Some Thoughts on the Principle of Revealed Preference[1]

Ariel Rubinstein

School of Economics, Tel Aviv University and

Department of Economics, New York University

and

Yuval Salant

Graduate School of Business, Stanford University

Economic theorists characterize an individual decision maker using three basic concepts:

(1) A collection of objects: The manner in which a decision maker perceives an object does not have to be objective. For example, one decision maker might think about a red triangle as a triangle while another might think about it as a red object.

(2) Mental preferences: These describe the mental attitude of an individual toward the objects. They can be defined in contexts which do not involve actual choice. In particular, preferences can describe tastes (such as a preference for one season over another) or can refer to situations which are only hypothetical (such as the possible courses of action available to an individual were he to become Emperor of Rome) or which the individual does not fully control (such as a game situation in which a player has preferences over the entire set of outcomes).

(3) Choice: It is customary to describe a choice situation using a set of objects the individual can choose from. A choice function spells out how the individual will respond to any choice situation he might face.

The standard economic approach assumes that a decision maker is rational in the sense that (i) in any choice situation within the domain of his choice function he objectively identifies the set of objects (ii) his choice function is consistent with maximization of some preference relation which we will refer to as the *behavioral preferences* and (iii) the behavioral preferences are identical to the mental preferences.

The Principle of Revealed Preference, as we understand it, is a methodological paradigm which follows the standard economic approach, whereby observed choices are used only to reveal the mental preferences of the individual over the set of objects as perceived by the modeler.

[1]We thank Douglas Gale, Andy Schotter and Rani Spiegler for most helpful comments.

In this short paper we wish to make three statements about the way that economists view this principle as a modeling guide.

Statement 1. There is no escape from including mental entities, such as the way in which an individual perceives the objects and his mental preferences, in economic models.

Statement 2. Economists should be also looking at models in which the observed choice leads to conclusions other than that the chosen element is always mentally preferred to the other elements in the set.

Statement 3. There is room for models in which the observable information about a choice situation is richer than just the set of available alternatives and the alternative chosen.

Before proceeding, we need to introduce some standard notation and definitions. Let X be a finite set of alternatives. A choice problem is a non-empty subset of X. Let D be the collection of all choice problems. A choice function c attaches to every choice problem $A \in D$ a single element $c(A) \in A$. A choice function informs us that the individual chooses the element $c(A)$ when facing the choice problem A. A choice correspondence C attaches to every $A \in D$ a non-empty subset $C(A) \subseteq A$. The interpretation of a choice correspondence is more subtle than that of a choice function. We follow the approach whereby $C(A)$ is the set of alternatives which are chosen from the choice problem A under certain additional circumstances which are not part of the model.

A choice function c satisfies the property of Independence of Irrelevant Alternatives (IIA) if $c(B) \in A \subseteq B$ implies that $c(A) = c(B)$. A choice correspondence C satisfies the Weak Axiom of Revealed Preference (WA) if $a, b \in A \cap B$, $a \in C(A)$ and $b \in C(B)$ imply that $a \in C(B)$. When $C(A)$ is always a singleton, WA is equivalent to IIA.

We say that a choice function c (correspondence C) is rationalizable if there exists a preference relation $\succsim$ such that $c(A)$ is the $\succsim$-maximal element in A ($C(A)$ is the set of all $\succsim$-maximal element in A) for every $A \in D$. A choice function (correspondence) is rationalizable if and only if it satisfies IIA (WA).

Statement 1. There is no escape from including mental entities, such as the way in which an individual perceives the objects and his mental preferences, in economic models.

If the individual in an economic model were treated as a robot who receives a description of a choice problem as input and produces a chosen element as output, then the assumption that his behavior is rationalizable would lack any mental meaning. It would be interpreted solely as a procedural property: the choices made by the individual are independent of the procedure he uses to make them. (When IIA is violated, the order in which the decision maker makes his choices becomes crucial in describing his behavior. For example, if an individual chooses a from $\{a, b, c\}$ and b from $\{a, b\}$, then his response to the task "choose from a, b and c" differs from his response to a two-stage task in which he first has the option of choosing c and, if he does not, he must then choose between a and b.)

However, as economists, we are interested not just in describing the behavior of individuals but also their well-being. When we analyze social mechanisms and make welfare statements we have in mind the individual's mental preferences, which reflect his well-being. We cannot find any a priori reason to assume that an individual's behavioral preferences, which describe his choices, fully represent or convey his mental preferences. On the contrary, there are reasons to assume that they don't.

First, there is often no objective specification of the outcome space. A decision maker may have in mind a description of the alternatives which differs from that of the modeler (see Rubinstein (1991)).

Example 1. Assume that a decision maker receives a pair of files of candidates A and B piled alphabetically and chooses A. An observer might conclude that the decision maker prefers A to B. However, assume that unlike the observer, the decision maker ignores the content of the files and pays attention only to the location of each file in the pile. He simply prefers the top location in the pile to the bottom location. In this case, the observer's interpretation that the individual has chosen the "best" candidate is incorrect.

Second, the decision maker might be operating in a very systematic way but not according to his mental preferences. Following is an extreme example (see Rubinstein (2006)):

Example 2. An individual has in mind a clear notion of utility which expresses his desires. Imagine that we are even able to measure his utility using an "ultimate happiness measure". However, the individual behaves in a way that is consistent with *minimizing* this measure of utility. This might be due to a mistake in his "operating system", due to some mental problem or simply because he applies a rule of thumb which has nothing to do with his mental preferences. Of course, in this case the individual's choice function is rationalizable, i.e. there exists a preference relation whose maximization describes the individual's behavior. However, this preference relation is clearly the opposite of the individual's mental preferences and it would be absurd to consider his behavioral preferences as an indication of his well-being.

The importance of referring to mental preferences is revealed when we consider the basic welfare concept of Pareto efficiency. Pareto efficiency is an intuitively appealing concept because everybody can be made better off by moving from a Pareto-dominated outcome to a Pareto-dominant one. However, this intuition is often based on viewing preferences as being mental.

One could argue that the meaning of a Pareto-inefficient outcome is that it is unstable even when defined with respect to behavioral preferences. According to this interpretation, an inefficient outcome is unstable since every individual will choose to support a move to the Pareto-dominant outcome. However, note that an individual's preference for a Pareto-dominant outcome over a Pareto-dominated one usually involves a change in the behavior of other individuals and therefore may not be observable in any choice situation. Thus, it must have an additional mental meaning.

Example 3: Consider a 2×2 coordination game with two actions $\{a, b\}$ available to each player. Assume that both players have the same mental preferences over the outcomes of the game: $(a, a) \succ (b, b) \succ (a, b) \sim (b, a)$. Thus, (a, a) and (b, b) are the two pure strategy equilibria of the game and (a, a) is Pareto-superior to (b, b). The rankings $(a, a) \succ (b, a)$ and $(b, b) \succ (a, b)$ are revealed by the actions of Player 1. However, the ranking between (a, a) and (b, b) is not revealed in any choice situation associated with the game since Player 1 does not control Player 2's actions. Thus, the statement "(b, b) is an undesirable equilibrium and (a, a) is a desirable one" is based on each player's preference for (a, a) over (b, b), a preference which is not revealed by the choices of the players.

Thus, even the basic welfare criterion of Pareto efficiency cannot be based solely on behavioral preferences without referring also to mental preferences.

Statement 2. Economists should be also looking at models in which the observed choice leads to conclusions other than that the chosen element is always mentally preferred to the other elements in the set.

Some choice procedures violate the weak axiom of revealed preference (or the IIA property) and thus are not consistent with maximizing a preference relation. In such cases, there is no basis to conclude from an observed choice that the chosen element is always preferred to the other elements in the set. Nevertheless, other conclusions about the properties of a choice procedure can be drawn. This is in fact the objective of axiomatic analysis of a choice procedure.

To demonstrate this point, consider the *Post-Dominance Rationality* (PDR) choice procedure discussed in Manzini and Mariotti (2007) and Rubinstein and Salant (2006b).[2] According to the PDR procedure, the decision maker first simplifies a given choice problem by eliminating any alternative which he feels is dominated in some sense by another alternative in the set. He then chooses the best alternative among those that remain. For example, consider an individual who chooses among hotel resorts in the following manner: He first eliminates any resort for which there is another with more stars and a lower per-night price. He then applies a complicated rule to choose from among the remaining resorts. Formally, the decision maker's choice procedure is characterized by two binary relations:

(i) A dominance relation R which is acyclic.

(ii) A post-dominance relation $\succ$ which is complete and transitive whenever restricted to sets of elements that do not dominate one another.

When facing a choice problem A, the decision maker first identifies the set of non-dominated elements according to R and then chooses the $\succ$-maximal element from among them.

Obviously, this choice procedure generates choices that may violate IIA. For example, let $X = \{a, b, c\}$, bRc and $a \succ b \succ c \succ a$. The PDR procedure based on these parameters violates IIA since a is chosen from $\{a, b, c\}$ but c is chosen from $\{a, c\}$.

The following behavioral property characterizes a choice function c induced by a PDR procedure: If adding an element a to a choice problem A implies that neither the previously chosen element $c(A)$ nor the new element a is chosen from the new set, then $c(A)$ is never chosen from a choice problem that includes a. Formally, a choice function c satisfies *Exclusion Consistency* if for every set A and for every $a \in X$, if $c(A \cup \{a\}) \notin \{c(A), a\}$ then there is no set A' which contains a such that $c(A') = c(A)$.

It is straightforward to verify that a PDR choice procedure induces a choice function that satisfies Exclusion Consistency. Indeed, consider a PDR choice procedure based on a dominance relation R and a post-dominance relation $\succ$. Then the chosen element from a set A is the $\succ$-maximal element among the R-maximal elements in A. We need to show that the induced choice function c satisfies Exclusion Consistency. Assume that the element a is chosen from the choice problem A and that the element $a' \notin \{a, b\}$ is chosen from $A \cup \{b\}$. It must be that bRa. Otherwise, the element a continues to be non-dominated in $A \cup \{b\}$ and the only (possibly) new non-dominated element is b,

[2] Manzini and Mariotti (2007) and Rubinstein and Salant (2006b) differ in the details of the axiomatization of the PDR choice procedure. We follow Rubinstein and Salant (2006b) here. For additional interesting examples, see Masatlioglu and Ok (2005, 2006).

which means that either a or b is chosen from $A \cup \{b\}$. By the definition of PDR, since b dominates a, the element a is never chosen from sets in which b appears.

One can also show that a choice function that satisfies Exclusion Consistency can be represented as a PDR choice procedure. The proof of this statement is important to our argument since it contains a construction of a dominance relation and a post-dominance relation based only on the choices of the individual. Thus, assume c satisfies Exclusion Consistency. We define the two binary relations R and $\succ$ as follows:

(i) aRb if there is a set A such that $c(A) = b$ and $c(A \cup \{a\}) \notin \{a, b\}$.

(ii) $a \succ b$ if $c(\{a, b\}) = a$.

The relation R is acyclic. If there were a cycle then by Exclusion Consistency no element could be chosen from the set of all elements in the cycle. The relation $\succ$ is asymmetric and complete. The relation $\succ$ is transitive when restricted to sets of elements that are not related to one another by R. Otherwise, assume that $a \succ b$, $b \succ c$ and $c \succ a$ and that a, b and c are not related by R. Without loss of generality, assume that $c(\{a, b, c\}) = b$. Then, since $c(\{a, b\}) = a$ we should have cRa which is a contradiction.

Since R is acyclic and $\succ$ is complete and transitive when restricted to sets of elements that do not dominate one another, the PDR procedure based on R and $\succ$ chooses exactly one element from every set A. It is not difficult to complete the proof and show that the element chosen by the procedure is identical to $c(A)$.

To conclude, an essential component of the principle of revealed preference is that one should be able to deduce the parameters of the choice procedure from behavior. With the rational man's choice procedure in mind, we elicit a single preference relation from a choice function by the inference that choosing a when b is available means that a is at least as good as b. Analogously, with the PDR choice procedure in mind, we elicit a dominance relation and a post-dominance relation. Of course, different "deduction rules" should be applied to different choice procedures. But, nonetheless, economic analysis based on observables can accommodate choice procedures other than the rational man's, in which the parameters of the procedure are elicited from observable information as in the case of the rational man.

Statement 3. There is room for models in which the observable information about a choice situation is richer than just the set of available alternatives and the alternative chosen.

Classical choice theory usually assumes that a researcher observes a pair (A, a) with the interpretation that the decision maker chooses the alternative a from the choice set

A. However, in many cases, additional information relevant to choice is available in the same sense that the set of alternatives and the chosen alternative are available. Accepting the idea that the analysis of the decision maker's behavior should depend on observables implies that we should use a model of choice that takes this information into account rather than a model that ignores it.

Consider, for example, the model of order-dependent choice in which the alternatives are presented to the decision maker in the form of a list of distinct elements of X. It is actually quite common that a choice problem is presented as a *list* rather than as a *set*. For example, when purchasing a product online, the alternatives are positioned in some order or when looking for a job offers are received sequentially. A decision maker who uses a systematic method to choose from lists may choose differently from two different lists that induce the same set of alternatives.

In Rubinstein and Salant (2006a) we investigate some properties of choice functions from lists which assign a chosen element to every list. In particular, we analyze the following property:

Partition Independence (PI): Dividing a list arbitrarily into several sublists, choosing an element from each and then choosing from the list of chosen elements yields the same result as choosing from the original list.

PI is satisfied by the rational procedure as well as by the satisficing procedure (Simon (1955)). According to the satisficing procedure, the decision maker classifies each element as either satisfactory or non-satisfactory and chooses the first satisfactory element from each list (if no such element exists, we assume that he chooses the last element in the list). In fact, we show that PI characterizes a larger class of choice functions from lists. In this class, each function is parameterized by a preference relation $\succsim$ over X and a labeling of every $\succsim$-indifference set by "First" or "Last". Given a list, the decision maker first identifies the set of $\succsim$-maximal elements within that list. He then chooses the first or the last element among them according to the label of the $\succsim$-indifference set they belong to. For example, in the satisficing procedure, there are two indifference classes of the preference relation: the class of satisfactory elements labeled by First, and the class of non-satisfactory elements labeled by Last. The family of functions satisfying PI naturally generalizes the class of preference-maximizing procedures in the context of standard choice functions.

We then relate the notion of a choice function from lists to the standard notion of a choice correspondence by assigning to every set all the elements chosen for *some* listing of that set. For example, a satisficing procedure induces a choice correspondence which chooses all the satisfactory elements from every set; if there are none, the correspondence chooses the entire set. We show that a choice function from lists satisfying PI induces a choice correspondence satisfying WA. Conversely, if a choice correspondence satisfies WA, it can be "explained" by a choice function from lists satisfying PI.

One might therefore argue that there is no need to study choice from lists since the outcome (in terms of choice correspondences) is indistinguishable from that of a correspondence satisfying WA. We would argue that this is not the case. The two terms are indistinguishable only if we choose to ignore the additional information which is often observable (especially when the list is generated by an exogenous mechanism, as in the case of entrees listed on a menu or products listed in a sales brochure). In such cases, the notion of a choice function from lists is typically richer than a standard choice correspondence and provides a more accurate description of behavior. So why should we ignore this additional information? As we remarked earlier, an essential component of the principle of revealed preference is that one should be able to deduce the parameters of the choice procedure based on behavior. But there is no reason to adopt a position which restricts the scope of the observable information to the set of alternatives and the actual choice.

References

Manzini, Paola and Marco Mariotti (2007), "Sequentially Rationalizable Choice", *American Economic Review*, forthcoming.

Masatlioglu, Yusufcan and Efe Ok (2005), "Rational Choice with Status Quo Bias", *Journal of Economic Theory*, 121, No. 1, 1-29.

Masatlioglu, Yusufcan and Efe Ok (2006), "Reference-Dependent Procedural Decision Making", mimeo.

Rubinstein, Ariel (1991), "Comments on the Interpretation of Game Theory", *Econometrica*, 59, No. 4, 909-924.

Rubinstein, Ariel (2006), "Comments on Behavioral Economics", forthcoming in *Advances in Economic Theory.*

Rubinstein, Ariel and Yuval Salant (2006a), "A Model of Choice from Lists", *Theoretical Economics*, 1, No. 1, 3-17.

Rubinstein, Ariel and Yuval Salant (2006b), "Two Comments on the Principle of Revealed Preferences", mimeo.

Simon, Herbert (1955), "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics*, 69, No. 1, 99-118.