

PHILOSOPHY OF ECONOMICS

REVEALED PREFERENCES?

REVEALED PREFERENCES For an economist, preferences are defined over sets \mathbf{X} of alternatives x_1, x_2, \dots, x_n by the following properties (where “ \succeq ” means “is weakly preferred to”, that is, “is strictly preferred to or is indifferent to”):

Reflexivity. For all x_i in \mathbf{X} , $x_i \succeq x_i$.

Completeness (or connectedness). For all x_i, x_j in \mathbf{X} , either $x_i \succeq x_j$, or $x_j \succeq x_i$, or both.

Transitivity. For all x_i, x_j, x_k in \mathbf{X} , if $x_i \succeq x_j$, and $x_j \succeq x_k$, then $x_i \succeq x_k$.

Continuity. For all x_j in \mathbf{X} , $\{x_i : x_i \succeq x_j\}$ and $\{x_i : x_i \preceq x_j\}$ are closed sets.¹

But how are preferences interpreted? Many philosophers of economics would say that they are *causes* of one’s choices, or *reasons* that justify one’s choices (cf. Reiss, 2013, 34). Interpretations such as these are possible if preferences (in the economist’s sense) are *not identical* to choices. This seemingly uncontroversial view is, however, disputed by those who would like to “free” demand theory from the concept of preference.

The ambition of the so-called revealed preference theory is to replace reference to unobservable mental entities with reference to manifest behaviour in order to make demand theory “scientifically more respectable” (cf. Sen, 1973, 242). The theory is driven by the following intuition:

If a collection of goods y could have been bought by a certain individual within his budget when he in fact was observed to buy another collection x , it is to be presumed that he has revealed a preference for x over y . (Sen, 1973, 241)

In Samuelson’s formulation, the theory was based on what has later become known as Weak Axiom of Revealed Preference. Let R_D mean “is directly revealed to be preferred to”. Then:

Weak Axiom of Revealed Preference. If $x_i R_D x_j$, then it is not the case that $x_j R_D x_i$.

Revealed preference theory is, to date, widely endorsed. Here is an influential statement:

Standard economics focuses on revealed preference because economic data come in this form. Economic data can—at best—reveal what the agent wants (or has chosen) in a particular situation. Such data do not enable the economist to distinguish between what the agent intended to choose and what she ended up choosing; what she chose and what she ought to have chosen. [...] If an economist proposes a new theory based on nonchoice evidence, then either the new theory leads to novel behavioral predictions, in which case it can be tested with revealed preference evidence, or it does not, in which case the modification is vacuous. In standard economics, the testable implications of a theory are its content; once they are identified, the nonchoice evidence that motivated a novel theory becomes irrelevant. (Gul and Pesendorfer, 2008, 8)

In his (1973) article “Behaviour and the concept of preference”, Sen argues against eliminating preferences in favour of revealed preferences and, more generally, against conflating preferences with revealed preferences outside demand theory, for instance, in normative economics.

PREFERENCES \neq CHOICES The economist Little says that “if an individual’s behaviour is consistent, then it must be possible to explain that behaviour without reference to anything other than behaviour” (242), that is, without reference to *underlying* preferences. Samuelson says that, by his market behaviour, the individual reveals his preference pattern “if there is such a consistent pattern” (241). Sen objects that interpreting the Weak Axiom as a requirement of consistency presupposes that, which revealed preference theorists want to avoid, namely (stable)

¹Continuity is necessary for preferences to be representable by a continuous utility function.

preferences underlying behaviour, and not identical with/reducible to behaviour:

The alleged inconsistency between (i) choosing x when y is available and (ii) choosing y when x is available, would seem to have something to do with the surmise about the person's preference underlying his choices. Preferring x to y is inconsistent with preferring y to x , but if it is asserted that choice has nothing to do with preference, then choosing x rather than y in one case and y rather than x in another need not necessarily be at all inconsistent. What makes them look inconsistent is precisely the peep into the head of the consumer, the avoidance of which is alleged to be the aim of the revealed preference approach. (243)

Alternatively, to avoid to peep in one's head, and still maintain that preferences are choices, one might interpret the Weak Axiom not as an assumption of consistency but as an empirically verified hypothesis. However, as a hypothesis, Sen argues, the Weak Axiom is *unverifiable*.

To check whether the Weak Axiom holds for the entire field of all market choices, we have to observe the person's choices under infinitely many price-income configurations. In contrast, the number of actual choices that can be studied is extremely limited.² (243)

Moreover, comparisons have to be made within a fairly short time “to avoid taste change” (243). The idea is that different tastes entail different preference orderings, and render tests of inconsistency inconclusive. But, Sen observes, “the concept of taste change is itself a preference-based notion, and the whole framework of revealed preference analysis of behaviour is steeped with implicit ideas about preference and psychology”. Sen concludes that the claim of explaining behaviour without reference to underlying preferences is “pure rhetoric”. The reason why the Weak Axiom is assumed is not because it is verified—such that preferences are reduced to revealed preferences—but because it is plausible that manifest behaviour (revealed preferences) provides strong evidence for (unobservable) preferences (cf. 258).

²Notice that much research has been done on how to infer preferences from a finite number of observations, starting with (Afriat, 1967). One may, however, object that, if this research is aimed at showing that preferences may be *confirmed* by the evidence but *not verified* by it, it is not directly relevant to the original problem of verifying the Weak Axiom.

If revealed preference is interpreted in this light, namely in terms of evidence for underlying preference, and if the Weak Axiom is interpreted as an axiom in the light of which consumer's choices are analysed and interpreted, rather than as a verifiable hypothesis, some of the additional axioms of revealed preference theory, such as the Strong Axiom of Revealed Preference, become redundant. The Strong Axiom explicitly states that if $x_i R x_j$, then it is not the case that $x_j R x_i$, where R is the transitive closure of R_D : given observed bundles x_i, x_j, x_k , if $x_i R_D x_j$ and $x_j R_D x_k$, then $x_i R x_k$, whether this is directly observed or not. Sen reports that the motivation for the Strong Axiom is that the Weak Axiom is assumed to hold of *pairs* of choices, and not to imply transitivity—which applies to *triples* of alternatives—such that the transitivity of preferences could not be verified. However, Sen objects, the Weak Axiom does imply transitivity (245). So, he conjectures, the best explanation of why the Strong Axiom is required is that the Weak Axiom entails transitivity only over observed choices, which are finite in number. However, *an infinite number* of triples of alternatives are, given divisible commodities, available to an individual (relative to given market prices and a given budget constraint). Stronger axioms would then be required to avoid that “the man can get away satisfying the Weak Axiom over all the cases in which his behaviour can be observed in the market and nevertheless harbour an intransitive preference relation” (246), that is, *unverifiable* intransitive preferences. But since the Weak Axiom cannot be verified anyway (see above), stronger axioms remain unjustified.

UNDERDETERMINATION OF PREFERENCES
In addition, Sen offers deeper reasons for scepticism about the determination of preferences.

First, the theory of revealed preferences uses observation to infer that one prefers, or is at least indifferent to, the chosen alternatives. However,

sometimes one chooses without their choice reflecting preferences. For instance, one may *have to* choose. More generally, “behaviour may not be based on systematic comparison of alternatives” (258), such that the assumption of completeness may be violated:

If a person chooses x rather than y , it is presumed that he regards x to be at least as good as y , and not that may be he has no clue about what to choose and has chosen x because he had to choose something. (248)

However, it is difficult to use observation to distinguish incompleteness from indifference.

A second difficulty Sen regards as even more fundamental:

[...] interdependence of different people’s choices [...] discredits individualistic rational calculus. Illustrated in terms of the game of the prisoners’ dilemma: even in the absence of a contract, the parties involved will be together better off following rules of behaviour that require abstention from the rational calculus which is precisely the basis of the revealed preference theory. People may be induced by social codes of behaviour to act as if they have different preferences from what they really have. This type of departure may also be stable for those codes since such behaviour will justify itself in terms of results from the point of view of the group as a whole. (258)

Let us assume that the payoff matrix for the two prisoners in the dilemma is

	Confess (2nd)	Not Confess (2nd)
Confess (1st)	-10, -10	0, -20
Not Confess (1st)	-20, 0	-2, -2

where the first number in each pair in the matrix represents the utility of the first prisoner, and the second number represents the utility of the second prisoner.

As the individual utilities indicate, each prisoner is better off by confessing, given the other’s choice. If the other prisoner does not confess, confession makes one go free. If the other prisoner confesses, confession gives one a reduced sentence. But suppose each prisoner acts not

based on this rational self-centred calculation but based on the norm of not letting the other person down irrespective of the consequences for himself. Then, neither prisoner will confess and they will both get a very mild sentence.

The problem for revealed preference theory is that, given such choices, the theory would infer to the wrong preferences:

[...] if there is anything in the assumption of revealed preference as it stands, it must be presumed that each prisoner prefers at least one of the possible outcomes resulting from his non-confession to what would have happened had he confessed, given other things. That is, either he prefers the consequence of his not confessing given the other prisoner’s non-confession, or the consequence of his not confessing given the other prisoner’s confession. But in fact neither happens to be true. The prisoner does not prefer to go to prison for twenty years rather than for ten; nor does he prefer a sentence of two years to being free. His choice has not revealed his preference in the manner postulated. (251)

In other words:

The behaviour pattern that will make each better off in terms of their real preferences is not at all the behaviour pattern that will *reveal* those real preferences. (252)

Notice that it is assumed that the norm does not change one’s preferences (as for instance would the concern for the other, if it were something that affects one’s utility). Rather, it interferes with the choice without altering the preferences.

Each is assumed to be self-centred and interested basically only in his own prison term, and the choice of non-confession follows *not* from calculations based on this welfare function, but from following a moral code of behaviour suspending the rational calculus. (251)

REVEALED PREFERENCE \neq WELFARE Individual preferences often enter economic analysis not only as determinants of behaviour but also as basis of welfare judgements. However, Sen notes, the two notions may not coincide:

Preference can be defined in such a way as to preserve its correspondence with choice, or defined so as to keep it in line with welfare as seen by the person in question, but it is not in general possible to guarantee both simultaneously. (259)

In particular, the non-identity between revealed and true preferences makes it difficult to justify the use of revealed (or “as if”) preferences as a ground for normative decisions:³

The difficulty arises in interpreting preference [as understood by revealed preference theory] with the property that if a person prefers x to y then he must regard himself to be better off with x than with y . (253)

People’s behaviour may still correspond to some consistent *as if* preference but the numerical representation of the *as if* preference cannot be interpreted as individual welfare. (254)

For instance, consider cases of free riding, such as non-returning bottles for recycling, trusting that most will return the bottles. Assume all have the following ordering of preferences: others return bottles but not me \succ all return bottles \succ no one does \succ I alone do. Then, no one will return bottles and yet all would have preferred that all did. Now, assume people have been persuaded that non-return is immoral, but still have the same view of their welfare, which generates a Prisoner’s Dilemma situation. That is, assume one could act “wrongly” and feel better off nonetheless, taking all factors into account, including the welfare of others (i.e., one’s utility is higher when non-returning), but the norm of acting morally (which is not, by assumption, a factor influencing one’s preference ordering) prevents one from doing as one pleases. Then, welfare functions/preference relations don’t change but behaviour does. In such a case, the theory of revealed preference will fail to reveal the “true” preferences. Evidence of revealed preferences cannot thus be safely used to devise policies that are meant to increase welfare. In Sen’s words:

There is a distinction from the point of view of social judgment between the relevance of a choice made under a moral sense of social responsibility and that made under a straightforward pursuit of one’s welfare (including any pleasure one takes in the happiness of others). The identification of welfare with as if preferences

blurs this distinction and withholds relevant information from the analysis of social welfare and collective choice. (256)

In particular, many policies rely on normative considerations such as Pareto optimality. Pareto optimality is defined as a state of resource allocation, such that it is impossible to reallocate resources so as to make any individual better off without making at least some individual worse off. If revealed preferences differ from true preferences, and have no straightforward bearing on what makes one better off, normative uses of Pareto optimality based on revealed preferences will be unjustified.

Notice that Sen identifies the “normal” notion of preference with self-centred and interested evaluation—that, which makes one better off (cf. 256). Hausman, however, disputes this interpretation:

Sen is mistaken to suggest that “the normal use of the word permits the identification of preference with the concept of being better off.” Expected advantage cannot be what people mean by preference, because there is no contradiction in maintaining that people’s preferences depend on many things that people do not expect to bear on their own well-being. (2012, 21)

Hausman views the normal notion of preference as “overall” comparative evaluation, namely evaluation of alternatives, which are interfered with, or constrained by, rules, norms, etc. He contrasts this notion with what he takes to be the economist’s notion, namely “total” comparative evaluation, that is, the evaluation of alternatives where all considerations that bear on these alternatives—including rules, norms, etc.—are factored in as influences on the preference ordering itself (see Hausman, 2012, 3-4).

But arguably, even under Hausman’s interpretation, Sen’s general point stands: it is difficult to use observation to distinguish between total and overall comparative evaluations; hence, revealed preferences, which are informative about

³Original proponents of the theory, such as Samuelson, would sharply distinguish positive from normative issues. However, it is not uncommon to find economists that use revealed preferences to justify policy issues. For instance, Gul and Pesendorfer (2008) state: “Economists use the revealed preference of individuals as a welfare criterion because it is the only criterion that can be integrated with positive economic analysis” (25).

the former but not about the latter, may not warrant justified welfare policies.

References

- Afriat, S. (1967). The construction of a utility function from expenditure data. *International Economic Review*, 8:67–77.
- Gul, F. and Pesendorfer, W. (2008). The case for mindless economics. In Caplin, A. and Schotter, A., editors, *The Foundations of Positive and Normative Economics*, pages 3–39. New York: Oxford University Press.
- Hausman, D. M. (2012). *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Reiss, J. (2013). *Philosophy of Economics. A Contemporary Introduction*. New York: Routledge.
- Sen, A. (1973). Behaviour and the concept of preference. *Economica*, 40(159):241–59.