

## Working paper n° 2-2015

### Sen's Critique of Revealed Preference Theory and Its "Neo-Samuelsonian" Critique: A Methodological and Theoretical Assessment

Cyril Hédoïn\*

\* Professeur des Universités en sciences économiques,  
REGARDS (Université de Reims Champagne-Ardenne)

#### Abstract

This paper evaluates how Amartya Sen's critique of revealed preference theory stands against the latter's contemporary, "neo-Samuelsonian" version. Neo-Samuelsonians such as Dowding (2002) or Ross [(2005); (2014)] have argued that Sen's arguments against revealed preference theory are innocuous, in particular once it is acknowledged that revealed preference theory does not assume away the existence of motivations or other latent psychological or cognitive processes. Sen's claims that preferences and choices need to be distinguished and that external factors need to be taken into account to analyze the act of choice then appear to be irrelevant. However, while it is true that contemporary revealed preference theory partially answers Sen's critique, I show that the latter is still relevant outside the restricted areas of consumer choice and markets dynamics. In particular, Sen's views regarding the importance of incomplete preferences and the multiplicity of levels of agency can hardly be integrated into the framework of contemporary revealed preference theory. This is a significant limit, given the imperialistic claims of some of the proponents of the latter.

**Mots clés :** Amartya Sen, revealed preference theory, neo-Samuelsonian economics, rationality, choice theory

*Les working papers d'économie et gestion du laboratoire Regards sont édités après présentation en séminaire et validation par deux relecteurs internes, sous la responsabilité du conseil de laboratoire.*



Laboratoire d'Économie et  
Gestion REGARDS (EA 6292)  
Université de Reims  
Champagne-Ardenne  
UFR de sciences économiques,  
sociales et de gestion  
57B Rue Pierre Taittinger  
51096 Reims

Directeur : **Martino Nieddu**

Responsable de l'édition des  
working papers : **Cyril Hédoïn**

# Sen's Critique of Revealed Preference Theory and Its "Neo-Samuelsonian" Critique: A Methodological and Theoretical Assessment

Cyril Hédoïn\*

*University of Reims Champagne-Ardenne*

**Abstract:** This paper evaluates how Amartya Sen's critique of revealed preference theory stands against the latter's contemporary, "neo-Samuelsonian" version. Neo-Samuelsonians such as Dowding (2002) or Ross [(2005); (2014)] have argued that Sen's arguments against revealed preference theory are innocuous, in particular once it is acknowledged that revealed preference theory does not assume away the existence of motivations or other latent psychological or cognitive processes. Sen's claims that preferences and choices need to be distinguished and that external factors need to be taken into account to analyze the act of choice then appear to be irrelevant. However, while it is true that contemporary revealed preference theory partially answers Sen's critique, I show that the latter is still relevant outside the restricted areas of consumer choice and markets dynamics. In particular, Sen's views regarding the importance of incomplete preferences and the multiplicity of levels of agency can hardly be integrated into the framework of contemporary revealed preference theory. This is a significant limit, given the imperialistic claims of some of the proponents of the latter.

**Keywords:** Amartya Sen – Revealed preference theory – Neo-Samuelsonian economics – Rationality – Choice Theory

## 0. Introduction

Besides his major contributions to social choice theory and normative economics, the Indian economist Amartya Sen has developed a sophisticated critique of one of the major approach to study consumer choice in economics, revealed preference theory (RPT). Sen's critique goes back to his articles "Behaviour and the concept of preference" (Sen 1973) and "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory" (Sen 1977), and has been extended and refined throughout the years. It culminates in the articles "Internal Consistency of Choice" (Sen 1993) and "Maximization and the Act of Choice" (Sen 1997), as well as in his book *Rationality and Freedom*

---

\* Full Professor of Economics, REGARDS research center.

Email address: [cyril.hedoin@univ-reims.fr](mailto:cyril.hedoin@univ-reims.fr)

Paper presented at the "Historical, Methodological and Philosophical Perspectives on Microeconomics" workshop, Nice, 27/03/2015. I thank the participants and especially Dorian Jullien, for their remarks. The usual disclaimer applies.

(Sen 2002). Through this critique, Sen also targets the two dominant conceptions of rationality within economics: rationality as internal consistency of choice and rationality as maximization of self-interest or personal welfare.

Sen's critique was originally directed toward what can be called "traditional revealed preference theory" (TRPT),<sup>1</sup> *i.e.* RPT as it has been originally developed by Samuelson (1938), Little (1949) and Houthakker (1950) in the context of the rise of behaviorism in psychology and logical positivism in philosophy. Accordingly, Samuelson's purported goal in his 1938 article was to "develop the theory of consumer's behaviour freed from any vestigial traces of the utility concept" (Samuelson 1938, 71). The dominant understanding of Samuelson's and his followers' contributions to TRPT is that the main objective was to make consumer theory (and more generally, economic theory) totally independent from any psychological assumption and to give a pure behaviorist interpretation to standard concepts in consumer theory such as indifference curves or marginal rates of substitution. Whether or not this understanding is accurate,<sup>2</sup> it is hardly disputable that Sen's critique has been partly fueled by this behaviorist reading of Samuelson's RPT. However, both the scientific aim and the philosophical foundations of RPT have evolved through the decades since Samuelson's pioneering contribution. This evolution culminates with recent developments in what can be called *contemporary* RPT (CRPT) and more particularly with the "neo-Samuelsonian" defense of RPT and of choice theory offered by Don Ross [(2005); (2011); (2014)]. Even though Sen has refined his critique of RPT throughout the years, it has been argued by Ross (2005) and others (*e.g.* (Dowding 2002); (Hands 2013)) that it is irrelevant against CRPT. My aim in this paper is to evaluate this claim.

I build on a basic postulate according to which the relevance of Sen's critique cannot be properly evaluated without taking into account how it is related with his general philosophy of economics and more particularly with his views on social choice theory and normative economics. The main idea I defend here is that while it is true that CRPT partially answers Sen's critique, the latter is still relevant outside the restricted areas of consumer choice and markets dynamics. In particular, Sen's views regarding the importance of incomplete preferences and the multiplicity of levels of agency can hardly be integrated into the framework of CRPT. This is a significant limit, given the imperialistic claims of some of the proponents of the latter. As a result, a foundational issue regarding the relevance of Sen's critique against CRPT is the *scope* of economics: if one arbitrarily restricts the scope of economics to consumer choice and markets dynamics, then CRPT evades most if not all of Sen's arguments. But since such a restriction seems hardly defensible, Sen's critique is still relevant.

The article is divided into seven sections. The first section briefly presents the basic tenets of RPT and characterizes the distinction between TRPT and CRPT. The second and third section present Sen's critique against RPT by focusing on two key ideas: first, the concept of commitment and the related idea of counter-preferential choice; second, the idea that internal consistency of choice cannot

---

<sup>1</sup> I am here following the classification proposed by Hands (2013) who also distinguishes an "empirical" version and a "contemporary" versions of RPT.

<sup>2</sup> Compare Giocoli (2003) and Ross (2014) on this issue.

properly capture the concept of rationality. The fourth section develops the neo-Samuelsonian's critique of Sen's critique. The fifth and sixth sections argue that CRPT and the neo-Samuelsonian approach cannot adequately deal with two key issues in Sen's philosophy of economics: the relationship of rationality with respectively incomplete preferences and levels of agency. A last section briefly concludes.

## 1. Revealed Preference Theory: Traditional and Contemporary

Samuelson's article "A note on the Pure Theory of Consumer's Behaviour" (1938) was not only a formal attempt to reconstruct consumer choice theory without any reference to psychologically-connoted concepts such as preferences or utility. It also opened the way to a whole research program for the development of new tools for investigating theoretically and empirically the behavior of consumers. As a consequence, RPT refers to more than a theory but rather to a whole set of approaches sharing some basic assumptions (Hand 2013). It is therefore possible to distinguish, inside the research program associated to RPT, several kinds of approaches with different scientific goals. Following Hands (2013), I distinguish three broad approaches inside the RPT research program: the *traditional* RPT pioneered by Samuelson (1938) and which mainly pursue the mathematical objective of the reformulation of ordinal utility theory; the *empirical* RPT (ERPT) which is applicable to finite sets of choice data and is thus able to produce empirical results regarding consumer behavior; the *contemporary* RPT which, while sharing several characteristics with ERPT, pursues the more ambitious philosophical and theoretical aim of defining a whole scientific method of inquiry in economics. While all three approaches are somehow related, they also have their specificities which make necessary to consider them separately.

Consider TRPT first. As noted above, Samuelson's (1938, 71) explicit aim was to "develop the theory of consumer's behaviour freed from any vestigial traces of the utility concept". Similarly, Little (1949, 97) argued that "it must be possible to explain the behavior without reference to anything other than behavior". Actually, Samuelson's and Little's behaviorist statements fit uneasily with the term "revealed preference" since the concept of preference is entirely defined in terms of choice and has no independent (read: mental or psychological) meaning. A brief review of Samuelson's axiomatic reasoning makes it clear that the preference concept is totally defined on the basis of choice behavior. Samuelson started with a set of given and continuous demand functions  $f_i$  defined for  $n$  goods,  $i = 1, \dots, n$ . The demand function specifies the quantity  $x_i$  of good  $i$  consumed given the set of prices  $p_1, p_2, \dots, p_n$  and the available monetary income  $M$ . Thus,

$$(1) \quad x_i = f_i(p_1, p_2, \dots, p_n, M) \text{ for all } i = 1, \dots, n.$$

Samuelson further postulated that all the functions  $f_i$  are homogenous of degree zero, *i.e.* they are independent of the units in which prices are expressed:  $f_i(\lambda p, \lambda M) = f_i(p, M)$  for all  $\lambda > 0$ , and that the consumer's budget constraint is satisfied,  $\sum_i x_i p_i = M$ . On this basis, Samuelson defined a consistency condition which is nowadays known as the *weak axiom of revealed preference* (WARP):

for two bundles  $x^a$  and  $x^b$  and two sets of prices  $p^a$  and  $p^b$ , if the consumer chose bundle  $x^a$  at prices  $p^a$  while  $x^b$  was affordable, then consistency implies that  $x^b$  cannot be chosen at prices  $p^b$ , unless  $x^a$  is unaffordable at those prices. Formally,

$$(2) \quad \sum_i p_i^a x_i^a \geq \sum_i p_i^a x_i^b \Rightarrow \sum_i p_i^b x_i^a > \sum_i p_i^b x_i^b$$

WARP can also usefully be expressed more generally on the basis of a choice function  $C(\cdot)$  specifying for a non-empty set of alternatives  $X$  and a non-empty class  $K$  of subsets  $S$  of  $X$ , a non-empty subset  $C(S)$  of *chosen* alternatives. Here, an alternative  $x \in X$  may correspond to a bundle of goods or to more abstract entities such as a “state of the world”, *i.e.* a complete description of everything that is relevant for the choice maker. This approach (used for instance in social choice theory) will be useful here since proponents of RPT have recently argued that the latter’s scope of validity extends beyond the restricted area of consumer choice. In this case, WARP is formally defined as follows:

$$(3) \quad \text{WARP: } [x, y \in S \text{ and } x \in C(S)] \Rightarrow x \in C(S'), \text{ for any } S' \text{ with } x, y \in S' \text{ and } y \in C(S')$$

On this basis, a weak revealed preference relation  $R$  can be defined as

$$(4) \quad xRy \Leftrightarrow \exists S: [x \in C(S) \ \& \ y \in S]$$

The important point here is that the preference relation  $R$  has no independent meaning. It is defined by the choice function and thus by the consumer’s behavior and nothing else. To weakly prefer  $x$  over  $y$  means nothing but to choose  $x$  when  $y$  is available. WARP does not guarantee that the revealed preference relation  $R$  is transitive.<sup>3</sup> As a consequence, it also does not guarantee the existence of an underlying utility function representing the binary relation defined by  $R$ . However, as Samuelson himself recognized, this is of secondary importance since the point of his article was precisely to dispense with the utility concept.

A few years later, Houthakker (1950) extended WARP from pair-wise choices to chain of choices, thus establishing the strong axiom of revealed preference (SARP). Using once again the formulation in terms of choice functions (Hansson and Grune-Yanoff 2012):

$$(5) \quad \text{SARP: for all alternatives } x_1, x_2, \dots, x_n \text{ and subsets } S_1, S_2, \dots, S_n, \\ \text{if}$$

$$x_1, x_2, \dots, x_n \in S_1$$

---

<sup>3</sup> Though it guarantees that  $R$  is complete, asymmetric and acyclic. Moreover, it guarantees that the choice function  $C$  is binary, *i.e.* that it is generated by the binary relation  $R$  which itself is defined by  $C$  according to (4). As we will see below, the binariness (or “rationalizability”) of the choice function is a key feature in Sen’s critique of rationality as internal consistency of choice. Note also that the binariness condition is weaker than WARP.

$$x_2, \dots, x_n \in S_2,$$

...

$$x_{n-1}, x_n \in S_{n-1}$$

$$x_n \in S_n$$

and

$$x_1 \in C(S_1), x_2 \in C(S_2), \dots, x_n \in C(S_n)$$

then for all  $T$  with  $x_1, x_2, \dots, x_n \in T$  and all  $x_i \in C(T)$ , we must have

$$x_1, x_2, \dots, x_{i-1} \in C(T)$$

Obviously, SARP is stronger than WARP. Notably, it guarantees that the revealed preference relation  $R$  corresponding to (4) is complete, asymmetric and transitive. As a consequence, a preference relation based on a choice function satisfying SARP can be represented by a maximizing utility function unique up to a monotonic transformation. However, for choice functions specifying a non-empty subset  $C(S)$  for all subsets  $S$  of alternatives composed of at most three elements, SARP and WARP are actually equivalent (Sen 1971).

Beyond the technical details, what matters most is to acknowledge what was the scientific project pursued by Samuelson, Houthakker and others who have developed TRPT. Obviously, the construction of choice functions depends on the available data about consumers' choices. No less obvious is the fact that these data always correspond to *finite* sets of choices. However, as (1) demonstrates, the starting point of Samuelson was a set of demand functions specifying what would be chosen for any set of prices and any income level. In other words, Samuelson's (and by extension, Houthakker's) RPT starts with a completely defined choice function and asks what are the consistency conditions required to obtain a mathematical isomorphism with ordinal utility theory. As Ross (2014, 63) notes, "Samuelson's 1938 paper was an exercise in formal definition: it was proof that purely formal microeconomic theory as conceived in the ordinalist tradition is not an application of psychological concepts". For this reason, TRPT can also legitimately be called "mathematical" RPT because it consists to apply some consistency condition (WARP or SARP) to hypothetical patterns over infinite choice sets (Ross 2014, 71). However, a few years later, Afriat (1967) proposed a more empirically oriented application of RPT. Afriat's approach separates from TRPT on a key point: it asks how to construct a utility function consistent with the choices observed on the basis of a *finite* set of data regarding consumers' choices. As a consequence, Afriat's approach is a truly constructive one and therefore also a much more suitable one for empirical studies (Varian 2006).

Afriat used a different consistency condition, nowadays known as the generalized axiom of revealed preference (GARP). Relatively to WARP, GARP adds two conditions regarding the monotonicity and the convexity of the revealed preference relation (Ross 2014, 66). Though it is slightly weaker

than SARP,<sup>4</sup> Afriat's consistency condition guarantees that it is possible to construct *some* utility function representing the preference relation revealed by the choice function. Thanks to Afriat's approach, empirical economists working with finite sets of data know that they can construct an utility function, as long as the consistency condition defined by GARP holds. As it is recognized both by Hands (2013, 1085) and Ross (2014, 66-72), the contemporary developments of RPT (what we call CRPT) can be seen as a partial extension of ERPT. More precisely, while CRPT makes use of the same consistency axioms and modeling tools than ERPT, it applies them to a much wider scope and grounds them on a broader philosophical framework. Hands (2013, 1087) characterizes it in the following way:<sup>5</sup>

- i) CRPT uses the consistency axioms to infer behavioral patterns on the basis of *finite* sets of choice data.
- ii) CRPT defines preferences solely in terms of choice and claims that it offers no causal explanation of the behavior of agent behavior on the basis of his preferences.
- iii) CRPT defends a kind of methodological imperialism, since it is claimed to offer a general framework for choice theory in economics.

Feature (i) explains why CRPT can be seen as an extension of ERPT: it is “a framework to make scientific inferences: given a consistent pattern identified within the initial choice data, it makes an inference about the new choices that would be associated with different prices and income” (Hands 2013, 1088). The point is significant because, according to Ross (2014, 71, emphasis in original), “Human behavior, in or out of markets, should *not* be expected to conform to [mathematical RPT]”. People's behavior can be expected to conform to consistency axioms such as GARP only in specific institutional environments and within a temporally limited scope. Feature (ii) refers to what Binmore (2009, 19) calls the “causal utility fallacy”: “In revealed-preference theory, it isn't true that Pandora chooses *b* rather than *a* because the utility of *b* exceeds the utility of *a*. This is the Causal Utility Fallacy. It isn't even true that Pandora chooses *b* rather than *a* because she prefers *b* to *a*. On the contrary, it is because Pandora chooses *b* rather than *a* that we say that Pandora prefers *b* to *a*, and assign *b* a larger utility” (emphasis in original). This seems to be a verbal statement of expression (4) which indicates that a preference of *b* over *a* consists in choosing *b* when *a* is available. However, there is a key difference: while (4) is a *definition* and makes no claim about the explanatory significance of RPT, (ii) is a proposition about the *explanatory* significance of the formal structure of RPT. To say things otherwise, while TRPT (through expression (4)) does not make any claim on whether and how agent behavior can be

---

<sup>4</sup> GARP guarantees that the revealed preference relation *R* is complete and transitive. However, it does not guarantee asymmetry of *R*, meaning that we can have both  $xRy$  and  $yRx$ . Obviously, SARP implies both GARP and WARP.

<sup>5</sup> Hands' characterization is based on a short list of selected writings: Binmore (2009), Bernheim and Rangel (2008) and Gul and Pesendorfer (2008). A fourth major source of CRPT is Ross [(2005); (2014)] but as Ross (2014, 200-201) recognizes himself, his “neo-Samuelsonian” defense of RPT has several distinguishing features. In particular, Ross takes some distance with Gul and Pesendorfer's methodological manifesto (Ross 2011), even though he adheres to their general overview on what economics is fundamentally about. I return on the specificities of the neo-Samuelsonian defense of RPT in the fourth section.

explained, Binmore's causal utility fallacy asserts that RPT *is not* about explaining agent behavior. As Binmore (2009, 20) notes, while the theory is no longer explanatory but mainly descriptive, "it is hard to criticize because it has little substantive content".

Feature (iii) is the most significant one. Proponents of CRPT have made multiple rhetorical claims that the latter is nothing but the core of the methodological approach constitutive of standard economics.<sup>6</sup> Actually, these are overstatements because CRPT can only reasonably pretend to be constitutive of *choice theory* in economics, but not of economics as a whole. Still, as I shall argue in sections 5 and 6, in view of Sen's critique of RPT, this claim seems hardly defensible even for the limited scope of choice theory. The point however here is to note that to its proponents, the methodological imperialism of CRPT is either a fact about what economics (or choice theory) is actually about (*e.g.* Gul and Pesendorfer 2008) or a normative claim about what economists should (or should not) do (*e.g.* Binmore's point about the causal utility fallacy). In any case, it refers to a methodological statement about the scope of economics and about the scientific inquiry should proceed within this scope.

## 2. Sen's Critique of RPT (I): Commitment and Counter-Preferential Choices

Sen started his attack on RPT with a series of papers published during the 1970s [(Sen 1973); (Sen 1977)]. RPT is grounded on the behaviorist postulate equating actual and observable choices to preferences: a person who chooses commodity *x* rather than commodity *y* is deemed to *prefer* commodity *x* to commodity *y*. Sen argues against this conflation and claims that choices made by rational agents may fail to reflect their personal welfare as represented by their preferences. He captures the idea of "counter-preferential choice" through the concept of *commitment*. In his article "Behaviour and the Concept of Preferences" (Sen 1973, 253), Sen writes: "I would argue that the philosophy of the revealed preference approach essentially underestimates the fact that man is a social animal and his choices are not rigidly bound to his own preferences only". The distinction between preference and choice is thus at the core of Sen's early critique of RPT.

Sen's argument for the disentanglement of preference and choice takes two forms. In his 1973 article, Sen essentially argues against the idea that preferences can be entirely defined without a "peep into the head of the consumer" (Sen 1973, 243), *i.e.* in behaviorist terms. Referring explicitly to Samuelson's RPT, Sen (1973, 243) remarks that since "the number of actual choices that can be studied is extremely limited", the "[f]aith in the axioms of revealed preference arises, therefore, not from empirical verification, but from the intuitive reasonableness of these axioms interpreted precisely in terms of preference". As a consequence, "the whole framework of revealed preference analysis of behaviour is steeped with implicit ideas about preference and

---

<sup>6</sup> See the quotations in Hands (2013, 1087). Gul and Pesendorfer's (2008) essay is probably the most significant attempt to make CRPT looking like the only "authentic" methodological and theoretical approach, not only in choice theory but in economics as a whole.

psychology”. Actually, Sen’s point is that the various consistency axioms of TRPT are conceptually meaningless unless preferences are understood as *mental states* referring to an evaluation of personal welfare. The fact that choice data are necessarily finite strengthen the point: we cannot possibly know if people are *actually* consistent; therefore the appealing character of the consistency axioms does not come from the fact they are empirically verified but rather because they are related to a reasonable concept of preference *psychologically understood*.<sup>7</sup>

Sen’s second point, developed both in his 1973 and 1977 articles, is that preferences cannot at the same time correspond to (actual or hypothetical) choices *and* personal welfare. According to Sen, the conflation of preference with choice is unacceptable because economic theory (and RPT in particular) also assume that preferences are identical to personal welfare. However, it cannot be assumed that what people choose is always uniquely determined by the consideration of their own welfare. He proposed to distinguish personal preferences (representing personal welfare), sympathy and commitment as a way to properly account for people’s choices, in particular in the case of choices guided by moral principles (Sen 1977). In some cases, an agent acting under commitment may choose contrary to his personal preferences. The nature of commitment is well illustrated with the example of the prisoner’s dilemma (see figure 1):

**Figure 1: The prisoner’s dilemma**

		Bob	
		Cooperate	Defect
	Cooperate	10 ; 10	0 ; 11
Ann	Defect	11 ; 0	1 ; 1

As this is well-known, in the prisoner’s dilemma to defect is the dominant strategy for each player. That means that (even if we abstain to making too strong assumptions regarding each player’s belief or knowledge regarding the other’s rationality) whatever the way the other player behaves, rationality entails that everyone has to defect. Indeed, according to RPT, defection is not a prediction based on the assumption that players are rational but rather a *logical* implication of the behaviorist postulate: if we are in a prisoner’s dilemma, players *have to* defect; the fact that players cooperate reveals that they are playing another game than the prisoner’s dilemma (Binmore 1994). According to Sen, this interpretation cannot be right because while many social interactions have the structure of a prisoner’s dilemma, social life would be

---

<sup>7</sup> Obviously, this argument is not valid against ERPT (and therefore against CRPT) since the latter explicitly deals with finite sets of choice data. Moreover, as noted by Hands (2013), the point that consistency requirements are empirically unverifiable is harmless in the context of falsificationism. Sen’s attack on the idea of rationality as internal consistency makes however a much stronger argument against the refutability of the consistency axioms. See below.

hardly possible if defection was in fact pervasive. The only way to escape this problem is to separate choices from preferences.

The problem of cooperation in the prisoner's dilemma does not merely resume to the wrong assumption that agents are enduring egoists who care for nothing but their self-interest. As Sen notes, nothing in the conceptual apparatus of RPT prohibits to endow the agents with other-regarding preferences. Ann's utility might well depend of Bob's welfare and it is perfectly right to construe Ann's utility function such as one of its term is reflecting Bob's payoffs. However, the possibility of altruism (or "sympathy" as Sen calls it) is only a partial explanation of cooperation in a prisoner's dilemma type of situation. Sympathy in the framework of RPT is still egoistic in some sense (Sen 1977, 326) because analytically each player is still maximizing *his* utility function. Thus, an altruistic agent is not acting against his personal welfare when cooperating in a prisoner's dilemma. The assumption of altruism allegedly fails to capture those cases where personal choices and personal welfare cease to be associated. These are cases of *commitments*: "One way of defining commitment is in terms of a person choosing an act that he believes will yield a lower level of personal welfare to him than an alternative that is also available to him... we can expand the definition the definition of commitment to include cases in which the person's choice, while maximizing anticipated personal welfare, would be unaffected under at least one counterfactual condition in which the act chosen would cease to maximize personal welfare" (Sen 1977, 327).

According to Sen's definition, there is commitment as soon as there is an actual or a counterfactual departure between an agent's preferences and his actual choices. The possibility of commitment has tremendous conceptual implications for RPT, in particular regarding the consistency conditions discussed in the preceding section. Indeed, the introduction of commitment breaks the identity between the choice function  $C(S)$  and the preference relation  $R$  since an agent acting under commitment may make a choice contradicting his personal preferences. Crucially, the very notion of commitment implies that one's choice is tied to the act of choice: commitment implies that someone is committed to something during the act of choice. Of course, to say that someone is "committed to his preferences" would make the commitment concept meaningless since this would be nothing but a recovery of the conflation between choice and preference. In particular, the possibility of commitment creates the possibility to distinguish what Sen calls *comprehensive outcomes* from *culmination outcomes* (Sen 1997, 161). Culmination outcomes only depend on the ultimate consequences of an agent's action. Quite the contrary, comprehensive outcomes include everything that leads to the consequences, including the choice process (who is choosing and from which set of alternatives). The acknowledgement of comprehensive outcomes leads to conclude that choice and preference are no longer identical since the choice act may lead one to make a choice that contradict one's preferences over outcomes: social norms and moral imperatives, reputation effects, conventional rule following, can all be explaining factors of the actual choice made by rational agents and thus can provide an explanation of the departure between the agent's choice and his personal preferences.

To characterize more fully commitment, Sen [(2002, 33-34); (1985, 207)] distinguishes three kinds of “privateness” for an agent’s preference ordering: *self-centered welfare*, *self-welfare goal* and *self-goal choice*. In the case of self-centered welfare, a person’s welfare is only a function of his own consumption and of other features contributing to the richness of his life. In the case of self-welfare goal, a person’s unique aim is to maximize his own welfare as represented by his personal preferences. In the case of self-goal choice, person’s choices are exclusively based on the pursuit of his own goals. Clearly, these three kinds of privateness are independent of each other and may be combined in various ways. A purely self-interested and self-regarding individual without any sympathy or antipathy for others satisfies the three criteria of privateness. A self-interested but partially other-regarding person satisfies self-welfare goal and self-goal choice but not self-centered welfare, since some of his goals may be to enhance the welfare of other persons. Things are quite different in the case of commitment, as suggested above. Commitment provides a person a reason for action that can be totally independent from his welfare and/or from his goals. As Philip Pettit puts it (Pettit 2008), commitment can be either goal-modifying or goal-displacing: goal-modification entails a violation of self-welfare goal since it leads one to take into account the externalities produced by one’s behavior on others; goal-displacement entails a violation of self-goal choice since it leads one to *recognize* other persons’ goals and to behave according to them, without incorporating them into one’s own goals.<sup>8</sup> Commitment is thus a loose concept designed to capture a range of reasons for action which cannot be meaningfully reduced to the satisfaction of some set of private preferences. The next section discusses the second part of Sen’s critique of RPT, namely the underlying concept of rationality as internal consistency of choice.

### **3. Sen’s Critique of RPT (II): Rationality as Internal Consistency of Choice**

Sen has latter extended his critique to the more general concept of rationality on which economics entirely relies.<sup>9</sup> I present this critique first since it helps to make more intelligible what Sen was trying to achieve with his concept of commitment.

In the introduction of his 2002 book *Rationality and Freedom*, Sen states that there are three different understandings of the rationality of choice assumption in economics (Sen 2002, 19): a) rationality as internal consistency of choice, b) rationality as self-interest maximization, c) rationality as maximization in

---

<sup>8</sup> See also Sen (2002, 35). The satisfaction of the three requirements entails traditional rational choice theory, which Sen calls “RCT-3” and where the agents are assumed to be purely egoistic. RCT-2 accepts the possibility that agents may have altruistic or social preferences, thus violating self-centered welfare. Finally, the more general form of rational choice theory is RCT-1. In this case, the only constraint set upon the agent’s behavior is that it must be sufficiently regular and meaningfully consistent to be seen as a maximizing behavior with a well-identified maximand. It implies none of the three criteria of privateness and in particular allows seeing commitment as a form of maximization. I return on this point below.

<sup>9</sup> Sen’s articles discussing the rationality concept as well as the issues surrounding the optimizing framework of economic theory have been compiled in the book *Rationality and Freedom* (Sen 2002). All page numbers in this section refer to the latter.

general. Sen's work offers a powerful critique of the first interpretation where rationality of choice is identified with a set of predefined axioms of consistency. As noted in the first section, RPT proposes a list of more or less strong axioms defining what it takes for choices to be consistent. For instance, the two following – relatively weak<sup>10</sup> – axioms guarantee the binariness of the choice function for finite sets of preference relations, *i.e.* the fact that the choice function generates a revealed preference relation that will in turn regenerate the choice function (Sen 1993, 128):

(6) *Independence of irrelevant alternatives*:  $[x \in C(S) \ \& \ x \in T \subseteq S] \Rightarrow x \in C(T)$

(7) *Basic expansion consistency*:  $[x \in \bigcap_j C(S_j) \text{ for all } S_j] \Rightarrow x \in C(\bigcup_j S_j)$ .

The axiom of independence of irrelevant alternatives (also known as the Chernoff condition, or axiom of “basic contraction consistency”) states that if some alternative  $x$  is chosen from a given set  $S$ , then it must be chosen from any subset  $T$  which contains  $x$ . The axiom of basic expansion consistency indicates that if an alternative  $x$  is chosen from different sets  $S_j$ , then it must be chosen from the union of these sets. Both axioms combined ensure that a revealed preference relation is generated by the choice function while at the same time an alternative  $x$  that is weakly preferred to all other alternatives in a set  $S$  will be chosen from  $S$  in the corresponding choice function:

(8) *Binariness of the choice function*: for every nonempty  $S$ ,

$$C(S) = [x | x \in S \ \& \ \forall y \in S: xRy].$$

The combination of the independence of irrelevant alternatives axiom and of the basic expansion consistency axiom guarantees the binariness of the choice function. These axioms are generally taken to be intuitive and reasonable criteria of consistency of choice. In an understanding where rationality is identified to consistency of choice, then they also defined what it is to be a rational agent. One of Sen's major points has been to argue that this definition of rationality as the fact of having consistent preferences (and thus making consistent choices) is highly problematic.

Sen's argument is a logical follow up of his critique of the conflation of preference and choice. Indeed, rationality as consistency of choice ultimately relies on this same conflation. To choose  $x$  when  $y$  is available in some context and to choose  $y$  when  $x$  is available in another context seem not properly speaking “inconsistent”; however, preferring  $x$  to  $y$  is inconsistent with preferring  $y$  to  $x$  (Sen 1973, 243). The difference is that preference relations are assumed to encompass everything that matter to the agent in terms of “betterness”. But as soon as one accepts that choice and preference are not two identical things, there is no reason that choosing sometimes  $x$  rather than  $y$  and

---

<sup>10</sup> I say “relatively weak” because these two axioms together do not entail that preferences are transitive. Transitivity of preferences is a necessary condition for the representation of a behavior as a maximization of some utility function. Obviously, the objections against these two axioms as criteria of rationality also hold against stronger requirements of transitivity such as WARP or SARP.

sometimes  $y$  rather than  $x$  is inconsistent. The reason for this is that there can be many factors that influence the choice one is making in a particular context. This argument casts doubts on the very relevance of internal consistency criteria such as the binariness of the choice function. For instance, consider the following general two choice acts:

$$(9) \quad C(\{x, z\}) = \{z\}$$

$$(10) \quad C(\{x, y, z\}) = \{x, y\}$$

If these two choices are made by the same person, then the criteria of consistency presented above should lead one to conclude on the inconsistency of the agent. Indeed, the combination of expressions (9) and (10) violates the binariness condition. However, this is not so straightforward. Suppose that you are invited at a table dinner and that there are two remaining apples in the fruit basket. You have the choice between taking one apple ( $x$ ), or the other apple ( $y$ ), or taking none of them ( $z$ ). Assuming that both apples are close to be identical, you are indifferent between  $x$  and  $y$  but you prefer to have one apple to the case you have none. Now, take the same situation with a slight difference: there is only one apple in the basket. You have to choose between taking that apple ( $x$ ) and not taking it ( $z$ ). Allegedly, if you are the same person with the same preferences, you prefer to have the apple to not having it. So you pick  $x$ , right? In fact, most persons in this last situation would prefer the option  $z$  while choosing  $x$  or  $y$  in the former situation because they follow some norm of ‘etiquette’. This is an ostensible violation of the binariness condition and thus of the stronger consistency axioms. But this example only indicates that the internal consistency of choice can only be recovered if we take into account the influence of a plurality of external conditions (Sen 1993, 129): most of the time, choices are influenced by external factors that are not properly accounted for by the choice function. The point is not that the behavior of the agent in the above example is inconsistent and thus irrational; rather, it is that this underscores that it is irrelevant to reduce rationality to conditions of *internal* consistency of a choice function.

Sen offers a list of mechanisms causing WARP, binariness of the choice function, and other conditions of internal consistency to be violated without calling into question the agent’s rationality [(1993, 130); (1997, 161-170)]. A first mechanism is *positional choice* which indeed corresponds to the above example: a choice may depend on preferences conditional of the position of the item chosen in the set of available items. A second mechanism is *epistemic value of menu*, corresponding to cases where one is uncertain about the proper content of the set of available alternatives. For instance, you may prefer to take beef when the only alternative in the menu is fish at a restaurant, but you may prefer fish rather than beef if a third alternative in the menu is dog meat. Here, the content of the menu gives relevant information over the true nature of the available alternatives and as a result may change one’s preference ordering. A third mechanism is *freedom to reject*: the fact that an alternative is available may affect one’s preference ordering compared to the case where it is not because the fact that one is free to choose this very alternative makes one of the other option less attractive. All these cases are instances of what Sen calls *menu-dependence*: the actual choice of the agent depends on the content of the

set of available alternatives in a way that makes internal consistency conditions meaningless. A fourth, slightly different case, is *chooser dependence*. A preference ordering of an agent may be dependent on the identity of who is actually making the choice. For instance, if two apples and one banana are available in the fruit basket, you may arguably rationally prefer to take one of the two apples if you have to make the choice yourself, while preferring the banana if the choice is made for you by someone else.

Menu dependence and chooser dependence indicate that a preference relation  $R_i$  of a person  $i$  is conditional on the chooser  $j$  and the set  $S$  from which the choice is being made:  $R_i^{j,S}$  (Sen 1997, 166-7). Therefore, the preference ordering of the person changes with the identity of the chooser and the set of available alternative. Obviously, this is in contradiction with any *inter-menu* condition of consistency. But this is precisely the point: the rationality-as-consistency fails to recognize that the choice of an agents and the way he ranks the available alternatives depend on a set of circumstances that are not reducible to the act of choice itself. There is more to a rational choice than the exhibition of some internal consistency.<sup>11</sup> Sen's concept of commitment, surveyed in the preceding section, is a way to account for the fact that consistency of choice cannot be properly accounted for without taking into account external factors influencing the agent's behavior.

#### 4. CTRP and the Neo-Samuelsonian Critique of Sen's Critique

Does Sen's critique against TRPT extend to CRPT? Hands (2013, 1096-7) argues that it is weakened by two facts: firstly, while Sen's argument that consistency axioms cannot be supported by infinite data sets on choice behavior is relevant as far as TRPT is concerned, it is no longer relevant against CRPT which, in the continuity of Afriat (1963) explicitly deals with finite data sets. The point is that Sen's claim that consistency axioms are unverifiable is innocuous if one accepts that verifiability has never been an adequate criterion for evaluating a scientific hypothesis. Secondly, Sen's critique as a whole relies on the assumption that preferences should/could be interpreted as referring to mental events (as this is the case for the concepts of desire or belief). Part of this assumption was sustained by the preceding argument (*i.e.* since no one observes all choices made for all prices and all income levels, the appealing character of the consistency axioms comes from the fact that preference is understood as a psychological concept) which is no longer valid against CRPT. Moreover, as I will explain below, this

---

<sup>11</sup> As Sen emphasizes, the conditions of internal consistency are not saved by the indexation of the preference relation to the identity of the chooser and to the set of alternatives. The various axioms of internal consistency have been defined as axioms of inter-menu consistency. Of course, it is still possible to define a richer – menu-*dependent* – preference ranking to obtain a menu-*independent* choice function (Sen 1997, 174, Theorem 3.1). But the identification of choice with preference is lost in the process. Moreover, as argued by Bhattacharyya et al. (2011), the fact that it is virtually always possible to refine the choice function and the underlying preference relation to account for seemingly inconsistent choices creates a problem of theoretical under-determination: given the observation of an inconsistent behavior on the basis of a choice problem described by the triple  $\langle X, K, C(\cdot) \rangle$ , it is impossible to determine whether this inconsistency is due to truly "irrational" behavior or to the fact that the choice problem is not appropriately specified.

(ontological) assumption is explicitly rejected by proponents of CRPT on more or less sophisticated philosophical grounds. For Sen's critique to have any value, it seems that one must accept the relevance of folk psychology. But the rejection of folk psychology is precisely at the core of the most sophisticated expression of CRPT, namely the neo-Samuelsonian interpretation.

CRPT can be associated with two distinct philosophical and methodological stances, beyond the common mathematical and theoretical apparatus surveyed in the first section. The first is the purely behaviorist interpretation defended by Gul and Pesendorfer (2008) in their manifesto against neuroeconomics. Gul and Pesendorfer argue for a strong separation between economics and psychology on the basis of the claim that economists are not interested in explaining the choices made by individuals. Indeed, economics is not about studying the causes of preferences. Moreover, economics (or at least choice theory) proceeds on the basis of a very peculiar and restricted kind of information about people's choice:

“A choice theory paper in economics must identify the revealed preference implications of the model presented and describe how revealed preference methods can be used to identify its parameters. Revealed preference earns such a central role in economics because this is the form of evidence that is available to economists – and not because of a philosophical stance against other forms of evidence” (Gul and Pesendorfer 2008, 36).

Despite their explicit denial of any philosophical commitment, Gul and Pesendorfer's identification of choice theory with RPT is actually well grounded on a particular philosophical stance. Firstly, as noted above in section 1, Gul and Pesendorfer make an imperialistic claim: they explicitly argue that all choice theory must use RPT. Secondly, their interpretation of the underlying mathematical structure of RPT follows a strong behaviorist line which is not logically implied by the latter. Indeed, the claim that economics is not concerned with explaining what causes persons' preferences (and thus what explains their choices) as well as the insistence that economists use (or should use) only a specific kind of evidence do not derive from RPT as such. These are core *philosophical* and *methodological* propositions that define CRPT as a research program.

This is not however the only philosophical stance through which CRPT can be defined. In fact, Ross (2014, 201) proposes an alternative, non-behaviorist characterization of CRPT that he calls “neo-Samuelsonian”. As Ross (2011, 218) notes, while Gul and Pesendorfer's essay makes a crucial methodological point regarding the specificity of economics (relatively to psychology) and what economists try to achieve,

“the claim that the domain of economics in principle excludes latent psychological variables is arbitrary methodological legislation that flies in the face of numerous precedents. Economic models abound with exogenous institutional, cultural, technological, political and other variables, and Gul and Pesendorfer provide no reasonable grounds for banning by ukase exogenous influences on ‘choice’ that might be inferred”.

Instead, the neo-Samuelsonian interpretation of RPT “interpret(s) preferences as summaries of, rather than causes of, economic choices” but recognizes “that

latent cognitive processes play roles in generating economic data” (Ross 2014, 201). The neo-Samuelsonian variant thus retains the constitutive features of CRPT: consistency axioms are used to make inferences on the basis of *finite* set of data of consumer choice, preferences are identified to choice patterns and are not defined as mental events and the methodology of RPT is claimed to apply to choice theory as a whole. An important difference though is that the neo-Samuelsonian approach does not ignore the role played by latent mental variables and presumably does not exclude to take them into consideration when it may help to refine the use of the theory.<sup>12</sup>

Sen’s critique of RPT has itself been attacked by neo-Samuelsonians on the basis that CRPT has the ability to take people’s motivations into account: “Sen’s criticism of RPT combines the correct claim that motivation should be endogenized within our models of agents, insofar as they are relevant to the agents’ behavior, with the claim that RPT inherently lacks the resources to do this” (Ross 2005, 132). Dowding (2002) claims that Sen’s criticism of RPT errs because it relies on a mistaken *internalism* about the mind. That is, Sen’s criticism assumes that preferences are intentional states and that intentional states (beliefs, desires) literally “exists” in the mind. Thus, Sen seems to be following a standard folk psychology account according to which individuals’ behavior can be explained by referring to folk concepts such a beliefs, desires or preferences. In this case, the twofold claim that preferences cannot be identified with choices (section 2) and that rationality cannot be meaningfully defined as internal consistency of choice (section 3) are hardly disputable. However, Ross and Dowding both insist that internalism is misguided and is contrary both to the current orthodoxy in the philosophy of mind and to the way the concept of preference should be interpreted in choice theory. Instead, they argue that the proper philosophical framework of RPT is *externalism* and in particular Dennett’s (1987) intentional functionalism. According to the latter, we explain the behavior of an entity (an individual, but also an animal or a machine) by *ascribing* to this entity some intentions. Intentions are external in the sense that they are the product of an interaction between the agent whose behavior is to be explained or understood and the observer’s willingness to explain the agent’s behavior.<sup>13</sup> In other words, intentions are not “real” mental states constitutive of an agent’s mind; they are theoretical constructions one is using to account for someone behavior.

The neo-Samuelsonian version of CRPT consists in interpreting the preference concept according to intentional functionalism. This means that preferences are not internal mental states, but rather *consistent patterns of choices* that the economist is able to extract from the mass of available data. These patterns eventually result from the interaction between the individuals’ latent cognitive processes and their physical and institutional environment, such as the actions physically available, the prices and opportunity costs, and so on. The ascription of preferences to an agent is thus a way to rationalize his behavior in a specific context, without necessarily assuming that these preferences refer to stable mental states. Under this externalist interpretation of the preference concept,

---

<sup>12</sup> Ross (2013, 201) cites Ken Binmore, Glenn Harrison, Charlie Plott and Vernon Smith as representative authors adhering to the neo-Samuelsonian variant of CRPT.

<sup>13</sup> This includes the case where an agent himself tries to account for his own behavior. In this case, the agent takes the intentional stance toward himself.

Sen's critique seems to collapse. Sen's claim that we have to refer to elements external to the act of choice to explain choice is vacuously true: preferences and more generally intentional descriptions of an agent's behavior are necessarily external to the agent's mind. The point of using axioms of consistency of choice is that they are tools to uncover the agents' motivations and how they frame the various aspects of the choice problem. Sen's examples illustrating the reasonable violation of WARP or of the binariness condition are then not convincing: what they show is that the apparent non consistency of choice is actually indicating that the way the economist is constructing the choice problem does not reflect the "intensionality" of its constitutive variables (Dowding 2002, 273).<sup>14</sup> Apparent inconsistency of choice then urges the economist to look for an alternative description of the choice problem. Once consistency has been recovered, the economist can claim to have found the agent's actual maximand and thus the motivations underlying his behavior in some specific context.

As noted above, Sen is obviously aware of the possibility to refine the description of a choice problem to recover some consistency condition. In particular, it is always possible to define *menu-dependent* choices functions and preference relations on the basis of which behavior is trivially consistent. This analytical strategy has led some critics to reject RPT and rational choice theory because they are "tautological" [(Basu 2000); (Hodgson 2012)]. More significantly, Bhattacharyya et al. (2011) argue that this possibility does not save RPT from Sen's critique. They show that for any violation of some consistency condition in a choice problem corresponding to the triple  $\langle X, K, C(\cdot) \rangle$ , it is always possible to generate an alternative description  $\langle X^*, K^*, C(\cdot)^* \rangle$  where the same consistency condition holds. However, they note that this leads to a problem of theoretical under-determination: it is impossible to determine whether the inconsistency observed under the first description comes from the fact that the agent's behavior is truly inconsistent or rather because the choice problem is inappropriately described. As a consequence, RPT as a theory of rational choice fails to meet a scientific standard of refutability.

This problem should not cause much trouble for neo-Samuelsonian economists though. A first reason is that RPT has never been conceived as a theory of rational *individual* behavior: "Consumer theory has never been tested by examining the behavior of *one* consumer. Rather the interpretation of prices given changes in demand and supply conditions have been tested by examination of mass consumer behavior" (Dowding 2002, 275, emphasis in original). Therefore, the fact that an agent's rationality is an assumption that cannot be falsified under RPT is not an objection: RPT has not been built to test the consumer's rationality but to make inference on finite aggregate data on choice behavior. RPT is a problem-solving device whose function is to permit the measurement of some class of coefficients (such as marginal rates of

---

<sup>14</sup> Philosophers of mind distinguish between the *extension* and the *intension* of a statement or proposition. The former refers to the property of two or several statements or propositions carrying the same meaning. For instance, writing 4 and  $\sqrt{16}$  has the same mathematical meaning and both expressions are thus extensionally equivalent. The latter refers to a statement or a proposition which fails to pass the test for extensionality. Sen's examples depend on such intensional features: the same variable  $x$  does not carry the same meaning for the decision maker across two different choice problems.

substitution or elasticities of substitution) for the purpose of explanation or prediction (Ross 2005, 142). A second, and more general reason is that economics (or choice theory) is not a theory of rationality but rather of the interactions of agents in a market context (Ross 2014). Whether or not agents are rational (and what is the meaning of being ‘rational’) is then an irrelevant question because what matters are only the aggregate behavioral patterns that result from the interactions between agents. Theoretical results showing that consistency and other properties of aggregate variables such as demand functions are not tied to the rationality of individual agents [(Becker 1962); (Gode and Sunder 1993)] support this interpretation of what economics is about.<sup>15</sup> Therefore, if we accept that choice theory or economics have nothing to say about individual rationality, then the neo-Samuelsonian variant of CRPT definitively defeats Sen’s critique. I shall argue however in the next two sections that at least two major themes in Sen’s philosophy of economics are well under the scope of choice theory and that the neo-Samuelsonian CRPT cannot deal properly with them. This is a major limit given the imperialistic ambitions of CRPT.

## 5. Rationality and Incomplete Preferences

Though Sen is critical toward the idea that rationality can be reduced to maximization of self-interest or personal welfare, he does not reject the concept of maximization *per se*. Indeed, once the requirement for introducing the influence of factors external to the personal preferences of the agent is met to account for the consistency of choices, most behaviors are amenable to a description in terms of a maximization of some maximand. Maximization is a *necessary* condition for a behavior to be qualified as “rational”. But Sen insists that it is not a *sufficient* condition and that rationality demands more than that (Sen 2002, 32, 39). Beyond maximizing some quantity (which implies that choices exhibit some consistency), rationality entails submitting one’s choice to reflexive scrutiny and reasoning:

“Rationality is interpreted here, broadly, as the discipline of subjecting one’s choices – of actions as well as of objectives, values and priorities – to reasoned scrutiny” (Sen 2002, 4).

“A person is not only an entity that can enjoy one’s own consumption, experience and appreciate one’s welfare, and have one’s goals, but also an entity that can examine one’s values and objectives and choose in the light of those values and objectives... We might or might not be much moved by moral concerns or by social reasons, but neither are we prohibited from entertaining these questions, in shaping our values and if necessary revising our objectives in that light” (Sen 2002, 36).

“Rationality of choice, in this view, is primarily a matter of basing our choices – explicitly or by implication – on reasoning that we

---

<sup>15</sup> Of course, the same is true with the “inverse” result, *i.e.* the theoretical result that well-behaved individual demand functions does not entail that the aggregate demand function is also well-behaved.

can reflectively *sustain* if we subject them to critical scrutiny” (Sen 2009, 180).

This idea of rationality as one’s ability to reason and to reflect on one’s choices is reminiscent of the critique of the Humean view of rationality developed by the American philosopher John Searle (2001). The Humean view wrongly assumes that an agent enters into an interaction with an already complete and well-behaved preference ranking (Searle 2001, 30). Indeed, Searle even argues that the ability to adjudicate between conflicting goals is the distinguishing feature of practical reason relatively to theoretical reason: well-ordered preferences are the *product* of practical reasoning, not a prior condition for it (Searle 2001, 253). In Sen’s philosophy of economics, this idea is related to two key themes that are also relevant for choice theory: the incompleteness of preference orderings on the one hand, the variety of levels of agency on the other. This section focuses on the former, the next one on the latter.

In standard utility theory, rationality depends on the assumption that an agent has a *complete* preference ordering. Completeness means that an agent must be able to relate any two states  $x$  and  $y$  by a relation of preference  $R$ . Incompleteness means that at least one state (which can be either a culmination outcome or a more inclusive comprehensive outcome) is not related to at least one other state. Incompleteness is not the same as indifference: for instance, the story of the Buridan’s ass that died of starvation because it was unable to make a choice between two haystacks is an authentic case of incompleteness (Sen 2002, 16). Indifference does not make choice impossible; any choosing device (such as throwing a coin) will be sufficient for the choice to be made because whatever the choice, one’s preferences will be equally satisfied. Incompleteness is completely different because the fact that one has no preference between two alternatives means that these two alternatives cannot be compared along common dimensions. Choosing between incomparable alternatives has no meaning in a revealed preference framework. Still, incompleteness cannot be assumed away for purely formal reasons in particular because incommensurability is a characteristic of many decision problems particularly when they involve moral matters or are about justice issues.

A significant part of Sen’s writings is devoted to show that incompleteness does not proscribe maximization. In his article “Maximization and the Act of Choice” (Sen 1997), Sen makes a rigorous distinction between *optimization* as the act of choosing the best alternative and *maximization* as the act of choosing an alternative that is not worse than any other available one. Incompleteness rules out optimization. For instance, assume that Ann must choose between three available alternatives  $x$ ,  $y$  and  $z$ . Ann’s preference ordering is the following

$$xRy$$

$$zRy$$

$$\neg xRz$$

$$\neg zRx$$

where  $R$  refers to the relation of weak preference. Ann’s preference ordering is incomplete because no preference relation exists between  $x$  and  $z$ . In this case,

Ann's behavior cannot optimize anything because Ann is unable to identify the best alternative. However, despite the incompleteness, both  $x$  and  $z$  are clearly no worse than any other available alternative. Therefore, no matter whether Ann chooses  $x$  or  $z$ , her behavior can be related to a meaningful maximization. This example illustrates a more general result showing that maximization has a wider scope than optimization (Sen 1997, theorems 5.1 to 5.6). In particular, for a weak preference relation  $R$  and for some set of alternatives  $S$ , Sen establishes that the optimal set  $B(R, S)$  is always a subset of the maximal set  $M(R, S)$ .

Moreover, it can be proved that for any maximizing choice function generating the maximal set  $M(R, S)$ , one can always define a binary relation  $R^+$  to generate an optimal choice function such that  $B(R^+, S) = M(R, S)$ . Crucially, the contrary is not true: it may not be possible to devise a preference relation  $R^+$  such that an optimizing choice function will be replicated by a maximizing choice function. The point is thus that any maximization can be seen as an optimization if incompleteness is assumed away by substituting the absence of relation between alternatives for a relation of indifference. But this formal trick does not work the other way around: an optimizing framework cannot be always seen as a maximizing framework. There are two reasons for this: firstly, menu-dependence may make impossible to define a menu-independent preference relation  $R^+$  mimicking the maximal set tied to the "real" preference relation  $R$ . Secondly, any pair of unconnected alternatives  $x$  and  $y$ <sup>16</sup> implies that  $B(\{x, y\}, R)$  is empty; but it is impossible to devise a relation  $R^+$  such that  $M(\{x, y\}, R^+)$  is also empty since that would imply that  $x$  is strictly preferred to  $y$  and  $y$  strictly preferred to  $x$  at the same time, which is a logical contradiction. In other words, incompleteness forbids reducing optimization to maximization.

These technicalities demonstrate the importance of incompleteness to account for the rationality of an agent's choices and behavior. If one takes maximization to be a necessary condition for rationality, and since incompleteness does not make impossible maximization, then incompleteness of preference rankings is not contradictory with rational behavior. A major implication is that the strong assumption that one must enter into a social interaction with an already complete preference ordering is unnecessary as a component of a theory of rational behavior. One may be unable to reveal a preference relation regarding two alternatives, for instance because he has not reflected properly on them, and still be seen as maximizing some quantity. Still, incompleteness is not necessarily ultimately insurmountable. Quite the contrary, rational agents may feel the urge to solve incompleteness in particular when it comes from the fact that they do not have sufficiently reflect on a particular issue. Therefore, it is useful to distinguish *assertive* from *tentative* incompleteness [(Sen 2002, 17); (Sen 2009, 108)]. Incompleteness is assertive when the incommensurability between two alternatives is proclaimed by an agent or, in the case of a social choice situation, through some aggregation mechanism. But in many cases, incompleteness stems from the fact that individuals do not have sufficiently reasoned on the issue related to these alternatives.

---

<sup>16</sup> This means that neither  $xRy$  nor  $yRx$  is true.

It can be hardly disputed that incompleteness of preferences is a significant theoretical problem in choice theory, in particular when the latter is applied to normative issues. For instance, axiomatic derivations of utilitarianism on the basis of expected utility theory and their applications to health problems or climate change crucially depend on the meaningfulness of the completeness axiom [(Broome 1991); (Broome 2004); (Harsanyi 1955)]. Most of the social choice literature also relies on the assumption that the social preference ordering must be complete and that the individual preference profiles are complete. Incompleteness thus rules out many of the theoretical results constitutive of choice theory. This is also the case as far as the relevance of RPT's consistency axioms is concerned. Indeed, WARP or the binariness axiom imply that the revealed preference relation  $R$  is complete. Incompleteness thus formally rules out these axioms (Sen 1997, 184). What are the implications for CRPT and in particular for its neo-Samuelsonian variant?

At first sight, the problem of incompleteness seems to be solvable within RPT in a similar way than the problem of inconsistency. For instance, the problem seems to be solved through the completion of the ordering by inserting an indifference relation  $I$  between two unconnected alternatives. As noted above, Sen highlights this formal strategy as a way to construct an *optimal* choice function on the basis of a *maximal* choice function. There are however at least two problems with this strategy. Firstly, at the formal level, in some cases the incompleteness of the weak preference ordering  $R$  also entails a violation of another axiom, namely the axiom of transitivity of the *negation* of the weak preference relation  $R$  (Putnam 2002, 82-83). The latter states that for three alternatives  $x$ ,  $y$  and  $z$ , if  $x$  is not weakly preferred to  $y$  ( $\neg xRy$ ) and  $y$  is not weakly preferred to  $z$  ( $\neg yRz$ ), then  $x$  should not be weakly preferred to  $z$  ( $\neg xRy \wedge \neg yRz \Rightarrow \neg xRz$ ). There is thus a formal difference between indifference and incompleteness that is not eliminated by the introduction of an indifference relation when connectedness is lacking. Secondly, as argued by Sen, at a more substantive level there is a clear difference between indifference and incompleteness. The latter may result either from incommensurability or vagueness.<sup>17</sup> Incommensurability means that two alternatives cannot be compared along the same dimensions, making them ultimately impossible to rank. Vagueness means that for some reason it is impossible to evaluate precisely the value of two or several alternatives such that it cannot be asserted that one alternative is better than the others. Clearly, neither incommensurability nor vagueness can be meaningfully reduced to indifference.

How neo-Samuelsonians can then handle the incompleteness issue? Acknowledging the fact that incompleteness does not negate the possibility of choice (Sen 2004) and is not in contradiction with the notion of maximization, this leaves them with some sort of conundrum. Obviously, the very notion of "incomplete preference ordering" has a dubious meaning in a revealed preference framework. Actually, it is probably meaningless (Mandler 2005): if preferences are defined as behavioral patterns resulting from the complex

---

<sup>17</sup> For the relationship between incompleteness and incommensurability, see Sen (2004). For the argument that incompleteness essentially results from vagueness, see Broome (2004, 165-186).

interactions between institutional structures and latent cognitive processes, then it is simply impossible to “observe” incompleteness. Two possibilities then remain: the first possibility for neo-Samuelsonian is simply to reject the concept of incompleteness on the ground that it relies on an “internalist” understanding of the preference concept and that it is outside the scope of economics. The second possibility consists in decoupling the concepts of choice and preference by enlarging the *informational basis* for determining the agents’ preferences. The former is problematical even in the terms of RPT: even if incompleteness is rejected as meaningless, it can result at the behavioral level in some form of instability or inconsistency, thus undermining RPT as an inferential and predictive tool. Though the latter possibility contradicts the strict behaviorist understanding of RPT, it does not imply the reduction of economics to psychology. More importantly, it can help to make better inferences at the aggregate level on the basis of finite choice data. But where neo-Samuelsonians to retain this latter option, they should then recognize the relevance of several of Sen’s concepts, such as the commitment concept.

## 6. Rationality and Levels of Agency

Incompleteness of preference orderings is not the only issue related to Sen’s definition of rationality as reasoning plus maximization that cannot be dealt with by CRPT. Another key issue, which this time directly involves the commitment concept, corresponds to what can be called the levels of agency problem. As for the incompleteness issue, the levels of agency problem is directly relevant for choice theory, in particular (but not only) for its application in the normative domain. And contrary to the incompleteness issue, the relevance of the levels of agency problem for economics is explicitly recognized by at least some neo-Samuelsonians (Ross 2005).

The possibility for a *person* to act on the basis of different levels of agency appears in Sen’s distinction between self-centered welfare, self-welfare goal and self-goal choice (see section 2). This distinction emphasizes the possibility for the same and unique person to act on different basis: her own self-interest, her personal welfare (which may not necessarily corresponds to her self-interest), her personal goals (which may not systematically consist into maximizing her personal welfare). In other words, *the unity of personality does not imply the unity of agency*. The concept of commitment makes the link between personality and agency by introducing a “fourth aspect of the self” in Sen’s writings (Davis 2008, 326):

“Commitment, consequently, is not an unreflective type of attachment to others, but rather a rational recognition of rules associated with social membership that can only be achieved by individuals who have distanced themselves from their own interest as self-scrutinizing individuals”.

The commitment concept captures the ability of individuals to reflect on the goals on the basis of which they act, but also to act on the basis of other persons’ goals or even of “external reasons” that do not depend on anyone’s goals (Sen 2008, 347-8):

“A person’s decision not to act only – or even mainly – on the basis of his or her own goals does not require that the person must then be acting according to some “other people’s goals.” One can take note of other people’s goals and

priorities and decide to constrain the unifocal pursuit of one's own goals with behavioral constraints and other restrictions, without that self-restraint being interpreted as the pursuit of the goals of others".

Commitment can thus be not only "goal-modifying" (*i.e.* it changes the goals one is pursuing) but also "goal-displacing" (*i.e.* it leads one to pursue others' goals or persons-independent goals). Interestingly, this latter property of commitment has been criticized on the basis that it fits uneasily within the framework of folk psychology (Pettit 2008). Indeed, it is almost an analytical truth according to folk psychology that one always act on the basis of his desires, beliefs, goals and any other relevant mental states. This may explain why the commitment concept has been criticized not only by revealed preference theorists but also by authors wholly sympathetic toward Sen's philosophy of economics.

Beyond the fact that this point casts some doubts on Ross' (2005) and Dowding's (2002) claim that Sen is endorsing internalism toward preferences, more significantly it emphasizes that for Sen social identity and the ability of persons to recognize themselves as the members of some community or group is essential to understand many economic phenomena, including market behavior. Actually, what Sen has in mind when speaking of a person acting neither on the basis of her own goals nor of other people's goals corresponds to what Wittgenstein (1953) identified as *rule-following behavior*. A distinguishing feature of humans is to be able to commit, as members of a community, to follow more or less abstract and impersonal rules not necessarily tied to a person (oneself or anyone else). As a member of a community, I can commit myself to act on the basis of the group's welfare defined in some way or the other.<sup>18</sup> But I can also act on the basis of rules which I recognize are holding in the community independently of any welfare consideration, as this is the case for many moral rules.

Rule-following behavior and more generally the notion of "goal displacement" is of course related to Sen's critique of rationality as internal consistency of choice. Rules are typically the kind of "external" objects that influence and constrain individual choice. In some cases, rules may also take the form of "self-imposed constraints" (Sen 1997, 189-192) that lead to a discrepancy between choice and personal preferences (corresponding to personal welfare and/or personal goals). Though these constraints are "internal", they generate the same kind of menu-dependent preferences than "truly" external constraints such as social rules. As it should be clear from what has been said in section 4, rule-following behavior and self-imposed constraints are not a problem for RPT at the formal level. Moreover, it can be argued that the fact that there are multiple levels of agency is also not a difficulty because RPT does not presuppose any particular conception of the individual (Ross 2005). Indeed, RPT is a tool for studying behavioral pattern which does not make any assumption about the "vehicle" of these behaviors: these can be persons, but also groups or "selves". Therefore, in complement with other tools such as game theory and overlapping generation models, RPT can be used to study

---

<sup>18</sup> See the emerging literature on "team-reasoning" in game theory which formalizes this possibility [(Bacharach 2006); (Sugden 2000)].

“intrapersonal bargaining” between selves as well the collective decisions of some group (Ross 2014).

However, Sen’s point is not only that people can act on the basis of others persons’ goals or of external rules related to some group or community. That people may have a plurality of motivations lying behind choice is not controversial. But he also emphasizes that reasoning, as a property of rationality, also entails that a same person may act on the basis of different and sometimes contradictory reasons. In other words, a person’s motivations and the agency level of a given behavior are changing as a function of times or other environmental consequences. That means that the same *person* may endows the identity of different *agents* as her behavior is revealing different preference orderings in different context. Though RPT is able to trace a consistent ordering of revealed preferences through any finite set of choice data, it is obviously unable to undercover the changes of agency levels that any person exhibits, sometimes under a very small timeframe. These changes will generate choice inconsistencies which can be eliminated by an appropriate reformulation of the choice problem. However, this reformulation will not necessarily reveal the fact that the same person was acting on the basis of different motivations.<sup>19</sup>

The issues of incompleteness and levels of agency<sup>20</sup> are deeply related to Sen’s latest interest for public reasoning in the understanding of justice (Sen 2009). Sen’s work on justice is strongly grounded on the postulate that any transcendental and idealistic approach of justice is deemed to fail. The justification of this postulate is partly practical (“ideal” or “perfect” institutional arrangements may not be implementable given the particular socioeconomic context) but essentially foundational: issues of social justice will typically involve the collusion between different and equally reasonable principles of justice. Different principles of justice are likely to lead to different prescriptions about what has to be done regarding, for instance, socioeconomic inequalities. As a result, it may be difficult or impossible to define an optimal set of social situations because the members of a population are collectively unable to rank two or more social alternatives. What is true in a social choice framework (*i.e.* when several individuals must define a single and collective ranking of social alternatives) may be also true at an individual level: a rational agent may be unable to state a preference between donating money for a foundation fighting against poverty and for a foundation collecting funds for research on treatments against cancer. The role of public reasoning, in particular in a democratic society, is precisely to help individuals and collectives to adjudicate between alternatives in such difficult issues. Public debates help the production and the sharing of information, but may also lead individuals to change their minds. In particular, public discussions may help

---

<sup>19</sup> An alternative is to treat separately the choices of each agent the person has endorsed the identity of. Then, there are as many agents as there are different consistent behavioral patterns. But this leads to an identification problem: to distinguish several agents we must know among the whole set of choice data what are the choices corresponding to each agent. But the latter necessitates to know already who are the agents making the choices. Davis (2003) makes a similar point against the conception of the individual in standard economics.

<sup>20</sup> Note that the former can be seen as a special case of the latter: a person can create a preference relation between two alternatives by endorsing the point of view of another person or of a group, that is by changing his level of agency.

individuals to ultimately agree on a specific moral or political issue and thus to finally make what was an incomplete social preference ordering a complete one.<sup>21</sup>

This ability to redefine and to complete a preference ranking is what both Sen and Searle (2001) refer to *reasoning*. Rationality is not only about maximizing; it is also about reflecting on what to maximize. Somehow, rational agents have the ability to “choose” the maximal choice function corresponding to their behavior. Sen’s important discussion of rankings of preference rankings, or “meta-rankings” (Sen (1977); (2002)) formalizes this intuition. But this should also lead to a reconsideration of the commitment concept. The notion of commitment implies that the agent is correlating his behavior to something that is not reducible to what he intrinsically prefers. This implies in turn that the agent has the ability to *recognize* a reason for action that is external to his intrinsic desires. This recognition is rational because it depends on the agent’s reasoning abilities. External reasons for action are almost surely never determined *before* social interaction. A hypothetical agent living alone on a desert island since his childhood could not have preferences corresponding to other factors than his biological needs and psychological desires. Therefore, the preference ordering of such hypothetical individuals would be necessarily incomplete: he could not express preference regarding issues involving “social” matters.

For instance, the rational agent who cooperates in the prisoner’s dilemma because he follows a social norm or a moral requirement must have, before choosing to cooperate, recognized the norm or the moral requirement. This recognition may be the product of a reasoning process fueled, for example, by discussions with moral philosophers explaining the “rationality” of cooperation. Or it may be caused by the fact the agent has interacted with other agents following the norm. In this case, reasoning and rational recognition imply a change in the preference ordering through the commitment to follow a norm. This is an instance of goal-modifying mechanism. In some other cases, commitment may resolve an incommensurability issue: two alternatives where seen as non-comparable by the agent, until his recognition of a moral requirement solves the comparability problem and makes the agent prefer one alternative to the other. In this case, it seems that we have an instance of goal-displacing mechanism: an alternative that was until now not ranked may suddenly become the preferred one for an individual because of his commitment to follow a moral requirement. The inability of RPT to reflect on the way persons make choices on the basis of reasoning is thus at the heart of Sen’s critique. CRPT and in particular its neo-Samuelsonian variant do not answer this concern. Since taking into account this ability might be necessary to be able to infer hypothetical choices from observed choices, it can be argued that Sen’s critique establishes that CRPT must fail as an imperialistic framework in choice theory.

---

<sup>21</sup> As Sen himself notes several times in his writings, this idea that the process of collective choice may lead individuals to change their preference orderings has been discussed very early by James Buchanan [(1954a) ; (1954b)].

## 7. Conclusion

Sen's writings on RPT and more generally on the concepts of preferences and rationality in economics do not merely developed a critique of standard choice theory. They are also insightful contributions advancing choice theory. I have suggested that Sen's critique remains relevant even with the development of CRPT and particularly of its neo-Samuelsonian variant. Neo-Samuelsonians are right to claim that RPT has the resources to deal with external factors determining choices. Moreover, their commitment to externalism and the resulting conflation of choice with preference do not imply the rejection of latent cognitive process as being outside the scope of economics. On these points, Sen's critique does not concern CRPT. However, my main claim has been that neo-Samuelsonian economics is unable to deal properly with two major themes in Sen's philosophy of economics: incomplete preferences and the agency levels issue. Since there are central issues in choice theory, this is a major flaw given the imperialistic ambitions of CRPT.

## References

- Afriat, S. N. 1967. "The Construction of Utility Functions from Expenditure Data." *International Economic Review* 8(1): 67–77.
- Bacharach, Michael. 2006. *Beyond Individual Choice: Teams and Frames in Game Theory*, edited by N. Gold and R. Sugden, Princeton University Press.
- Basu, Kaushik. 2000. *Prelude to Political Economy: A Study of the Social and Political Foundations of Economics*. Oxford University Press.
- Becker, Gary S. 1962. "Irrational Behavior and Economic Theory." *Journal of Political Economy* 70: 1–13.
- Bernheim, B. Douglas and Antonio Rangel 2008. "Choice-Theoretic Foundations for Behavioral Welfare Economics." In A. Caplin & A. Schotter (eds.), *The Foundations of Positive and Normative Economics*, Oxford University Press, 155-192.
- Bhattacharyya, Aditi, Prasanta K. Pattanaik, and Yongsheng Xu. 2011. "Choice, Internal Consistency and Rationality." *Economics and Philosophy* 27(2): 123–49.
- Binmore, Kenneth G. 1994. *Game Theory and the Social Contract: Playing Fair*. MIT Press.
- . 2009. *Rational Decisions*. Princeton University Press.
- Broome, John. 1991. *Weighing Goods: Equality, Uncertainty and Time*. Wiley.
- . 2004. *Weighing Lives*. Oxford University Press.
- Buchanan, James. 1954a. "Social choice, Democracy, and Free Markets". *Journal of Political Economy*, 62(2): 114-23.
- . 1954b. "Individual Choice in Voting and Market". *Journal of Political Economy*, 62(3): 334-43.

- Davis, John B. 2003. *The Theory of the Individual in Economics: Identity and Value*. Routledge.
- . 2008. “Identity and Commitment: Sen's Fourth Aspect of the Self.” In F. Peterand & H.B. Schmidt (eds.), *Rationality and Commitment*, Oxford University Press, 313-334.
- Dennett, Daniel. 1989. *The Intentional Stance*. MIT Press.
- Dowding, Keith. 2002. “Revealed Preference and External Reference.” *Rationality and Society* 14(3): 259–84.
- Giocoli, Nicola. 2003. *Modeling Rational Agents: From Interwar Economics to Early Modern Game Theory*. Edward Elgar Publishing.
- Gode, Dhananjay K., and Shyam Sunder. 1993. “Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality.” *Journal of Political Economy* 101(1): 119–37.
- Gul, Faruk B. and Wolfgang Pesendorfer. 2008. “The Case for Mindless Economics.” In A. Caplin & A. Schotter (eds.), *The Foundations of Positive and Normative Economics*, Oxford University Press, 3-39.
- Hands, D. Wade. 2013. “Foundations of Contemporary Revealed Preference Theory.” *Erkenntnis* 78 (5): 1081–1108.
- Hansson, Sven Ove, and Till Grune-Yanoff. 2012. “Preferences.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2012.
- Harsanyi, John C. 1955. “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility.” *Journal of Political Economy* 63(4): 309–21.
- Hodgson, Geoffrey M. 2012. “On the limits of rational choice theory.” *World Economics Journal*.
- Houthakker, H. S. 1950. “Revealed Preference and the Utility Function.” *Economica*, New Series, 17 (66): 159–74.
- Little, Ian M. D. 1949. “A Reformulation of the Theory of Consumer's Behaviour.” *Oxford Economic Papers*, New Series, 1(1): 90–99.
- Mandler, Michael. 2005. “Incomplete Preferences and Rational Intransitivity of Choice.” *Games and Economic Behavior* 50(2): 255–77.
- Putnam, Hilary. 2002. *The Collapse of the Fact/value Dichotomy and Other Essays*. Harvard University Press.
- Pettit, Philip. 2008. “Construing Sen on Commitment.” In F. Peterand & H.B. Schmidt (eds.), *Rationality and Commitment*, Oxford University Press, 28-48.
- Ross, Don. 2005. *Economic Theory And Cognitive Science: Microexplanation*. MIT Press.
- . 2011. “Estranged Parents and a Schizophrenic Child: Choice in Economics, Psychology and Neuroeconomics.” *Journal of Economic Methodology* 18 (3): 217–31.

- . 2014. *Philosophy of Economics*. Palgrave Macmillan.
- Samuelson, Paul A. 1938. “A Note on the Pure Theory of Consumer’s Behaviour.” *Economica* 51(17).
- Searle, John R. 2001. *Rationality in Action*. MIT Press.
- Sen, Amartya. 1971. “Choice Functions and Revealed Preference.” *Review of Economic Studies* 38 (115): 307–17.
- . 1973. “Behaviour and the Concept of Preference.” *Economica* 40(159): 241–59.
- . 1977. “Rational Fools: A Critique of the Behavioral Foundations of Economic Theory.” *Philosophy and Public Affairs* 6(4): 317–44.
- . 1993. “Internal Consistency of Choice.” *Econometrica* 61(3): 495–521.
- . 1997. “Maximization and the Act of Choice.” *Econometrica* 65 (4): 745–79.
- . 2002. *Rationality and Freedom*. Harvard University Press.
- . 2004. “Incompleteness and Reasoned Choice.” *Synthese* 140 (1-2): 43–59.
- . 2008. “Rational Choice: Discipline, Brand Name and Substance.” In F. Peterand & H.B. Schmidt (eds.), *Rationality and Commitment*, Oxford University Press, 339-360.
- Sugden, Robert. 2000. “Team Preferences.” *Economics and Philosophy* 16(2): 175–204.
- Varian, Hal. 2006. “Revealed Preference.” In M. Szenberg, L. Ramrattan and A.A. Gottesman (eds.), *Samuelsonian Economics and the 21<sup>st</sup> Century*, Oxford University Press.
- Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. John Wiley & Sons.

