Article

Path-Independent Consideration

Juan Lleras [1], Yusufcan Masatlioglu [2]*, Daisuke Nakajima [3], Erkut Ozbay[4]

[1] Deloitte; jslleras@gmail.com
[2] University of Maryland; yusufcan@umd.edu
[3] Otaru University of Commerce; nakajima@res.otaru-uc.ac.jp
[4] University of Maryland; ozbay@umd.edu
* Correspondence: yusufcan@umd.edu;

Abstract: In the context of choice with limited consideration, where the decision-maker may not pay attention to all available options, the consideration function of a decision maker is *path-independent* if her choice cannot be manipulated by the presentation of the choice set. This paper characterizes a model of choice with limited consideration with path independence, which is equivalent to a consideration function that satisfies both the attention filter and consideration filter properties from [1] and [2], respectively. Despite the equivalence of path-independent consideration with the consideration structures from [1] and [2], we show that in order to have a choice with limited consideration that is path-independent, satisfying both axioms on the choice function that characterize choice limited consideration with attention and consideration filters unilaterally (from [1] and [2]) is necessary but not sufficient.

Keywords: Revealed Preferences; Limited Attention; Consideration Set; Choice Reversals; Path Independence

Citation: Lleras, J.; Masatlioglu, Y.; Nakajima, D.; Ozbay E. Path-Independent Consideration. *Games* **2021**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

1. Introduction

Individuals do not always compare all alternatives in making a decision, particularly when a decision problem is complex or contains many alternatives. For instance, [3] report that 22% of new car customers consider only one brand out of more than 100 car brands available in USA. Instead, a decision maker (DM) often forms a consideration set, a subset of her actual feasible set, and ignores the rest (see e.g., [4]).

The issue of limited consideration shakes the main principle of revealed preferences and raises the following question: how can we identify the DM's preference by observing her choice under limited attention? It is not straightforward to answer this question by the standard revealed preference tools since the revealed preference implicitly relies on the knowledge of consideration, which is not observable in real life. [1] and [2] managed to provide answers the above question when she picks her most preferred item from her consideration set, not from her actual entire feasible set. These two papers impose certain assumptions on the consideration set and employ the feasible set variations to reveal preferences.

The conditions on the formation of consideration sets contemplated by [1] and [2] are called *attention filter* and *competition filter*. According to the attention filter, the consideration set is not affected when overlooked alternatives are removed from her feasible set; whereas the competition filter requires that if she ignores some alternatives, she will also do so when her feasible set expands. Attention filter is plausible when her inattention is based on unawareness, while for a competition filter, the alternatives are competing for DM's attention (such as in a large supermarket).

The consideration set formation in both models are vulnerable to manipulation. To illustrate a manipulation possibility in a simple example, consider the following consideration set formation which is an attention filter (but not a competition filter). There are three products x, y, and z, where z is considered only when both x and y are present.[1] The consideration set of such DM is $\{x, y, z\}$ if everything is presented at once. On the other hand, when x and z are offered first and then y, z is never considered. Hence, the consideration set of a DM may be manipulated by way the choice problem is presented.[2]

More precisely, if instead of presenting a large menu, A, the choice problem is divided into sub-menus, S and T where $S \cup T = A$ and first the sub-menu S is presented. Once a DM forms her consideration based on S, then the sub-menu T is presented. Now, the DM forms her consideration based on the sub-menu T and the consideration set of S. If this consideration set is different than presenting the large menu A as a whole, a firm, that is responsible for presenting DMs with menus, may have the ability to manipulate the DMs' consideration sets.

In this paper, we are interested in situations where the formation of consideration sets are free from this kind of manipulation. This non-manipulability requirement is captured by the well-known property of choice theory, *path-independence*, proposed by [5], which imposes a "consistency" requirement on how consideration sets are determined across comparable situations. Path-independence requires that the consideration set cannot be manipulated by changing the presentation of the set of alternatives.

Our model rests on the realistic assumption of path-independence and thus offers a better understanding non-manipulation in the consideration set formation. This property not only will eliminate manipulation examples mentioned above but also makes the revealed preference analysis more complete.

In the literature, the formation of the consideration sets has been motivated by behavioral reasons, such as shortlisting in [6][3], rationalization in [16], and categorization in [17]. At first glance, the path-independence property may be *normatively* plausible but does not sound *behaviorally* plausible. However, the path-independence property is equivalent to consideration set formation satisfying both *the attention filter* by [1] and *the competition filter* by [2]. Therefore, the path-independence property, though it appears to be demanding, has a behavioral background, too. In Section 2, we provide a list of heuristics generating a consideration set satisfying the path-independence property.

The organization of this paper is as follows: Section 2 introduces notations and the model, Section 3 provides our characterization, Section 4 analyzes the revealed preference.

2. Model

We denote the set of alternatives by X, which is an arbitrary non-empty finite set. $\mathcal{X}$ denotes the set of all non-empty subsets of X with cardinality at least 2. Each subset of X is a choice problem. Let c be a choice function: $c : \mathcal{X} \to X$ and $c(S) \in S$ for all $S \in \mathcal{X}$. Our DM assigns a unique alternative for each choice problem S. Let $\succ$ be a complete, transitive and antisymmetric binary relation (a linear order) over X. $\max(\succ, S)$ represents the best element in S with respect to $\succ$.

Let $\Gamma : \mathcal{X} \to \mathcal{X}$ be a self-map on $\mathcal{X}$, that is, $\Gamma(S) \subset S$ for all $S \in \mathcal{X}$. $\Gamma(S)$ represents the consideration set under $S \in \mathcal{X}$; that is, the set of alternatives is considered when the DM is facing feasible set S. Since the DM can only consider options that are available, $\Gamma(S)$ must be a subset of S.

The next definition describes the behavior of DM with limited consideration. that is, the DM picks her most preferred alternative within her consideration set, not the entire feasible set.

[1] This example is from page 2198 in [1] called "pairwisely unchosen".

[2] A DM who has a competition (but not an attention) filter can be manipulated in a similar manner. For instance, suppose she ignores y when x is present and z when y is given. This is a competition filter but not an attention filter. She considers only x when all of the three are presented at once. On the other hand, she will considers x and z when she first sees x and y (so discards y) and then is given z, as y has been already gone.

[3] For the extensions of the shortlisting procedure, see also [7], [8], [9], [10], [11], [12], [13], [14] and [15].

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

$$c(S) = \max(\succ, \Gamma(S))$$

The LC model is very broad. Indeed, without further condition on Γ, any choice behavior has an LC representation. [1] and [2] propose two conditions on consideration sets. [1] requires that a DM who overlooks some feasible alternative, has the same consideration set if that alternative is removed. This property is called Attention Filter (AF).[4]

$$\textbf{AF}: \text{If } x \notin \Gamma(S \cup x) \text{ then } \Gamma(S \cup x) = \Gamma(S)$$

[2] imposes that when the size of the opportunity set gets larger, DMs tend to overlook more options. This property is called Competition Filter (CF).[5]

$$\textbf{CF}: \text{If for all } x \text{ and } y, x \notin \Gamma(S) \text{ then } x \notin \Gamma(S \cup y).$$

It turns out that if a consideration function satisfies both AF and CF, then it also satisfies the well-known path independence condition. In consideration context, path independence would mean that the final consideration set is independent of the way the alternatives were initially presented.

$$\textbf{PI}: \ \Gamma(S \cup T) = \Gamma(\Gamma(S) \cup T) \text{ for all } S \text{ and } T.$$

Path independent imposes a consistency property on how the consideration sets are determined. An important benefit of PI is the elimination of manipulation possibilities. That is, the consideration set of our DM cannot be altered by different ways of presenting the available set of alternatives. We now define our model formally.

Definition. *We say c is a π-LC model if c is a choice with limited consideration and Γ satisfies path independence.*

Before we state our main result, we first state the result of [28] stating that AF and CF together is equivalent to PI. Then we provide a list of examples of consideration set formations satisfying path independence.

Theorem 1. *([28]) A consideration set Γ is path independent if and only if it satisfies AF and CF.*

There are several examples of non-manipulable heuristics. For example, in elimination by aspects of ([29]), at each stage, the DM selects an aspect perceived and eliminates alternatives lacking that attribute. The DM continues selecting aspects and eliminating products. The process stops at stage n when there is no alternative left. The consideration set is the set of alternatives survived at stage $n-1$. Additional examples can be listed as follows: The DM considers

- only the three cheapest suppliers in the market ([30]).
- the products that appear on the first page of the web search and/or sponsored links ([31]).
- the first N available alternatives according to an exogenously given order ([32]).
- only a job candidate if she is the best in a program. Or consider the top-two job candidates from all first-tier schools and the top candidate from second-tier schools.

4 Equivalent properties have been considered in the literature on choice functions: Postulate 5* of [18], Axiom 2 of [19], the strong superset's axiom of [20], the Outcast Property of [21], Axiom 2 of [22], and the irrelevance of rejected contracts of [23].

5 In the choice theory literature, this property is mainly known as Sen's α axiom ([24]), Postulate 4 of [18], C3 of [25], the Heritage property of [21], or the Heredity property of [26]. See also [27] for a detail discusuion for these properties on choice functions.

- only the cheapest car, the safest car, and the most fuel-efficient car on the market.

3. Characterization

In this section, we provide necessary and sufficient conditions for our model where the consideration set mapping satisfies path independence. In other words, we ask how one could decide whether choice data is consistent with the π-LC model.

The weak axiom of revealed preference (WARP) characterizes the preference maximization. However, WARP does not distinguish between "being feasible" and "being considered." Therefore, one cannot decide that an alternative is chosen from a choice problem without confirming that the alternative is considered. The question is how we can infer that an alternative is considered. The answer for this question depends on the structure imposed on the consideration set. [1] and [2] provided two axioms characterizing their models. We now state each of them.

In [1], if removing an alternative from a set changes the DM's choice, they infer this alternative is considered. which is the additional requirement for x^* to be chosen from T. [1] introduced the following axiom.

WARP-AF: For any nonempty S, there exists $b^* \in S$ such that, for any T including b^*,

$$c(T) = b^* \text{ whenever } \quad \begin{array}{l} (i)\ c(T) \in S, \textit{ and} \\ (ii)\ c(T) \neq c(T \setminus b^*) \end{array}$$

An alternative way to state this axiom is through revealed preferences. Whenever the choices change as a consequence of removing an alternative, the initially chosen alternative is preferred to the removed one. Formally, for any distinct x and y, define:

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

WARP-AF indeed guarantees that the binary relation P^{AF} defined in (1) is acyclic and it fully characterizes the class of choice functions generated by an attention filter. The lemma from [1] states that WARP-AF is equivalent to the fact that P^{AF} has no cycle.

Lemma 1 ([1])**.** *P^{AF} is acyclic if and only if c satisfies WARP-AF.*

Given this result, we now illustrate that the just observing the following two choice reversals falsifies WARP-AF.

$$c(S_1) = x \neq c(S_1 \setminus y), \text{ and } c(S_2) = y \neq c(S_2 \setminus x).$$

These observations reveal that $xP^{AF}y$ and $yP^{AF}x$, which is a cycle. By Lemma 1, WARP-AF is violated.

In [2], if an alternative is considered in larger set, then it must be considered in a smaller set. That is, if $b^* = c(T')$ for some $T' \supset T$, then $b^* \in \Gamma(T')$ since a necessary condition for choice is that the b^* is considered. Since Γ is a competition filter, $b^* \in \Gamma(T)$. The following axiom of [2] summarizes this discussion.

WARP-CF For any nonempty S, there exists $b^* \in S$ such that for any T including b^*,

$$c(T) = b^* \text{ whenever } \quad \begin{array}{l} (i)\ c(T) \in S, \textit{ and} \\ (ii)\ b^* = c(T') \text{ for some } T' \supset T \end{array}$$

Similar to above, we can state this axiom through revealed preference. Whenever a DM's choices from a small set and a larger set are inconsistent, the former reflects her true preference under CF more than the latter. Formally, for any distinct x and y, define the following binary relation:

$$xP^{CF}y \text{ if } x = c(S) \text{ and } y = c(T) \text{ such that } \{x, y\} \subseteq S \subset T \tag{2}$$

Similar to [1], the binary relation P^{CF} defined in (2) is acyclic and it fully characterizes the class of choice functions generated by an attention filter. The lemma from [2] states that WARP-CF is equivalent to the fact that P^{CF} has no cycle.

Lemma 2 ([2])**.** *P^{CF} is acyclic if and only if c satisfies WARP-CF.*

Only three observations can falsify this axiom. For example, consider the following choice pattern:

$$c(\{x, y, z, t\}) = y, \;\; c(\{x, y, z\}) = x \text{ and } c(\{x, y\}) = y.$$

The first two observations imply that $xP^{CF}y$. Similarly, the last two observations indicate $yP^{CF}x$, which leads a cycle of two. By Lemma 2, the axiom is violated.

Since each axiom characterizes the corresponding model, and the path-independence is equivalent to an AF and CF, it is tempting to claim that WARP-AF and WARP-CF would characterize the π-LC model where the consideration structure satisfies both AF and CF. However, that is not the case. The following example satisfies the both axioms but it cannot be represented a pair of preference and consideration set satisfying path independence.

Consider the following "choosing pairwisely unchosen" pattern (the chosen alternative from $\{x, y, z\}$ is never chosen in any binary comparisons):

$$c(\{x, y, z\}) = x, \quad c(\{x, y\}) = y, \quad c(\{y, z\}) = y, \quad c(\{x, z\}) = z.$$

First, note that c satisfies both WARP-AF and WARP-CF. To see this, note c involves just two choice reversals: when y or z is removed from $\{x, y, z\}$. Therefore, the revealed preference generated by AF is just $xP^{AF}y$ and $xP^{AF}z$ and the one by CF is only $yP^{CF}x$ and $zP^{CF}x$. Neither of them contains a cycle so c satisfies the both axioms. Nevertheless, x must be best in AF and worst in CF, which is not compatible so c cannot be represented by Γ satisfying AF and CF simultaneously. In other words, there is no $(\Gamma, \succ)$ pair that can rationalize this data, where Γ satisfies PI.

The axiom we propose is a stronger version of both WARP-AF and WARP-CF. Remember that both axioms requires that every set S has the "best" alternative x^* and it must be chosen from any other decision problem T as long as it attracts attention. Remember that, with an attention filter, an alternative, say x^*, attracts attention at a choice set, T, when removing it changes the choice, i.e., $c(T) \neq c(T \setminus x^*)$. Now that we assume that the consideration set is path independent, we can also conclude it when we know x^* is paid attention to at some bigger decision problem $T' \supset T$ by observing $c(T') \neq c(T' \setminus x^*)$. Therefore, we need to state that if the removal of x^* changes the choice in some superset of T, then it attracts attention at T.

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

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

It turns out that WARP-PI is the necessary and sufficient condition for the π-LC model.

Theorem 2. *(Characterization) A choice function satisfies WARP-PI if and only if it is a π-LC model.*

Theorem 2 characterizes a special of class of choice behavior we studied earlier. The characterization involves a single behavioral postulate which is stronger than both WARP-AF and WARP-CF. We show that this model has higher predictive power, which comes with diminishing explanatory power: "choosing pairwisely unchosen" is no longer within the model.

We finalize this section by considering Attraction Effect type choice pattern:

$$c(\{x,y,z\}) = y,\ \ c(\{x,y\}) = x,\ \ c(\{y,z\}) = y,\ \ c(\{x,z\}) = x.$$

It is routine to verify that this choice behavior satisfies WARP-PI.[6] Hence Theorem 2 implies that it is consistent with a π-LC model. The choice reversal between $\{x,y,z\}$ and $\{x,y\}$ yields that her preference must be $x \succ y \succ z$. This implies that we can uniquely pin down preference for this choice behavior. Note that this is not true for the models of [1] and [2].

In addition to unique preference, we can also reveal the unique consideration set mapping. To see this, consider the set $\{x,y,z\}$. First of all, the choice, which is y, must be in the consideration set. Since removing z changes the choice, therefore z is also in it (attention filter). Finally, we know x is better than the choice from above discussion, x does not belong the consideration set of $\{x,y,z\}$. Hence $\Gamma(\{x,y,z\}) = \{y,z\}$. In addition, path independence requires that y and z attract attention whenever they are available, which pins down the consideration set mapping uniquely for this example.

$$\Gamma(\{x,y,z\}) = \{y,z\},\ \ \Gamma(\{x,y\}) = \{x,y\},\ \ \Gamma(\{y,z\}) = \{y,z\}, \text{ and } \Gamma(\{x,z\}) = \{x,z\}.$$

Theorem 2 states that it is possible to test our model non-parametrically from observed choice behaviour even when the consideration sets themselves are unobservable.[7]

4. Revealed Preference

In this section, we discuss the revealed preference of our model. One might suspect that $P^{AF} \cup P^{CF}$ should be the revealed preference of this model. The following example illustrates that this is not the case. The example shows that there is an additional preference revelation, which cannot be captured even by the transitive closure[8] of $P^{AF} \cup P^{CF}$.

Example 1. [Hidden Revelation] Consider the following behavior with four alternative x,y,z,t: A DM chooses z whenever z is available except in two occasions $\{x,z,t\}$ and $\{z,t\}$, from which the DM chooses t. When z is not available, the DM chooses t whenever t is available. Lastly, the DM chooses x from $\{x,y\}$. It is routine to show that this choice behaviour satisfies WARP-PI, hence it is a π-LC model.[9]

Let X be $\{x,y,z,t\}$. Consider the following choice behaviour on X: A DM chooses z from whenever z is available except from $\{x,z,t\}$ and $\{z,t\}$, x from $\{x,y\}$, and t from the rest of decision problems.

The DM exhibits only one choice reversal: $c(\{x,y,z,t\}) = z \neq c(\{x,z,t\}) = t$. This implies that we must have $zP^{AF}y$ and $tP^{CF}z$. This implies that t must be better than z and z is better than y (of course, t is better than y). However, there is no revelation between x and y according to $P^{AF} \cup P^{CF}$.

We now illustrate that in our model, we reveal that x must be better than y. To see this, $c(x,y,z,t) = z$ and $c(x,z,t) = t$ implies $y \in \Gamma(\{x,y,z,t\})$. Then we must have $y \in \Gamma(\{x,y\})$. Since x is chosen from $\{x,y\}$, x must be better than y. Yet, this is not captured by either P^{CF} or P^{AF}.

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

7 There are two recent papers which provide similar characterizations under stronger data requirements. While [33] assumes that ex-ante menu preferences are observable, [34] assumes both preferences and choices are observable.

8 We can conclude x must be preferred to z when $xP^{AF}y$ and $yP^{CF}z$. The revealed preference illustrated in this example is not captured even by this process.

9 The following $\succ$ and Γ represent the choice behavior: $t \succ z \succ x \succ y$ and $\Gamma(X) = \{x,y,z\}$, $\Gamma(\{y,z,t\}) = \{y,z\}$, and $\Gamma(S) = S$ for all other S. Clearly, Γ satisfies AF and CF property.

Given this observation, we provide a characterization for the revealed preference when Γ is known to be path independent. To do this, we consider cyclical choice behavior: $c(\{x,y,z\}) = x, \quad c(\{x,y\}) = x, \quad c(\{y,z\}) = y, \quad c(\{x,z\}) = z$. Here, we can uniquely pin down the preference for the cyclical choice example when Γ is path independent. To see this, first note that $c(\{x,y,z\}) = x$ implies that the DM pays attention to x at $\{x,y,z\}$, so she does at $\{x,z\}$ (revealed attention due to competition filter). Since she picks z from $\{x,z\}$, we can conclude that she prefers z over x (revealed preference). Since $c(\{x,y,z\}) \neq c(\{x,z\})$, y must attract attention at $\{x,y,z\}$ (revealed attention due to attention filter). Since she picks x from $\{x,y,z\}$, we can conclude that she prefers x over y (revealed preference). Therefore, her preference is uniquely pinned down: $z \succ x \succ y$.

Now we generalize this observation. Suppose $c(T) \neq c(T \setminus y)$ and $c(T) \neq y$. Then we conclude that y must be paid attention to at T, hence $c(T) \succ y$. Since Γ is path independent, $c(T)$ must attract attention at any decision problem S smaller than T including $c(T)$. Therefore, if $c(S) \neq c(T)$, $c(S)$ is revealed to be preferred to $c(T)$, hence $c(S) \succ c(T) \succ y$. Formally, for any distinct pair of x and y define:

$$xP^{PI}y \text{ if there exist } S \text{ and } T \text{ such that} \quad \begin{array}{l} (i)\ \ \{x,y\} \subset S \subset T \text{ and } x = c(S) \\ (ii)\ c(T) \neq c(T \setminus y) \end{array}$$

Note that the second condition in the definition of P^{PI} holds trivially when y is equal to $c(T)$. This implies that $c(T)$ must have been considered not only at T but also at any decision problem S smaller than T including $c(T)$ since Γ satisfies PI. Therefore, whenever $c(T) \subseteq S \subset T$ and $c(T) \neq c(S)$, we have $x = c(S) \succ c(T) = y$.

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

Proposition 1. *Suppose c is a π-LC model. Then, x is revealed to be preferred to y if and only if $xP_R^{PI}y$.*

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

Finally, note that P^{PI} must include both P^{AF} and P^{CF}, but it might include more. To show this, we revisit Example 1 and illustrate that P^{PI} captures x is better than y, which was missed by both P^{AF} and P^{CF}. Let $T = \{x,y,z,t\}$ and $S = \{x,y\}$. Since $c(T) \neq c(T \setminus y)$ and $c(S) = x$, we must have $xP^{PI}y$. Hence our model reveals more preference information than the combined models of [1] and [2].

Author Contributions: Authors contributed equally to this work.

Funding: Yusufcan Masatlioglu gratefully acknowledges financial support from the National Science Foundation through grant SES-1628883.

Acknowledgments: We would like to thank Emel Filiz-Ozbay, Paola Manzini, Marco Mariotti, Neslihan Uler for their helpful comments.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Proofs

The Proof of Theorem 2

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

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

Lemma A1. *xP^{PI} is acyclic if and only if c satisfies WARP-PI.*

The proof of Lemma A1 is completely analogous to the proofs of Lemmas 1 and 2 (see [1,2]), hence we skip it here.

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

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

Lemma A2. *For any S,*

(i) $c(S)$ is a minimum block of S. There is no other minimum block that includes $c(S)$.
(ii) If B is a minimum block of S other than $c(S)$, then $c(S) \succ x$ for all $x \in B$.
(iii) If $c(T) \neq c(S)$ and $T \supsetneq S$, then T has a minimum block that is a subset of $T \setminus S$.

Proof. Part (i) and (iii) are trivial. For Part (ii), let $B' = B \setminus x$ (it may be empty). Then we have

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

Therefore, we have $c(S)P^{PI}x$ so it must be $c(S) \succ x$. $\square$

Claim 1. *Γ is path independent.*

Proof. Γ satisfies CF by construction so we shall prove that Γ satisfies AF. Suppose not, i.e., $x \notin \Gamma(S)$ and $\Gamma(S) \neq \Gamma(S \setminus x)$. Since Γ satisfies CF, we must have $\Gamma(S) \subseteq \Gamma(S \setminus x)$. Hence $\Gamma(S) \subsetneq \Gamma(S \setminus x)$, that is, there exists $y \in S$ such that $y \notin \Gamma(S)$, but $y \in \Gamma(S \setminus x)$. Then there exists $T \supset S$ such that (i) $T \setminus x$ has a minimum block B and y is the worst element in B and (ii) none of elements in B are included in $\Gamma(T')$ for any $T' \supsetneq T \setminus x$.

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

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

Therefore, by Lemma A2 (iii), T has a minimum block that is a subset of $x \cup B$ so at least one element in $x \cup B$ must be in $\Gamma(T)$, which is a contradiction. $\square$

Now we want to show that $(\succ, \Gamma)$ represents c. Since Lemma A2 (i) implies that $c(S) \in \Gamma(S)$, all we need to show is that $c(S) \succ y$ for all $y \in \Gamma(S) \setminus c(S)$.

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

Proof. Since $y \in \Gamma(S)$, there exists $T \supset S$ such that $y \in \Gamma(T)$. Furthermore, T has a minimum block B where y is the worst element and none of elements in B is in $\Gamma(T')$ for any $T' \supsetneq T$. There are three easy cases: (i) if $c(S) = c(T)$ then by Lemma A2 (ii) we have $c(S) = c(T) \succ y$, (ii) if $y = c(T)$ then we have $c(S)P^{PI}y$ so it must be $c(S) \succ y$, and finally (iii) if $c(S) \in B$, then $c(S) \succ y$ by construction. Therefore, we only need to investigate the case when $y \neq c(T) \neq c(S)$ and $c(S) \notin B$. Note that $c(T) \succ y$ in this case by Lemma A2 (ii).

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

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

By Lemma A2 (iii), S has a minimum block B' that is a subset of $S \setminus S'' \subset B$. Since $c(S) \notin B'(\subset B)$, every element in B' is worse than $c(S)$ by Lemma A2 (ii). Since y is the worst element in B that is a superset of B', we conclude $c(S) \succ y$.

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

Since $y \neq c(T) = c(T \setminus \{B \setminus y\}) \neq c(T \setminus B)$, and $c(S \setminus \{B \setminus y\}) \in \{T \setminus \{B \setminus y\}\}$, we have $c(S \setminus \{B \setminus y\})\ P^{PI}\ y$. Therefore, $c(S) \succ y$ because of $c(S \setminus \{B \setminus y\}) = c(S)$. □

References

1. Masatlioglu, Y.; Nakajima, D.; Ozbay, E.Y. Revealed Attention. *American Economic Review* **2012**, *102*, 2183–2205.
2. Lleras, J.S.; Masatlioglu, Y.; Nakajima, D.; Ozbay, E.Y. When More Is Less: Limited Consideration. *Journal of Economic Theory* **2017**, *170*, 70–85.
3. Lapersonne, E.; Laurent, G.; Le Goff, J.J. Consideration sets of size one: An empirical investigation of automobile purchases. *International Journal of research in Marketing* **1995**, *12*, 55–66.
4. Wright, P.; Barbour, F. Phased decision strategies: Sequels to an initial screening. In *Studies in Management Sciences, Multiple Criteria Decision Making*; Starr, M.; Zeleny, M., Eds.; North-Holland, Amsterdam, 1977; pp. 91–109.
5. Plott, C. Path independence, rationality and social choice. *Econometrica* **1973**, *41*, 1075–1091.
6. Manzini, P.; Mariotti, M. Sequentially Rationalizable Choice. *American Economic Review* **2007**, *97*, 1824–1839.
7. Dutta, R.; Horan, S. Inferring rationales from choice: identification for rational shortlist methods. *American Economic Journal: Microeconomics* **2015**, *7*, 179–201.
8. Au, P.H.; Kawai, K. Sequentially rationalizable choice with transitive rationales. *Games and Economic Behavior* **2011**, *73*, 608–614.
9. Tyson, C.J. Behavioral implications of shortlisting procedures. *Social Choice and Welfare* **2013**, *41*, 941–963.
10. Horan, S. A simple model of two-stage choice. *Journal of Economic Theory* **2016**, *162*, 372–406.
11. Yildiz, K. List-rationalizable choice. *Theoretical Economics* **2016**, *11*, 587–599.
12. Cuhadaroglu, T. Choosing on influence. *Theoretical Economics* **2017**, *12*, 477–492.
13. Matsuki, J.; Tadenuma, K. Choice via grouping procedures. *International Journal of Economic Theory* **2018**, *14*, 71–84.
14. Horan, S.; Sprumont, Y. Two-stage majoritarian choice **2020**.
15. Geng, S.; Ozbay, E.Y. Shortlisting with a Limited Capacity. *Forthcoming: Journal of Mathematical Economics* **2020**.
16. Cherepanov, V.; Feddersen, T.; Sandroni, A. Rationalization. *Theoretical Economics* **2013**, *8*.
17. Manzini, P.; Mariotti, M. Categorize then choose: Boundedly rational choice and welfare. *Journal of the European Economic Association* **2012**, *10*, 1141–1165.
18. Chernoff, H. Rational Selection of Decision Functions. *Econometrica* **1954**, *22*, 422–443.
19. Fishburn, P.C. Semiorders and choice functions. *Econometrica* **1975**, *43*, 975–976.
20. Bordes, G. Some more results on consistency, rationality and collective choice. *Aggregation and revelation of preferences* **1979**, pp. 175–197.
21. Aizerman, M.A.; Aleskerov, F. *Theory of choice*; Vol. 38, North Holland, 1995.
22. Jamison, D.T.; Lau, L.J. Semiorders and the Theory of Choice. *Econometrica* **1973**, *41*, 901–912.
23. Aygün, O.; Sönmez, T. Matching with Contracts: Comment. *American Economic Review* **2013**, *103*, 2050–51.
24. Sen, A. Choice Functions and Revealed Preferences. *Review of Economic Studies* **1971**, *38*, 307–317.
25. Arrow, K. Rational Choice Functions and Orderings. *Economica* **1959**, *26*, 121–127.
26. Aleskerov, F.; Bouyssou, D.; Monjardet, B. *Utility maximization, choice and preference*; Vol. 16, Springer Science & Business Media, 2007.
27. Monjardet, B. Statement of precedence and a comment on IIA terminology. *Games and Economic Behavior* **2008**, *62*, 736–738.
28. Aizerman, M.; Malishevski, A. General theory of best variants choice: Some aspects. *IEEE Transactions on Automatic Control* **1981**, *26*, 1030–1040.
29. Tversky, A. Elimination by Aspects: A Theory of Choice. *Psychological Review* **1972**, *79*, 281.
30. Dulleck, U.; Hackl, F.; Weiss, B.; Winter-Ebmer, R. Buying online: An analysis of shopbot visitors. *German Economic Review* **2011**, *12*, 395–408.
31. Hotchkiss, G.; Jensen, S.; Jasra, M.; Wilson, D. The Role of Search in Business to Business Buying Decisions A Summary of Research Conducted. Enquiro White Paper.
32. Salant, Y.; Rubinstein, A. (A, f): Choice with Frames. *Review of Economic Studies* **2008**, *75*, 1287–1296.
33. Kopylov, I.; Yang, E. Revealed Delegation and Persuasion **2020**.
34. Ridout, S. A Model of Justification. *arXiv preprint arXiv:2003.06844* **2020**.