>and</u> y. The identification method of Masatlioglu *et al*. relies on choice reversals, so the fewer choice reversals the DM makes, the fewer identifications of her preference we can make. In the extreme case where her choice is completely consistent (her choice is compatible with a preference maximization with full attention), we cannot make any inference about her preference based on the CLA model. For instance, if her choices *appear* to be maximizing a certain preference $x > y > z$, the CLA model can attribute her choices entirely to preference or entirely to her inattention ($\Gamma(S) = \{c(S)\}$). In the latter case, any preference is compatible with her choices, so there is no way to identify, even partially, her true preference.

How should we conduct a welfare analysis of the DM following the CLA model when the model cannot identify her preference? In what follows we present three distinct approaches.

4. Compromise to model-free approach

In the standard model, we first look for a preference that rationalizes behaviours of a DM and use such a preference for the welfare analysis. This is not an easy task when the DM is not fully rational because there is no preference explaining her choice.

Bernheim and Rangel (2009) argue that it is not necessary to rationalize choices to make a welfare judgment (Model-Free approach). Instead, they propose a very conservative but reasonable criterion: x is welfare improving over y if y is *never* chosen while x is present and x is chosen *at least once* when y is available.

Masatlioglu *et al*. point out the problem with Bernheim and Rangel's approach. Example 1 of Masatlioglu *et al*. illustrates that their approach may result in a welfare ranking completely opposite to the revealed preference elicited through the CLA model. Masatlioglu *et al*. argue that such an example highlights the importance of making a model in welfare analysis even when the DM is boundedly rational.

The problem is that even after constructing a model to explain the DM's choice, we very often have only partial identification of her preference. In such cases, (Masatlioglu *et al*., 2012, p. 2200) suggest a possibility of compromising to model-free approaches.

> Now imagine that a policymaker knows/believes a consumer behaves according to our model [the CLA model]. Then, he should first elicit her preference based on our method. Admittedly, it only provides an incomplete ranking (and empty if the choice data satisfies WARP). If the policymaker is forced to make a complete welfare judgment with a risk of making

mistakes, he can apply the other methods [model-free approaches] with the constraint of respecting the revealed preference generated by our model. In other words, these methods should be used to break the incompleteness of our revealed preference.

The idea of Masatlioglu *et al*. is that we should first model how a DM makes a choice and, as much as possible, identify her preference based on the model. The model-free approach should be used only to resolve the incompleteness of the revealed preference. For instance, suppose that she follows a CLA model but her choice is perfectly consistent with maximizing some preference, say $x \succ y \succ z$. We can use this for the welfare analysis, but only because the revealed preference based on the (CLA) model is empty.

Accepting this idea, the question is, how should we actually implement such analysis? Notice that Bernheim and Rangel's criterion is not particularly powerful because their criterion is also often incomplete. One can easily see that their criterion does not make any judgment in Example 1. Thus, we need to seek a model-free approach that generates a complete welfare ranking.

4.1 Scoring method

One such approach is a scoring method, which gives points to an alternative every time it is chosen (and sometimes subtracts points when it is unchosen). Then, it makes a welfare judgment based on their points. Chamber and Hayashi (2012) axiomatized such scoring methods. Clearly, the scoring method generates a complete welfare ranking over alternatives.

The problem of this method is that, even after discarding parts of scoring method's welfare rankings that contradict the revealed preference of the CLA model, we may still face the inconsistency. Consider the following example:

Example 2: *There are five elements x, y, z, a and b and the DM's true (unobservable) preference is according to this order. The DM has an attention filter where*

- *If a, b ∉ S, Γ(S) = S except Γ(xy) = y, and Γ(yz) = z.*
- *If a ∈ S or b ∈ S,*
  - *Γ(S) = S \ {x, y} when z ∈ S.*
  - *Γ(S) = S \ {x} when z ∉ S and y ∈ S.*
  - *Γ(S) = S when y, z ∉ S.*

It is straightforward to verify that Γ is an attention filter. The DM's choice exhibits a cycle when a or b is not available (c(xyz) = x, c(xy) = yc(yz) = zc(xz) = x). Her choice seems to maximize y ≻ x ≻ a ≻ b when a or b is present.

Let us consider how to make a welfare ranking among *x*, *y* and *z*. Because choice reversals happen only when *a* or *b* is removed or *z* is removed from *xyz*, the CLA model can only identify $x \succ z$ (as c(xyz) = x <u>and</u> c(xy) = y). Thus, we need to rely on the scoring method to find the ranking of *y* in relative to *x* <u>and</u> *z*. Let us consider the simplest scoring rule which assigns one point to an alternative every time it is chosen and each feasible set is observed just once.

Here, *x* is chosen from 5 sets (*xyz*, *xz*, *xa*, *xb* and *xab*). *y* is selected from 7 sets (*xy*, *yab*, *ya*, *yb*, *xyab*, *xya*, and *xyb*). *z* is chosen from 13 sets (*yz*, and 12 other sets (every set including *z* and "either *a* or *b*"). Therefore, the scoring method ranks *z* over *y* and *y* over *x*.

Notice that the CLA model has already identified *x* to be preferred to *z* so, combining it with the scoring rule result, we have a cycle. Because our policy is to use the scoring rule only for the secondary criterion, we must drop one of its judgments. Dropping *z* over *y*, we conclude that *y* is the best and *z* is the worst. Ignoring *y* over *x* instead, we judge that *x* is the top and *y* is the bottom. The problem is that there is no compelling standard to choose which of scoring method's judgments to drop.

In principle, any variant of the scoring method can encounter this problem. Therefore, the scoring method is not a good approach for completing the revealed preference of the CLA model.

4.2   Measuring the distance between choice and preference

Another model-free approach is to measure a "distance" between choice data and a preference over all alternatives, then use the preference with the shortest distance to the given choice data for a welfare analysis. This method is proposed by Apesteguia and Ballester (2014) along with the axiomatization of how to measure such a distance.[8]

Their model-free approach can naturally complete the revealed preference of Masatlioglu *et al*.: instead of using the closest (to a given choice data) preference among *all possible* preferences, pick the closest one among *only those that are consistent with the revealed preference of Masatlioglu et al*. For instance, in Example 2, we should look for the preference closest to the choice data among those including *x* > *z* (and *x*, *y*, *z* are all preferred to *a* and *b*), instead of all preferences over these five alternatives.

In general, any model-free approach generating a ranking of *all* alternatives at once may be used to complement the identification of Masatlioglu *et al*. simply by including the constraint that the ranking must respect the revealed preference. However, another type of model-free approach that evaluates the welfare of each alternative one by one (like the scoring method) has the potential problem found in Example 2.

5.   Utilizing exogenously obtained preference/attention/inattention information

The basic principle of the method of Masatlioglu *et al*. is to conclude that *x* is preferred to *y* only when we know that the DM chooses *x* after considering *y*. In doing so, Masatlioglu *et al*. check whether removing *y* shifts the DM's choice from *x* to something else.

The method of Masaltlioglu *et al*. method does not necessarily exclude the possibility that the DM's attention information can be obtained through different channels, such as eye-tracking, functional magnetic resonance imaging, the tracking system in Internet

---

[8]   They also propose the shortest distance as the measure of the DM's rationality but we do not necessarily agree with them in this regard. According to the CLA model, the DM who considers all alternatives and the DM who looks at only one alternative are not distinguishable. Nevertheless, both of them exhibit choices satisfying the weak axiom of the revealed reference. Apesteguia and Ballester would measure the distance between such choices and the corresponding preference to be zero so they would consider both as perfectly rational.

commerce, or even surveys. We now illustrate how combining this exogenously obtained information into the model of Masatlioglu *et al.* strengthens the identification power.

Suppose that the model of Masatlioglu *et al.* is silent whether or not *a* attracts the DM's attention at *S* (i.e. no revealed attention or inattention). Now imagine that we exogenously learn that the DM pays attention to *a* at *S*. This immediately reveals that $c(S)$ is preferred to *a*. Not only that, but the extra identification may give us more information about the DM's preference/attention. First, any alternative that is known to be better than $c(S)$, say *b*, is now revealed preferred to *a* thanks to the transitivity of her preference. Furthermore, if she chooses *a* in some other decision problem where *b* is present, we learn that she does not pay attention to *b* at that time. This illustrates how a piece of exogenously obtained preference information generates more revealed preference, attention and inattention.

Formally, let P_R be the revealed preference according to Theorem 1. Now we add the extra information $c(S) \succ a$ to P_R to create the new revealed preference P'. That is, $xP'y$ if and only if xP_Ry or $x = c(S), y = a$. Then, take the transitive closure of P', denoted by P'_R. This P'_R is now the revealed preference by utilizing the extra attention information. Using P'_R, we now characterize the revealed attention and inattention as well. The following proposition states that the extra information makes not only the revealed preference but also the revealed attention and inattention more complete because P'_R is more complete than P_R.

Proposition 1 (Revealed Preference, Attention and Inattention with Exogenously Obtained Attention Information): *Suppose c is a CLA and x does not have revealed attention or inattention at S according to Theorem 2. Now we exogenously learn that it must be $x \in \Gamma(S)$. Then,*

1. *x is revealed preferred to y if and only if xP'_Ry.*
2. *x is revealed not to attract attention at S if and only if $xP'_Rc(S)$.*
3. *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 $z \in T \setminus S$.*

Proof: See the proofs of Masatlioglu *et al.* (2012) for Theorems 1 and 2 and replace P_R with P'_R. ∎

Now imagine that we exogenously learn that a DM does not pay attention to some alternative. It does not appear to help us in identifying her preference. However, a careful investigation will help to reveal her preference. Consider the following example.

S	xyz	xy	yz	xz
c(S)	x	y	y	x
Revealed attention	xz	x	y	x

In this example, the revealed preference of Masatlioglu *et al.* is only xPz, elicited by $c(xyz) = x$ <u>and</u> $c(xy) = y$. There is no revealed inattention at all as she never chooses an alternative revealed inferior to something else. Now imagine that we exogenously learn that she does not pay attention to *z* at *xz*.

Notice that we have already identified that the DM pays attention to z at xyz; otherwise she would not exhibit a choice reversal when z is removed so we know she considers x <u>and</u> z at xyz but we are not sure whether she considers y at this point. Suppose that y was not considered at xyz. Then the property of an attention filter requires that removing y does not affect the DM's consideration set. Thus, she should still consider both x <u>and</u> z at xz, but the extra information tells us that this is not the case. Therefore, we must conclude that she pays attention also to y at xyz. Then, we learn that her choice <u>of</u> x at xyz is made under full attention, which reveals she prefers x to y. Now, look at xy. She chooses the revealed inferior alternative y so we identify that she does not pay attention to x at that time.

In sum, the extra revealed inattention information ($x \notin \Gamma(xz)$) at a smaller decision problem (xz) gives us the extra revealed attention at a larger problem ($y \in \Gamma(xyz)$), the extra revealed preference ($x \succ y$) and the extra revealed inattention in another decision problem ($x \notin \Gamma(xy)$).

We do not have the characterization theorem of the revealed preference, attention and inattention with exogenously obtained inattention information. Nevertheless, the above example illustrates how to find extra revelations in such a case. With inattention information <u>of</u> x at a decision problem S, we need to look at decision problems of S and one more element, say $S \cup a$, where we are sure that x is considered but unsure whether or not a is considered. If we find such a problem, then we can conclude that a must be considered there (otherwise x would be considered at S, a contradiction) so $c(S \cup a)$ must be preferred to a. The extra revealed preference can be used to find more revealed attention/inattention somewhere else.

6. Utilizing auxiliary choice data

6.1 Framing-dependent attention filter

The CLA model does not consider the possibility that some seemingly irrelevant factors affect the DM's consideration set (Tversky and Kahneman, 1981). Consider the example of Masatlioglu *et al.* of an attention filter generated by a search engine. Even when the DM is following the same procedure to narrow down alternatives, say considering only items appearing on the first page of the search results, she will form a different consideration set when she switches her search engine from Google to Bing.

Each search engine can cause choice reversals on different occasions. For instance, while using Google, a choice reversal may occur when y is removed from xyz. Using Bing, the DM may exhibit a choice reversal when z is taken away from ayz. Because each choice reversal provides a revelation of her preference, we can expect to improve the identification power by taking into account which search engine is used in each decision problem.

Formally, consider the following model. In each decision problem, the DM faces $S \in \mathcal{X}$ and a framing $f \in \mathcal{F}$.[9] A framing affects her consideration set. Examples of framings include:

- a search engine the DM uses to find/narrow down alternatives
- the amount of time she can allocate to make a search
- the order of alternatives to be presented

---

[9] This is due to Salant and Rubinstein (2008).

We call a pair of a feasible set and a framing an *extended decision problem*. Her choice at extended decision problem (S, f) is denoted by $c(S, f)$. Formally, define an extended choice function as $c : \mathcal{X} \times \mathcal{F} \to X$, where $c(S, f) \in S$ for all $(S, f) \in \mathcal{X} \times \mathcal{F}$.

As in the CLA model, she has a complete and transitive preference $\succ$ over X. Facing an extended decision problem (S, f), she only considers alternatives in $\Gamma(S, f) \subset S$, from which she picks the $\succ$-best element. Each framing $f \in \mathcal{F}$ generates an attention filter $\Gamma(\cdot, f)$. That is, $\Gamma(S \setminus x, f) = \Gamma(S, f)$ whenever $x \notin \Gamma(S, f)$. We call such Γ a *framing-dependent attention filter*. Our question is how to identify her preference $\succ$ through her extended choice function $c(\cdot, \cdot)$.[10]

Definition 4: *An extended choice function c is a choice with framing-dependent limited attention (FDLA) if there exists an asymmetric and transitive preference over X, denoted by $\succ$ and a framing-dependent attention filter Γ such that $c(S, f)$ is the $\succ$-best element in S for all $(S, f) \in \mathcal{X} \times \mathcal{F}$.*

Given the FDLA model, we illustrate how to identify the DM's preference through her choice. Because $c(\cdot, f)$ is a CLA for each fixed $f \in \mathcal{F}$, we can partially elicit her preference according to Theorem 1. Whenever we observe a choice reversal under framing f, that is $x = c(S, f) \neq c(S \setminus y, f)$, we learn that she prefers x to y. Denote such revelation by $xP^f y$. Repeat this process for different framings and collect all of such revelations. That is, xPy if $xP^f y$ for some $f \in \mathcal{F}$. Taking P's transitive closure, denoted by P_R, we learn that she prefers a to b whenever $aP_R b$. The following proposition states that it is the necessary and sufficient condition to identify her preference under the FDLA model.

Proposition 2 (Revealed Preference of the Framing-dependent Limited Attention Model): *Suppose an extended choice function c is a FCLA. Then x is revelaed to be preferred to y if and only if $xP_R y$.*

Proof: The if part is straightforward, so we only prove the only-if part. Suppose $xP_R y$ does not hold. We shall show that there exists a preference $\succ$ with $y \succ x$ that represents c with some framing-dependent attention filter.

Since P is a subset of the DM's true preference, it must be acyclical. If xPy does not hold, there exists a complete and transitive preference over X, denoted by $\succ$ with $y \succ x$. Given $\succ$, define Γ as

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

Clearly, $c(S, f)$ is the $\succ$-best element in $\Gamma(S, f)$. Thus, we only need to verify that $\Gamma(\cdot, f)$ is an attention filter for each fixed $f \in \mathcal{F}$. Suppose $a \notin \Gamma(S, f)$. By construction, it is $a \succ c(S, f)$ so $c(S, f)Px$ does *not* hold. Therefore, $c(S \setminus a, f)$ must be the same as $c(S, f)$. Thus, we have $\Gamma(S, \setminus a, f) = \Gamma(S, f)$ so $\Gamma(\cdot, f)$ is an attention filter. ■

Finally, we illustrate how to test the FDLA model. One may suspect that the implication of the FDLA model is no more than $c(\cdot, f)$ is a CLA model for every framing f. Actually, it is more. Observing $x = c(S, f_1) \neq c(S \setminus y, f_1)$ and $y = c(S, f_2) \neq c(S \setminus x, f_2)$, we have to conclude that $x \succ y$ and $y \succ x$, although each of $c(\cdot, f_1)$ and $c(\cdot, f_2)$ can be compatible with

---

[10] Iwata's (2013) generalized attention filter model is a special case of the FDLA model. In his model, the DM's consideration set is affected by saliences of alternatives, which corresponds to framings.

the CLA model. Thus, the FDLA model implies the consistency of the revealed preference not only within each framing but also *inter*framing-ly.

Proposition 3: *There exists a pair of a preference $>$ and a framing-dependent attention filter Γ that represents an extended choice function c if and only if P has no cycle.*

Proof: The only-if part is obvious. The if part directly follows from the proof of Proposition 2's only-if part. ∎

6.2 Using more information about framings

We would like to emphasize that the FDLA model is just a framework. The model alone, without any knowledge about how each frame relates to each others, the FDLA model simply provides $|\mathcal{F}|$ independent CLA models with a common preference. Thus, the revealed preference of the FDLA model alone is nothing more than a combination of the revealed preference gathered from each CLA model induced by each frame.

Indeed, we learn nothing from any choice reversals caused by framings (i.e. $c(S, f) \neq c(S, f')$). With some kinds of framings, it sounds too conservative. For instance, let f stand for the amount of time she can allocate for searching. She can consider more alternatives under $f = 10$ than under $f = 5$, so her choice under $f = 10$ should reflect her preference more that that under $f = 5$.

To capture this idea, let us consider the ordered set of framings $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$. Here, f_n is the "best" framing which allows the DM to consider the largest number of alternatives while f_1 is the "worst" that blocks the most of alternatives from her consideration. Formally,

Definition 5: *An framing-dependent attention filter Γ is ordered if the set of framings is ordered $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$ and $\Gamma(S, f_j) \subset \Gamma(S, f_i)$ for any S and $j < i$.*

An extended choice function c is a choice with ordered FDLA (OFDLA) model if there exists an asymmetric and transitive preference over X, denoted by $>$ and a ordered framing-dependent attention filter Γ such that $c(S, f)$ is the $>$-best element in S for all $(S, f) \in \mathcal{X} \times \mathcal{F}$.

That is, if a DM, facing a feasible set S, can consider x under an inferior framing f_j, she can do so under a superior framing f_i ($j < i$).[11]

With the OFDLA model, a choice reversal caused by framings also reveals her preference. Imagine she picks x facing feasible set S under a superior framing f_i but y under an inferior framing f_j ($j < i$) from the same feasible set. We can conclude she prefers x over y. This is because, $y = c(S, f_j)$ reveals that she considers y at (S, f_j) so she must have considered y also at (S, f_i) while selecting x.

This example suggests that we can identify her preference through two different types of choice reversals: one caused by removing an irrelevant alternative and one caused by changing framings. The latter type of the identification is possible only with the OFDLA model.

Formally, define xPy as

---

[11] A similar idea without framings is introduced by Lleras *et al*. (2015). In their model, if a DM can consider x in a larger decision problem, she can do so in a smaller decision problem.

- $x = c(S, f_i) \neq c(S \setminus y, f_i)$ for some (S, f_i) or
- $x = c(S, f_i)$ and $y = c(S, f_j)$ for some S and $i > j$

and let P_R denote P's transitive closure. The next proposition states that xP_Ry is the necessary and sufficient condition to identify $x \succ y$ with the OFDLA model.

Proposition 4 (Revealed Preference of the OFDLA model): *Suppose an extended choice function c is an OFDLA. Then x is revealed to be preferred to y if and only if xP_Ry.*

Proof: The only-if part is obvious, so we shall prove the if part here. Suppose xP_Ry does not hold. We shall show that x is not revealed preferred to y.

Since c is an OFCLA, P must include DM's true preference so P must have no cycle. Thus, there exists $\succ$ that includes P and $y \succ x$. We shall show that $\succ$ represents c with some ordered framing-dependent attention filter.

Define

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

Clearly, $c(S, f)$ is the $\succ$-best element in $\Gamma(S, f)$ so we only need to show that Γ is an ordered framing-dependent attention filter. Because we are working on a special class of FDLA model, Γ is an attention filter (see proof of Proposition 2).

Thus, we only need to show that Γ is ordered. Let $i > j$ and $a \in \Gamma(S, f_j)$. If $c(S, f_i) = c(S, f_j)$, then $a \in \Gamma(S, f_j) = \Gamma(S, f_i)$. If, $c(S, f_i) = c(S, f_j)$, then $c(S, f_i)Pc(S, f_j)$ so $c(S, f_i) \succ c(S, f_j)$. Therefore, $a \in \Gamma(S, f_j) = \Gamma(S, f_i)$. Thus, Γ is ordered. Therefore, c can be represented by an OFCLA model with $y \succ x$ so x is not revealed preferred to y. ∎

One may want to test whether framings a DM may face are, indeed, ordered. The following proposition provides a behavioural test to see if extended choice data are compatible with an OFDLM model. It states that the data is consistent with an OFDLM model if and only if the revealed preference has no contradiction.

Proposition 5: *There exists a pair of preferences and an ordered framing-dependent attention filter Γ that represents an extended choice function c if and only if P has no cycle.*

Proof: The only-if part is obvious. The if part directly follows from the proof of Proposition 4's only-if part. ∎

Indeed, this proposition is useful in determining which framing is superior to the other. We end the discussion by presenting such an example.

Imagine that there are two search engines, f_A and f_B. f_A is a superior search engine for some people while the others find f_B superior. Thus, we know that a DM has an ordered framing-dependent attention filer but are unsure which one is superior. That is, we are not sure if it is $\Gamma(S, f_A) \supset \Gamma(S, f_B)$ or $\Gamma(S, f_A) \subset \Gamma(S, f_B)$ for every S.

Suppose the DM makes the following choices:

	f_A:			
S	xyz	xy	yz	xz
$c(S)$	z	y	z	z

S	f_B:			
	xyz	*xy*	*yz*	*xz*
$c(S)$	*y*	*x*	*y*	*x*

Her choice is perfectly consistent under f_A but exhibits a reversal under f_B when z is removed from xyz. Thus, one may be tempted to conclude that f_A is superior to f_B so her choice under f_A reflects her preference: $z \succ y \succ x$ appears to be her true preference.

Actually, it is completely wrong and an OFDLA model will identify that f_B is superior and her preference is completely opposite. To see this, look at her choice reversal when z is removed from xyz under f_B. This reveals that she pays attention to z at xyz from which y is chosen. Thus, it reveals $y \succ z$.

Now observe the choice reversal when x, y and z are all available but framings are switched. She picks a revealed preferred alternative (y) under f_B compared to z under f_A. Thus, we learn that f_B is a superior framing for her. Now, consider another choice reversal due to framings when x <u>and</u> y are available. Becaus f_B is known to be superior, we understand that she prefers x over y. Therefore, we completely identify her preference: $x \succ y \succ z$.

Indeed, we can also identify the DM's attention and inattention completely in this example. Under f_A, she always pays attention only to what she chooses because she is always choosing her least preferred alternative. Because what is considered under f_A also attracts her attention at f_B, she considers yz at xyz (not x because x is known to be the best alternative while y is chosen) and all feasible alternatives at every pairwise choice. It is easy to verify that her attention is consistent with an OFDLA model with f_B being superior to f_A.[12]

This example, like Example 1 of Masatlioglu *et al*., illustrates the importance of the knowledge about the underlying choice procedure when we conduct the welfare analysis. In other words, welfare analysis is a more delicate task than it looks.

7. Conclusion

We believe Masatlioglu *et al*. (2012) have two major problems in their identification. The first problem is the incompleteness. We provide remedies for this problem: (i) compromise with model-free approaches; (ii) incorporating exogenously obtained attention/inattention information; and (iii) utilizing auxiliary information that affects the DM's consideration set (framings).

Thus, the future direction of the research should address the second problem of the framework of Masatlioglu *et al*.: the incompleteness of choice data. The identification method of Masatlioglu *et al*. relies on choice shifts. Thus, we must observe the same DM's choices from different feasible sets (and different framings when we work on the FDLA model) several times. In some situations, it is not possible because we may not be able to keep track of the DM's identity across different decision problems. In such situations, we

---

[12] Indeed, this example is created by the following story. While her preference is $x \succ y \succ z$, z is the most and x is the least advertised product. She pays attention only to the most advertised product (among feasible ones) under f_A and the top two most advertised under f_B.

only have market share data (or stochastic data) generated by many consumers' (DM) each of whom follows the CLA model with different parameters. For instance, we only learn percentages of consumers buying x under decision problem S (or extended decision problem (S, f)) and those under different decision problems. However, we do not have panel data that include who changes her choices between two decision problems.

Manzini and Mariotti (2014) is the first paper to work on stochastic limited attention. In their model, each alternative is considered with a certain probability that is *independent of whether other alternatives are considered or not*. Although their model creates stochastic choice data, we are unsure whether it can be viewed as a natural generalization of the deterministic attention filter of Masatlioglu *et al*. Dealing with such data to learn consumers' preferences is a difficult task but is necessary for many empirical studies.

Acknowledgements

REFERENCES

Apesteguia, J. and M. A. Ballester (2014) "A measure of rationality and welfare", mimeo.

Aumann, R. J. (2005) "Musings on Information and Knowledge", *Econ Journal Watch*, Vol. 2, No. 1, pp. 86–96.

Bernheim, D. B. and A. Rangel (2009) "Beyond Revealed Preference: Choice- Theoretic Foundations for Bebehavior Welfare Economics", *Quarterly Journal of Economics*, Vol. 97, No. 2, pp. 464–470.

Chambers, C. P. and T. Hayashi (2012) "Choice and Individual Welfare", *Journal of Economic Theory*, Vol. 147, No. 5, pp. 1818–1849.

de Clippel, G. and K. Rozen (2014) "Bounded rationality and limited datasets", mimeo.

Hauser, J. R. and B. Wernerfelt (1990) "An Evaluation Cost Model of Evoked Sets", *Journal of Consumer Research*, Vol. 16, No. 4, pp. 383–408.

Iwata, Y. (2013) "Generalized revealed attention", mimeo.

Lapersonnea, E., G. Laurent and J. L. Goffc (1995) "Consideration Sets of Size One: An Empirical Investigation of Automobile Purchases", *International Journal of Research in Marketing*, Vol. 12, No. 1, pp. 55–65.

Lleras, J. S., Y. Masatlioglu, D. Nakajima and E. Y. Ozbay (2015) "When More is Less: Choice by Limited Consideration", mimeo.

Manzini, P. and M. Mariotti (2014) "Stochastic Choice and Consideration Sets", *Econometrica*, Vol. 82, No. 3, pp. 1153–1176.

Masatlioglu, Y., D. Nakajima and E. Y. Ozbay (2012) "Revealed Attention", *American Economic Review*, Vol. 102, No. 5, pp. 2183–2205.

Salant, Y. and A. Rubinstein (2008) "(A, F): Choice with Frames", *The Review of Economic Studies*, Vol. 75, No. 4, pp. 1287–1296.

Samuelson, P. A. (1938) "A Note on the Pure Theory of Consumers' Behavior", *Econometrica*, Vol. 5, No. 17, pp. 61–71.

Tversky, A. and D. Kahneman (1981) "The Framing of Decisions and the Psychology of Choice", *Science*, Vol. 211, No. 4481, pp. 453–458.

Wright, P. and F. Barbour (1977) "Phased Decision Strategies: Sequels to An Initial Screening", in M. K. Starr and M. Zeleny, eds, *Studies in Management Sciences, Multiple Criteria Decision Making*, Amsterdam: North-Holland, pp. 91–109.