Revealed cognitive preference theory

Eric DANAN *
EUREQua, Université de Paris 1

December 15, 2003

Abstract

Revealed preference theory is generally thought of as deriving an agent's *tastes* from her observed *choice behavior*. However, a careful analysis shows that the potential *incompleteness* of tastes has been overlooked, and consequently the derived preferences cannot sensibly receive such a cognitive interpretation. I propose a solution based on the natural link between tastes' incompleteness and the concept of *preference for flexibility*, which enables an interpretation of derived preferences as modelling an agent's tastes. Furthermore, these derived preferences exist under weaker axiom than those commonly assumed in revealed preference theory. It is thus possible to conduct *welfare analysis* in a theory that can be tested by means of *behavioral data*.

Keywords. Revealed preference theory, cognitive preferences, incomplete preferences, welfare, preference for flexibility.
JEL Classifications. D11, B41.

*Any comment is welcome at danan@univ-paris1.fr. Thanks to Peter Klibanoff and Jean-Marc Tallon for helpful comments and suggestions. Financial support from the French Ministry of Research and the Franco-American Commission for Educational Exchange, as well as hospitality of Northwestern University, are gratefully acknowledged.

1 Introduction

Revealed preference theory (henceforth RPT) was born with Samuelson's (1938) seminal paper. Unlike the traditional approach to consumer theory taking as primitive an ordinal utility function (e.g. Pareto 1906), he directly started from a demand function. Thereby, he based the theory "upon those elements which must be taken as *data* by economic science" (Samuelson 1938, p71), i.e. upon *choice behavior*, as opposed to utility judgements. Samuelson (1938) not only disposed of utility as a primitive of consumer theory, but completely removed the concept of utility from the theory. As he noted himself (p71), such a purely behaviorist theory is silent about any notion of *tastes* or *welfare*.

Welfare economics is, however, a fundamental part of economic theory: most of economic analysis is not solely aimed at determining equilibrium behavior, but also seeks to assess the optimality of the identified equilibria, and to design policies or mechanisms that correct suboptimalities. Subsequent developments of RPT (e.g. Houthakker 1950) are generally thought of as grounding consumer theory on observed behavior without giving up welfare analysis, by *deriving* tastes from choice behavior.[1] Furthermore, RPT has disposed of the consumption setup to deal with arbitrary choice situations (e.g. Richter 1966, Sen 1971). Thus, Kreps (1988), introducing RPT in such a general framework, motivates it as follows.

> "we would like to start with [the concept of] choices made rather than preferences expressed. That is, from a descriptive point of view what we see is an individual's choice behavior – we have to connect this behavior at best we can with his preferences which are never directly expressed." (p11)

In this paper, I shall argue that interpreting preferences derived by RPT in terms of welfare is somewhat problematic, because this derivation overlooks an intrinsic feature of tastes: their potential *incompleteness*. I shall then propose a way of deriving an agent's tastes from her choice behavior which takes into account tastes' incompleteness.

The primitive concept of RPT is an agent's *choice function*, which describes her choice behavior in various choice situations. From this choice function are derived preferences, which express paired comparisons between alternatives. The

[1] On this historical evolution of revealed preference theory, see e.g. Mongin (2000).

derivation is achieved by means of the assumption that preferences *rationalize* choice behavior, i.e. that in any choice situation, the agent chooses an alternative if and only if she (weakly) prefers this alternatives to any available one.[2] From this assumption, it directly follows that an alternative a is preferred to an alternative a' if and only if a is chosen over a'. Therefore, preferences rationalizing choice behavior can naturally receive a *behavioral* interpretation, i.e. can be interpreted as modelling the agent's choice behavior in binary choice situations.

Thus, if one does not push their interpretation further, the derived preferences only enable to deduce the agent's behavior in arbitrary choice situations from her choice behavior in binary choice situations. However, as explained above, RPT seeks to derive preferences modelling the agent's tastes, and hence relevant for welfare analysis. Can this *cognitive* interpretation be applied to preferences rationalizing the agent's choice function? If so, then, necessarily, tastes and choice behavior always agree. That is to say, an agent likes a better than a' if an only if she chooses a over a'. This, however, need not be the case, as an agent may well be unable to determine which of two alternatives she desires the most, and still be forced to choose between them. In other words, the potential incompleteness of tastes undermines the validity of the rationalization assumption.

Consequently I shall weaken the rationalization condition by assuming preferences to be *consistent* with choice behavior: if the agent likes an alternative at least as much as any available one, then she chooses it, while if she likes an alternative strictly less than some available one, then she does not choose it. Consistency is equivalent to rationalization when preferences are complete, but is in general weaker, and the weakening allows for tastes' incompleteness: if the agent is unable to determine which of two alternatives she desires the most, then she can still choose between them without violating consistency.

On the other hand, consistency complicates the derivation of preferences from the agent's choice function. More precisely, while rationalizing preferences are straightforwardly *unique*, several preferences are always consistent with a given choice function. Intuitively, whenever the agent is observed choosing a over a', it is possible that she likes a better than a', or that she is unable to determine which these two alternatives she likes the most. To solve this uniqueness problem, I shall assume alternatives to be *opportunity sets* (or *menus*), and link tastes'

[2]Some authors (e.g. Kreps 1988) adopt a slightly different, but equivalent approach, see Section 3.

incompleteness to the concept of *preference for flexibility* introduced by Koopmans (1964). More precisely, I shall assume that the agent is unable to determine which of two alternatives she likes the most if and only if she would like to postpone her choice between them. This *learning-then-acting* assumption, which I have introduced and extensively discussed in Danan (2001), formalizes the intuitive justification for preference for flexibility in terms of "uncertainty about tastes" given by Koopmans (1964) and Kreps (1979). Consistency and learning-then-acting together yield uniqueness of derived preferences.

Besides uniqueness, an important question is that of *existence* of derived preferences.[3] Sen (1971) showed that rationalizing preferences exist if and only if the agent's choice function satisfies Axioms "α" and "γ". I shall identify axioms characterizing existence of consistent preferences satisfying the learning-then-acting property, and show that they are weaker than α and γ. Thus, cognitive/consistent preferences exist under weaker axioms than behavioral/rationalizing preferences. This result is interesting from a descriptive viewpoint, as the axioms of RPT are known to be violated by observed behavior.[4]

In Danan (2001), I have analyzed the derivation of cognitive preferences from choice behavior, taking as primitive concept that of behavioral preferences. This amounts to the special case of the present analysis in which the agent's choice function satisfies Axioms α and γ. The present results are more general, as they show that Axioms α and γ are not necessary for the existence of cognitive preferences. In the same paper, I have also taken into account another issue in the derivation of tastes from behavior, that of "unobservable indifference": if "choosing" means selecting one, and only one alternative, then indifference in tastes cannot be observed. This issue is not taken into account here: as in the standard approach, I allow the agent's choice function to select several alternatives at once, and consistency (as rationalization) is underlain by the assumption that if the agent likes two alternatives equally, then she selects them both. My conjecture is that this issue can be solved the same way it is solved in Danan (2001) (i.e. by adding monetary payoffs to the alternatives), but that this would complicate the analysis without bringing new insights, because existence of derived preferences would then be characterized by axioms that are neither weaker nor stronger than α and γ.

[3] RPT has focused on the existence issue, because rationalization straightforwardly yields uniqueness.

[4] See e.g. Sippel (1997) and the references therein.

Another related work is that of Eliaz and Ok (2003). This is, as far as I know, the only existing attempt to derive cognitive/incomplete preferences from a choice function. They consider a weakening of rationalization which is stronger than consistency and, in short, reveal tastes' incompleteness through intransitivities in choice behavior. This approach, however, is normatively questionable because it makes incomplete tastes irrational *per se* according to the "money pump" criterion,[5] a drawback from which the present approach is exempt (see Danan 2002). Moreover, Eliaz and Ok's (2003) derived preferences only exist under stronger axioms than the present model's (and even stronger than α and γ). I shall discuss their work in more detail in Section 4.

The paper is organized as follows. Section 2 introduces the formal setup. Section 3 briefly reviews Sen's (1971) analysis of the derivation of rationalizing/behavioral preferences, which serves as a benchmark for the derivation of cognitive preferences, presented in Section 4.

2 Setup

Given a set $\mathcal{S}$, $\#\mathcal{S}$ denotes the cardinality of $\mathcal{S}$, $2^{\mathcal{S}}$ the set of all subsets of $\mathcal{S}$, and $\mathcal{P}(\mathcal{S}) = 2^{\mathcal{S}} \setminus \emptyset$ the set of all nonempty subsets of $\mathcal{S}$. Given a set J, the notation $(S_j)_{j \in J} \in \mathcal{S}$ refers to a class of subsets of $\mathcal{S}$.

Let $\mathcal{A}$ be a nonempty set of mutually exclusive *alternatives*. An element A of $\mathcal{P}(\mathcal{A})$ is interpreted as a *choice situation*. The agent's observed choice behavior is modelled by a *choice function* C on $\mathcal{A}$, i.e. a function $C : \mathcal{P}(\mathcal{A}) \to 2^{\mathcal{A}}$ such that $\forall A \in \mathcal{P}(\mathcal{A})$, $C(A) \subseteq A$. Call *support* of C the set $\Sigma(C) = \{A \in \mathcal{P}(\mathcal{A}) : C(A) \neq \emptyset\}$. $A \in \Sigma(C)$ is interpreted as "the choice situation A is observed", and $a \in C(A)$ as "a is chosen out of A". Given a positive integer n, say that C is *n-decisive* if $\forall A \in \mathcal{P}(\mathcal{A})$, $\#A \leq n \Rightarrow A \in \Sigma(C)$. Define the choice function C^+ on $\mathcal{A}$ by $\forall A \in \mathcal{P}(\mathcal{A})$,

$$C^+(A) = A,$$

i.e. C^+ is the maximal (with respect to set inclusion) choice function on $\mathcal{A}$

Preferences over $\mathcal{A}$ are modelled by a binary relation $\succsim$ on $\mathcal{A}$ (i.e. a subset of $\mathcal{A} \times \mathcal{A}$), with $a \succsim a'$ being interpreted as "a is *weakly preferred* to a'". Given $\succsim$,

[5] See e.g. Davidson, McKinsey, and Suppes (1955).

define the binary relations $\succ$, $\sim$, and $\bowtie$ on $\mathcal{A}$ by $\forall a, a' \in \mathcal{A}$,

$$a \succ a' \Leftrightarrow a \succsim a' \not\succsim a, \qquad a \sim a' \Leftrightarrow a \succsim a' \succsim a, \qquad a \bowtie a' \Leftrightarrow a \not\succsim a' \not\succsim a.$$

$a \succ a'$ is interpreted as "a is *strictly preferred* to a'", $a \sim a'$ as "a and a' are *indifferent*", and $a \bowtie a'$ as "a and a' are *incomparable*". Say that $\succsim$ is

— *reflexive* if $\forall a \in \mathcal{A}$, $a \sim a$,

— *complete* if $\nexists a, a' \in \mathcal{A}$ such that $a \bowtie a'$,

— *transitive* if $\forall a, a', a'' \in \mathcal{A}$, $a \succsim a' \succsim a'' \Rightarrow a \succsim a''$.

Clearly, completeness implies reflexivity. Define the binary relations $\succsim^+$ and $\succsim^-$ on $\mathcal{A}$ by

$$\succsim^+ = \mathcal{A} \times \mathcal{A}, \qquad\qquad \succsim^- = \{(a,a) : a \in \mathcal{A}\},$$

i.e. $\succsim^+$ is the maximal (with respect to set inclusion) binary relation on $\mathcal{A}$ and $\succsim^-$ is the minimal reflexive binary relation on $\mathcal{A}$.

As explained in the introduction, there are two different kinds of preferences one might be interested in: *behavioral preferences* and *cognitive preferences*. Binary relations modelling them are respectively denoted by $\succsim_B$ and $\succsim_C$, with $a \succsim_B a'$ being interpreted as "the agent *chooses* a over a'" and $a \succsim_C a'$ as "the agent *desires* a at least as much as a'". Throughout the sequel, I shall restrict attention to complete behavioral preferences, and reflexive cognitive preferences, i.e. I shall assume that the agent can be forced to choose between any two alternatives and, although she may not cognitively rank any two alternatives, she desires any alternative exactly as much as itself. Hence the following definition.

Definition. *A* ***behavioral preference relation*** *is a complete binary relation. A* ***cognitive preference relation*** *is a reflexive binary relation.*

Finally, two particular choice functions, constructed from preferences, shall be of interest: given a binary relation $\succsim$ on $\mathcal{A}$, define the choice functions $\Gamma(., \succsim)$ and $\Delta(., \succsim)$ on $\mathcal{A}$ by $\forall A \in \mathcal{P}(\mathcal{A})$,

$$\Gamma(A, \succsim) = \{a \in A : \forall a' \in A,\ a \succsim a'\},$$
$$\Delta(A, \succsim) = \{a \in A : \forall a' \in A,\ a' \not\succ a\}.$$

$\Gamma(A, \succsim)$ and $\Delta(A, \succsim)$ are the sets of $\succsim$-*dominating* and $\succsim$-*undominated* alternatives in A, respectively. Note that $\Gamma(A, \succsim) \subseteq \Delta(A, \succsim)$, and that $\succsim$ is complete if and only if $\Gamma(., \succsim) = \Delta(., \succsim)$. Also, $\Delta(., \succsim)$ is 2-decisive for any binary relation $\succsim$, while $\Gamma(., \succsim)$ is 2-decisive if and only if $\succsim$ is complete.

3 Revealed behavioral preferences

In this section, I shall briefly review standard RPT, as presented by Sen (1971). This will serve as a benchmark for next section analysis. From an agent's choice function, Sen (1971) seeks to derive preferences satisfying the following condition.

Definition. *A binary relation $\succsim$ on $\mathcal{A}$* ***rationalizes*** *a choice function C on $\mathcal{A}$ if $\forall A \in \Sigma(C)$,*

$$C(A) = \Gamma(A, \succsim).$$

Rationalization asserts that an alternative a is chosen out of A if and only if a is weakly preferred to any alternative in A. Note that if C is 2-decisive, then it follows that $\succsim$ is complete and $\forall a, a' \in \mathcal{A}$, $a \succsim a' \Leftrightarrow a \in C(\{a, a'\})$. Thus, preferences rationalizing a 2-decisive choice function can naturally receive a behavioral interpretation, but not a cognitive one because of the potential incompleteness of tastes. An alternative definition of rationalization would be that $\forall A \in \Sigma(C)$,

$$C(A) = \Delta(A, \succsim). \tag{1}$$

I shall refer to (1) as Δ-*rationalization.* Since rationalization of a 2-decisive choice function implies completeness of $\succsim$, it also implies Δ-rationalization. The converse, however, does not hold, because Δ-rationalization does not force completeness of derived preferences. For example, one can check that both $\succsim^+$ and $\succsim^-$ Δ-rationalize C^+, but only $\succsim^+$ rationalizes it. As I shall discuss in the next section, Δ-rationalization has recently been used in an attempt to derive cognitive/incomplete preferences from a choice function. In standard revealed preference theory, Δ-rationalization has been used in an approach that makes it equivalent to rationalization.[6] This approach consists in taking strict preference (instead of weak preference) as primitive, and then defining weak preference by $a \succsim a' \Leftrightarrow a' \not\succ a$. Clearly, this makes $\succsim$ complete by definition, and hence Δ-rationalization boils down to rationalization.

[6] E.g. Kreps (1988).

Preferences rationalizing a 2-decisive choice function are obviously unique. Sen (1971) provides necessary and sufficient axioms on a choice function for the existence of preferences rationalizing it.

Axiom (α). $\forall A, A' \in \Sigma(C)$,

$$A \subseteq A' \Rightarrow A \cap C(A') \subseteq C(A).$$

Axiom (γ). $\forall (A_j)_{j\in J} \in \Sigma(C)$ *such that* $\bigcup_{j\in J} A_j \in \Sigma(C)$,

$$\bigcap\nolimits_{j\in J} C(A_j) \subseteq C(\bigcup\nolimits_{j\in J} A_j).$$

Axiom α asserts that if a is chosen out of A', then a is also chosen out of any subset of A' containing a. Axiom γ asserts that if a is chosen out of each A_j, for $j \in J$, then a is also chosen out of $\bigcup_{j\in J} A_j$. These axioms can be illustrated by an example taken from Sen (1969). Assume $\mathcal{A}$ is the set of (male) marathon runners in the world, and each choice situation A is a race between the runners in A, whose winner the agent has to bet on (all bets yield the same monetary prize $m > 0$ if won, and 0 if lost, say). Note that the fact that $C(A)$ might contain more than one element does not mean that the agent is allowed to simultaneously bet on several runners, but rather that there are several hypothetical betting behaviors she "might" adopt. According to Axiom α, if the agent bets on a Pakistani in the Asian championship, then she also bets on him in the Pakistani championship. According to Axiom γ, if she bets on the same French runner in both the European and the Mediterranean championship, then she also bets on him in the Euro-Mediterranean championship.[7]

In this example, existence of a binary relation $\succsim$ on $\mathcal{A}$ rationalizing C simply means that the agent bases her betting behavior in arbitrary races on her betting behavior in duels: in a race A, she bets on a runner a if and only if she would bet on a in any duel against an opponent in A. It is intuitive that such behavior satisfies Axioms α and γ. Formally, one can check that given any binary relation $\succsim$ on $\mathcal{A}$, $\Gamma(., \succsim)$ satisfies Axioms α and γ by definition.

Theorem 1. *Let C be a 2-decisive choice function on $\mathcal{A}$. Then C satisfies Axioms α and γ if and only if there exists a (unique) behavioral preference relation $\succsim_B$ on $\mathcal{A}$ rationalizing C.*

[7] I.e. the (virtual) race involving all European or Mediterranean runners.

Proof. See Sen's (1971) Theorem 9 for the existence part. Uniqueness is obvious by 2-decisiveness. □

4 Revealed cognitive preferences

I now turn to the problem of revealing an agent's cognitive preferences from her choice function. As explained above, cognitive preferences cannot be derived by means of the rationalization condition, because of their potential incompleteness. Still, it makes sense that $\Gamma(A, \succsim) \subseteq C(A)$, i.e. if the agent desires a at least as much as any alternative in A, then she chooses a out of A. What is not compelling any more is the converse inclusion. For example, it is plausible that $\succsim^-$ be an agent's cognitive preference relation while C^+ be her choice function. Thus, one has to weaken rationalization.

Eliaz and Ok (2003) use Δ-rationalization in order to derive cognitive preferences. As shown by the previous section's example, Δ-rationalization does not yield uniqueness of derived preferences. The way they achieve uniqueness essentially consists in assuming that $\succsim$ is transitive and that $\forall a, a' \in \mathcal{A}$,

$$a \bowtie a' \Rightarrow [\ \exists a'' \in \mathcal{A} \text{ such that } a \bowtie a'' \succ a'\ ]. \tag{2}$$

As one can check, (2) alone implies that $\succsim^+$ is the only binary relation Δ-rationalizing C^+ (one can easily construct examples in which transitivity is needed as well).

Unfortunately, (2) makes incomplete preferences irrational *per se*, in that an agent whose choice function is Δ-rationalizable by an incomplete binary relation is always vulnerable to money pumps (unless one deviates from the standard money pump argument by introducing a concept of *status quo*, e.g. Burros 1974). To illustrate this point, suppose that C is rationalizable by a behavioral preference relation $\succsim_B$ (this is always the case in Eliaz and Ok's (2003) model), and let $\succsim$ be an incomplete binary relation Δ-rationalizing C. Then by (2), $\exists a, a', a'' \in \mathcal{A}$ such that $a' \bowtie a \bowtie a'' \succ a'$, and hence by Δ-rationalization, $a' \sim_B a \sim_B a'' \succ_B a'$, so $\succsim_B$ is intransitive, i.e. vulnerable to money pumps.

Thus, the conjunction of Δ-rationalization and (2) forces an agent whose tastes are incomplete to adopt an irrational choice behavior. As for the other assumption in Eliaz and Ok's (2003) model (transitivity), Danan (2002) shows that it is sufficient, but not necessary, for the agent to have the possibility of adopting a rational behavior without "contradicting" her tastes. This implies that Eliaz and Ok's (2003) model is normatively questionable, as it forces agents who could behave rationally to behave irrationally. The transitivity assumption alone is problematic, since it has no normative justification, and even if it had, one would like, from a descriptive viewpoint, to test it, which is impossible if the derivation of cognitive preferences relies on it. Finally, Δ-rationalization is itself questionable. For example, the cognitive preference relation $\succsim^-$ models totally undeterminated tastes, and hence one would like that an agent having such tastes be free to adopt any choice behavior, yet the only choice function that is Δ-rationalized by $\succsim^-$ is C^+. Consequently, I shall neither assume (2) nor transitivity of derived preferences, and I shall weaken Δ-rationalization as follows.

Definition. *A binary relation $\succsim$ on $\mathcal{A}$ is **consistent** with a choice function C on $\mathcal{A}$ if $\forall A \in \Sigma(C)$,*

$$\Gamma(A, \succsim) \subseteq C(A) \subseteq \Delta(A, \succsim).$$

Consistency asserts that if a is weakly preferred to any alternative in A, then a is chosen out of A, while if some alternative in A is strictly preferred to a, then a is not chosen out of A. Clearly, this condition is weaker than Δ-rationalization, and hence than rationalization, and these three properties are equivalent if and only if $\succsim$ is complete. In the special case where C is 2-decisive and satisfies Axioms α and γ, consistency with C is equivalent to the following condition, which I introduced in Danan (2001) (under the name "strong consistency").

Definition. *A binary relation $\succsim$ on $\mathcal{A}$ is **consistent** with a behavioral preference relation $\succsim_B$ on $\mathcal{A}$ if $\succsim \subseteq \succsim_B$ and $\succ \subseteq \succ_B$.*

Lemma 1. *Let C be a choice function on $\mathcal{A}$, and $\succsim_B$ be a behavioral preference relation rationalizing C. Then a binary relation $\succsim$ on $\mathcal{A}$ is consistent with $\succsim_B$ if and only if it is consistent with C.*

Proof. Assume $\succsim$ is consistent with $\succsim_B$. Then $\forall A \in \mathcal{P}(\mathcal{A})$, $\Gamma(A, \succsim) \subseteq \Gamma(A, \succsim_B)$ and $\Delta(A, \succsim_B) \subseteq \Delta(A, \succsim)$. Hence $\succsim$ is consistent with C since $\forall A \in \Sigma(C)$, $C(A) = \Gamma(A, \succsim_B) = \Delta(A, \succsim_B)$.

Conversely, assume $\succsim$ is consistent with C. Then $\forall a, a' \in \mathcal{A}$, $a \succsim a' \Rightarrow a \in C(\{a, a'\}) \Leftrightarrow a \succsim_B a'$, and $a \succ a' \Rightarrow a' \notin C(\{a, a'\}) \Leftrightarrow a' \not\succsim_B a \Leftrightarrow a \succ_B a'$. $\square$

Clearly, consistency does not yield uniqueness of derived preferences, even in the special case of Lemma 1. For example, both $\succsim^+$ and $\succsim^-$ are consistent with

C^+. Intuitively, it is always possible that the agent's tastes agree with her choice behavior, or that they be undetermin. In order to achieve uniqueness, I shall make an assumption about the nature of alternatives. Namely, I shall assume that $\mathcal{A} = \mathcal{P}(\mathcal{X})$, for some set $\mathcal{X}$, where $x \in \mathcal{X}$ is interpreted as an *option*, and $X \in \mathcal{A}$ as an *opportunity set* (or *menu*).[8] That is to say, X is the commitment to choose an option $x \in X$ at some given later date. This is the standard structural framework for modelling the concept of *preference for flexibility*, introduced by Koopmans (1964). Note that this structural assumption is essentially unrestrictive, in the sense that any set $\mathcal{X}$ of alternatives can be extended to the set $\mathcal{P}(\mathcal{X})$ of menus.

Given a binary relation $\succsim$ on $\mathcal{P}(\mathcal{X})$, define the binary relation $\parallel$ on $\mathcal{P}(\mathcal{X})$ by $\forall X, X' \in \mathcal{P}(\mathcal{X})$,

$$X \parallel X' \Leftrightarrow [\ X \cup X' \succ X \text{ and } X \cup X' \succ X'\ ].$$

$X \parallel X'$ is interpreted as "the agent has a *preference for flexibility* at $\{X, X'\}$". Koopmans (1964) and Kreps (1979) justify preference for flexibility by "uncertainty about future tastes". In the present framework, this intuition is naturally modelled as follows.

Definition. *A cognitive preference relation $\succsim_C$ on $\mathcal{P}(\mathcal{X})$ satisfies the* ***learning-then-acting*** *property if $\bowtie_C = \parallel_C$.*

In Danan (2001), I provide an extensive discussion of the learning-then-acting property, and I show how this property yields uniqueness of cognitive preferences consistent with given behavioral preferences. Combined with Theorem 1 and Lemma 1, this result yields the following theorem.

Theorem 2. *Let C be a 2-decisive choice function on $\mathcal{P}(\mathcal{X})$. Then C satisfies Axioms α and γ if and only if there exists a (unique) behavioral preference relation $\succsim_B$ on $\mathcal{P}(\mathcal{X})$ rationalizing C and a (unique) cognitive preference relation $\succsim_C$ on $\mathcal{P}(\mathcal{X})$ consistent with C that satisfies the learning-then-acting property.*

Proof. By Theorem 2, C satisfies Axioms α and γ if and only if there exists a (unique) behavioral preference relation $\succsim_B$ on $\mathcal{P}(\mathcal{X})$ rationalizing C. Hence by Lemma 1, it is sufficient to show that there exists a (unique) cognitive preference

[8] As in Danan (2001), this structural assumption could be weakened by only assuming the existence of an "idempotent", "commutative", and "associative" operator on $\mathcal{A}$.

relation $\succsim_C$ on $\mathcal{P}(\mathcal{X})$ consistent with $\succsim_B$ that satisfies the learning-then-acting property. This directly follows from Danan's (2001) Theorems 1 and 2. $\square$

Thus, whenever a 2-decisive choice function satisfies Axioms α and γ (and alternatives are menus), one can not only derive behavioral preferences, but also cognitive preferences, from observed choice behavior. This enables to deal with both equilibrium and welfare analysis in a revealed preference framework.

Now, if one is interested in deriving cognitive preferences alone, then Axioms α and γ are no longer necessary. For example, let $\mathcal{X} = \{x_1, x_2, x_3\}$, and define the choice function C on $\mathcal{A} = \mathcal{P}(\mathcal{X})$ by

$$\begin{cases} C(\{\{x_1\},\{x_2\},\{x_3\}\}) = \{\{x_2\}\}, & \\ C(\{\{x_1\},\{x_j\}\}) = \{\{x_1\}\} & \text{for } j = 2,3, \\ C(A) = \{X \in A : \forall X' \in A,\ \#X \geq \#X'\} & \text{otherwise.} \end{cases}$$

Then C violates Axioms α (because $\{x_2\} \in C(\{\{x_1\},\{x_2\},\{x_3\}\})$ but $\{x_2\} \notin C(\{\{x_1\},\{x_2\}\})$ as well as Axiom γ (because $\{x_1\} \in \bigcap_{j\in\{2,3\}} C(\{\{x_1\},\{x_j\}\})$ but $\{x_1\} \notin C(\{\{x_1\},\{x_2\},\{x_3\}\})$), but one can check that the cognitive preference relation $\succsim_C$ on $\mathcal{A}$ defined by $\forall X, X' \in \mathcal{A}$,

$$X \succsim_C X' \Leftrightarrow X' \subseteq X$$

satisfies the learning-then-acting property and is consistent with C. Intuitively, the violations of Axioms α and γ only involve cognitively incomparable alternatives.

In order to characterize existence of derived cognitive preferences, define, given a choice function C on $A = P(\mathcal{X})$, the choice function $\Psi(.,C)$ on $\mathcal{P}(\mathcal{X})$ by $\forall A \in \mathcal{P}(\mathcal{A})$,

$$\begin{aligned} \Psi(A,C) = \{X \in A : \exists X' \in A \setminus \{X\} & \text{ such that} \\ & C(\{X, X\cup X'\}) = C(\{X', X\cup X'\}) = X \cup X'\}. \end{aligned}$$

$\Psi(A,C)$ is interpreted as the set of menus $X \in A$ such that, for some $X' \in A\setminus\{X\}$, the agent "chooses flexibility at $\{X, X'\}$". Looking ahead to the cognitive preferences to be derived, these X' shall be those which are cognitively incomparable to X.

Axiom (F-α). $\forall A, A' \in \Sigma(C)$,

$$A \subseteq A' \Rightarrow A \cap C(A') \subseteq C(A) \cup \Psi(A, C).$$

Axiom (F-γ). $\forall (A_j)_{j\in J} \in \Sigma(C)$ *such that* $\bigcup_{j\in J} A_j \in \Sigma(C)$,

$$\bigcap_{j\in J}(C(A_j) \setminus \Psi(A_j, C)) \subseteq C(\bigcup_{j\in J} A_j).$$

Axiom F-α asserts that if $X \in A \subseteq A'$ is chosen out of A' and the agent does not choose flexibility at $\{X, X'\}$ for any $X' \in A \setminus \{X\}$, then X is also chosen out of A. Axiom γ asserts that if X is chosen out of A_j and the agent does not choose flexibility at $\{X, X'\}$ for any $X' \in A_j \setminus \{X\}$, for $j \in J$, then X is also chosen out of $\bigcup_{j\in J} A_j$. Clearly, Axiom F-α is weaker than Axiom α, and Axiom F-γ is weaker than Axiom γ.

In the marathon example, opportunity sets can be incorporated as follows. All races are to be held at some given later date, and postponing betting means that the agent can wait until the moment just before the race to make her bet. Note that this does not mean that she will be able to simultaneously bet on several runners, but only that she has the possibility to wait until the last moment to make her decision. Also, she must not be able to gather objective information about the runners (such as their results in past races), for otherwise she may well choose flexibility without being uncertain about her tastes. Thus, postponing her decision only enables her to learn about her tastes. According to Axiom F-α, if the agent bets on a Pakistani x in the Asian championship, then either she also bets on him in the Pakistani championship, or there is another Pakistani x' such that she postpones her bet in the duel between x and x'. According to Axiom F-γ, if the agent bets on the same French runner x in both the European and the Mediterranean championship, and she does not postpone her bet in any duel between x and any other European or Mediterranean runner, then she also bets on x in the Euro-Mediterranean championship.

Theorem 3. *Let C be a 2-decisive choice function on $\mathcal{P}(\mathcal{X})$. Then C satisfies Axioms F-α and F-γ if and only if there exists a (unique) cognitive preference relation $\succsim_C$ on $\mathcal{P}(\mathcal{X})$ consistent with C that satisfies the learning-then-acting property.*

Proof. *Necessity.* Assume there exists a cognitive preference relation $\succsim_C$ on $\mathcal{A} = \mathcal{P}(\mathcal{X})$ consistent with C that satisfies the learning-then-acting property. Let

$A, A' \in \mathcal{A}$ such that $A \subseteq A'$ and $X \in A \cap C(A')$. Then $\forall X' \in A'$, $X' \not\succ_C X$, i.e. $X \succsim_C X'$ or $X \bowtie_C X'$. Hence $X \in C(A) \cup \Psi(A, C)$ since C is 2-decisive, so C satisfies Axiom F-α. Now let $(A_j)_{j \in J} \in \Sigma(C)$ such that $\bigcup_{j \in J} A_j \in \Sigma(C)$, and let $X \in \bigcap_{j \in J}(C(A_j) \setminus \Psi(A_j, C))$. Then $\forall X' \in \bigcup_{j \in J} A_j$, $X' \not\succ_C X$ and $X \not\bowtie_C X'$, i.e. $X \succsim_C X'$. Hence $X \in C(\bigcup_{j \in J} A_j)$, so C satisfies Axiom F-γ.

Sufficiency. Assume C satisfies Axioms F-α and F-γ. Define the binary relation$\succsim_B$ on $\mathcal{A} = \mathcal{P}(\mathcal{X})$ by $\forall X, X' \in \mathcal{P}(\mathcal{X})$,

$$X \succsim_B X' \Leftrightarrow X \in C(\{X, X'\}). \tag{3}$$

Then $\succsim_B$ is a behavioral preference relation since C is 2-decisive. Hence by Danan's (2001) Theorem 2, there exists a cognitive preference relation $\succsim_C$ on $\mathcal{P}(\mathcal{X})$ consistent with $\succsim_B$ that satisfies the learning-then-acting property. Let $A \in \mathcal{P}(\mathcal{A})$ and $X \in A$. If $\forall X' \in A$, $X \succsim_C X'$, then $X \in C(\{X, X'\}) \setminus \Psi(\{X, X'\}, C)$, and hence $X \in C(A)$ by Axiom F-γ. If $\exists X' \in A$ such that $X' \succ_C X$, then $X \notin C(\{X, X'\}) \cup \Psi(\{X, X'\}, C)$, and hence $X \notin C(A)$ by Axiom F-α.

Uniqueness. By Lemma 1, any cognitive preference relation that is consistent with C is also consistent with the behavioral preference relation $\succsim_B$ defined by (3) (this part of the lemma's proof does not make use of Axioms α and γ. Hence by Danan's (2001) Theorem 1, there exists at most one such cognitive preference relation satisfying the learning-then-acing property. $\square$

Thus, derived cognitive preferences exist under weaker axioms than those characterizing the existence of derived behavioral preferences.

5 Conclusion

In this paper, I have shown how cognitive preferences, modelling an agent's tastes, can be derived from her observed choice behavior. The difficulty lied in the potential incompleteness of tastes, a phenomenon which has been overlooked by revealed preference theory, but which causes discrepancies between tastes and choice behavior. The solution I have proposed is based on the natural link between tastes' incompleteness and preference for flexibility.

It is thus possible to conduct welfare analysis in a theory that can be tested by means of behavioral data. Furthermore, the model I have proposed relies on weaker axiom than those of standard revealed preference theory. As far as I know, only much stronger axioms, like the Weak Axiom of Revealed Preference, have been empirically tested (and mostly rejected). These axioms ensure not only the existence of complete preferences rationalizing choice behavior, but also some normative properties such as transitivity. From a purely descriptive viewpoint, it would therefore be natural to investigate how much observed behavior can be explained by the general model presented here.

References

BURROS, R. (1974): "Axiomatic analysis of non-transitivity of preference and of indifference," *Theory and Decision*, 5, 185–204.

DANAN, E. (2001): "A behavioral model of individual welfare," mimeo, Université de Paris 1.

——— (2002): "The cognitive implications of behavioral rationality," mimeo, Université de Paris 1.

DAVIDSON, D., J. MCKINSEY, AND P. SUPPES (1955): "Outlines of a formal theory of value, I," *Philosophy of Science*, 22(2), 140–160.

ELIAZ, K., AND E. OK (2003): "Indifference or indecisiveness: choice-theoretic foundations of incomplete preferences," mimeo, New York University.

HOUTHAKKER, H. (1950): "Revealed preference and the utility function," *Economica, New Series*, 17(66), 159–174.

KOOPMANS, C. (1964): "On the flexibility of future preferences," in *Human judgements and rationality*, ed. by M. Shelley, and G. Bryan. John Wiley and Sons.

KREPS, D. (1979): "A representation theorem for preference for flexibility," *Econometrica*, 47, 565–577.

——— (1988): *Notes on the theory of choice*. Westview Press.

MONGIN, P. (2000): "Les préférences révélées et la formation de la théorie du consommateur," *Revue Economique*, 51(5), 1125–1152.

PARETO, V. (1906): *Manuale di economia politica*. Società Editrice Libraria.

RICHTER, M. (1966): "Revealed preference theory," *Econometrica*, 34(3), 635–645.

SAMUELSON, P. (1938): "A note on the pure theory of consumer's behaviour," *Economica, New Series*, 5(17), 61–71.

SEN, A. (1969): "Quasi-transitivity, rational choice and collective decision," *Review of Economic Studies*, 36(3), 381–393.

——— (1971): "Choice functions and revealed preference," *Review of Economic Studies*, 38(3), 307–317.

SIPPEL, R. (1997): "An experiment on the pure theory of consumer's behaviour," *Economic Journal*, 107, 1431–1444.

SUZUMURA, K. (1976): "Remarks on the theory of collective choice," *Economica, New Series*, 43(172), 381–390.