Sen on Rationality, Commitment and Preferences

Cyril Hédoin*
REGARDS - University of Reims Champagne-Ardenne

2nd Draft – This version: 15/02/2013

Abstract: This article discusses Sen’s critique of the behavioral implications of revealed preference theory. Sen proposes the concept of commitment as a way to account for the possibility of counter-preferential choice. More generally, it is part of an argument against the thesis that rationality may be reduced to some criteria of internal consistency of choices. The main purpose of this paper is to show that though commitment is more often discussed in the context of Sen’s critique of various consistency axioms of choice such as the transitivity axiom, this concept is also deeply related to the Sen’s equally significant evaluation of the completeness axiom. Comparing Sen’s account to John Searle’s critique of the classical model of rationality, I suggest that Sen uses the commitment concept as a more general critique of the Humean paradigm where “reason is the slave of passions”. Sen’s latest concerns with public reasoning and partial orderings indicate an attempt of his part to go beyond the Humean paradigm of rationality and to reflect on the way preferences are formed and constructed through rational deliberation.

Keywords: Amartya Sen, John Searle, preferences, rationality, consistency of choice, incompleteness of preference orderings

* cyril.hedoin@univ-reims.fr
Sen on Rationality, Commitment and Preferences

1. Introduction

Amartya Sen is one of the earliest and more forceful critics of the behavioral implications of revealed preference theory. He develops a sophisticated critique of the behaviorist postulate equating actual and observable choices to preferences. More generally, Sen expressed skepticism toward the possibility and the relevance of conflating into a unique preference set all the factors causally responsible for the actual behavior of a rational agent. According to Sen, preferences are insufficient to explain choices. Sympathy and commitment provide two other fundamental reasons for action. In particular, commitment implies that an agent’s choice may not reflect his personal welfare. In subsequent writings, Sen gave more substance to his concept of commitment as counter-preferential choices. In particular, commitment either leads an agent to modify her goals or, more radically, to displace them. Sen’s concept of commitment is related to his claim that it is impossible to reduce rationality to some conditions of internal consistency of choice. However, it has been criticized on the ground that a proper understanding of the preference concept as used in contemporary economics makes commitment useless as a theoretical concept. The main purpose of this paper is to show that though commitment is more often discussed in the context of Sen’s critique of various consistency axioms of choice such as the transitivity axiom, this concept is also deeply related to the Sen’s equally significant evaluation of the completeness axiom.

Basically, I emphasize one important parallel and make one key claim. Regarding the parallel, it is interesting to note that Sen’s critique of revealed preference theory almost perfectly matches John Searle’s critique of the “classical model of rationality” (Searle 2003). Indeed, Searle’s notion of “desire-independent reason for action” is equivalent to Sen’s commitment. The key claim is that the parallel between Sen’s and Searle’s critiques reveals that Sen’s target is less revealed preference theory than the Humean paradigm of rationality. In the Humean paradigm, “reason is the slave of passions”, i.e. one cannot rationally deliberate on preferences prior entering into a decision problem. Sen’s latest concerns with public reasoning and partial orderings indicate an attempt of his part to go beyond the Humean paradigm of rationality and to reflect on the way preferences are formed and constructed through rational deliberation.

This article is divided into six sections. The next section presents Sen’s critique of revealed preference theory and in particular the impossibility to define rationality according to some criteria of internal consistent of choices. In the third section, I introduce the concepts of commitment developed by Sen to account for counter-preferential choices. The fourth section draws a parallel between Sen’s critique of revealed preference theory with Searle’s attack on the “classical model of rationality”. I argue in the fifth section that both critiques can be met with an appropriate definition of preferences as total comparative evaluations. However, I show in the sixth section that Sen’s and Searle’s relevant target is the assumption that rational
agents enter in social interactions with prior and complete preference orderings. I briefly conclude in a last section.

2. Sen’s Critique of Revealed Preference Theory and of Internal Consistency Conditions

Sen started his attack on revealed preference theory with a series of papers published during the seventies ((Sen 1971); (Sen 1973); (Sen 1977)). Revealed preference theory 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 on which I return in the next section. However, Sen has latter extended his critique to the more general concept of rationality on which economics entirely relies.\(^1\) I present this critique first it helps to make more intelligible what Sen was trying to achieve with his concept of commitment.

In the introduction of his 2004 book *Rationality and Freedom*, Sen states that there are three different understandings of the rationality of choice assumption in economics (Sen 2004, p.19): a) rationality as internal consistency of choice, b) rationality as self-interest maximization, c) rationality as maximization in 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. Revealed preference theory proposes a list of more or less strong axioms defining what it takes for choices to be consistent. An example is the Weak Axiom of Revealed Preference (WARP) which states that, given that two commodities \( x \) and \( y \) are available, if a person reveals a strict preference for \( x \) over \( y \), then he must not also reveal a strict preference for \( y \) over \( x \). Formally,

\[
\text{Weak axiom of revealed preference: for any two } x \text{ and } y, [x \in C(S) \& x, y \in S \& y, z \in S] \Rightarrow x \in C(S') \text{ where } x \in C(S) \text{ accounts for “} x \text{ is chosen from the set } S \text{ of alternatives.}
\]

As Sen (1973) notes, while WARP is a criteria of consistency of two choices only, it also guarantees the stronger condition of the transitivity of preferences, *i.e.* to prefer \( x \) over \( y \) and \( y \) over \( z \) entails preferring \( x \) over \( z \).\(^2\) Transitivity of preferences is, along with the axiom of completeness, a constitutive axiom of ordinal utility theory (Hausman 2011). As such, it constitutes a normative benchmark against which the rationality of an agent is assessed; anyone whose behavior violates transitivity is considered as “irrational”.\(^3\) Two weaker axioms

\(^1\) 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 2004). All page numbers in this section refer to the latter.

\(^2\) Non-transitivity of preferences necessarily entails a violation of WARP. Assume that a person \( i \) has revealed the following preferences: \( x \succ_i y \succ_i z \) and \( z \succ_i x \), thus violating transitivity. The same person having to choose between \( x \), \( y \) and \( z \) will necessarily violate WARP: if he chooses \( x \), then he has demonstrated at the same time a preference for \( x \) over \( z \) and for \( z \) over \( x \), and the same inconsistency results for the choice of \( y \) or \( z \).

\(^3\) As Hausman (2011) notes, microeconomics textbooks diverge regarding the interpretation of the axioms of ordinal utility theory. Some textbooks regard these axioms as true descriptions of people’s actual preferences,
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, p.128):

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

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

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 choose from any subset \( T \) which contains \( x \). The axiom 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 an alternative \( x \) that is weakly preferred to all other alternatives in a set \( S \) will be chosen from \( S \):

**Revealed preference** \((R_c)\): \( xR_cy \Leftrightarrow \exists S: [x \in C(S) \& y \in S] \)

The binariness of the choice function guarantees that an alternative that is weakly preferred from a set of alternatives will be chosen from that set:

**Binariness of the choice function**: for every nonempty \( S \),
\[
C(S) = \{ x | x \in S \& \forall y \in S: xR_cy \}. \]

The combination of the independence of irrelevant alternatives axiom and of the basic expansion consistency guarantees the binariness of the choice function. These axioms, as well as WARP, 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 is it to be a rational agent. One of Sen’s major contributions has been to show that this definition of rationality as the fact of having consistent preferences (and thus making consistent choices) is highly problematic.

Rationality as consistency of choice ultimately relies on the conflation of choice with preference. Indeed, 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, p.243). The difference is that preference relations are assumed to encompass everything that matter to the agent. 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 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

---

while others take the axioms as normative criteria of rationality: a rational person must have a behavior conforming to the axioms of completeness and transitivity.
context. This argument cast doubts on the very relevance of the internal consistency criterion of binariness of the choice function. Take the following general two choice acts:

\[(1) \quad C(\{x, z\}) = \{z\}\]
\[(2) \quad C(\{x, y, z\}) = \{x\}\]

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 (1) and (2) violates both WARP and 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 have 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 WARP and thus of the transitivity axiom. 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, p.129): most of the time, choices are influenced by external factors that are not properly accounted 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, p.130); (1997, p.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 in 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 well
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, p.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 alternatives available to him is dependent from 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. Sen’s concept of commitment, which is also part of his critique of revealed preference theory, goes a step farther in separating choice and preference and in underscoring the importance of factors external to the agent to explain the latter’s choice.

3. Commitment and Counter- Preferential Choice

In his article “Behaviour and the Concept of Preferences” (Sen 1973), 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” (p.253). The distinction between preference and choice is at the core of Sen’s early critique of revealed preference theory. He proposed to distinguish personal preference (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 morality principles (Sen 1977). In some cases, an agent acting under commitment may choose contrary to his personal preferences. Interestingly, Sen started to develop his theory of commitment and counter-preferential choice before his elaborated critique of internal consistency conditions. However, Sen’s concept of commitment makes much more sense (and avoids to falling victim of several misunderstandings) once one acknowledges that it is a logical complement to axioms of inter-menu consistency.

The nature of commitment is well illustrated with the example of the prisoner’s dilemma (see figure 1):

---

4 As I will indicate below, and as Sen as emphasized, 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, p.174 Theorem 3.1). But the identification of choice with preference is lost in the process.
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 make 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. In fact, under the frame of revealed preference theory, 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 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 agent are enduring egoists who care for nothing but their self-interest. As Sen notes, nothing in the conceptual apparatus of revealed preference theory prohibits to endow the agents with other-regarding preferences. Ana’s utility might well depend of Bob’s welfare and it is perfectly right to construe Ana’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 revealed preference theory is still egoistic in some sense (Sen 1977, p.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, p.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 revealed preference theory, 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 \( R_c \) 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 to say 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 the **comprehensive outcome** from the **culmination outcome** (Sen 1997, p.161). Culmination outcome only depends on the ultimate consequences of an agent’s action. Quite the contrary, comprehensive outcome includes 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.

Once again, take the example of the fruit basket. Assume that my preference ordering is the following:

\[
AP, BP, N
\]

with A for apple, B for banana, N for nothing and \(P_i\) for “strongly preferred to”

If I am presented a fruit basket with two apples and two bananas, my preference must correspond to me choosing one apple. Formally,

\[
C([A_1, A_2, B_1, B_2]) = \{A_1\} \text{ (or } \{A_2\})
\]

However, if there is only one apple and two bananas in the basket, it is perfectly sensible to expect that I will choose one of the two bananas:

\[
C([A_1, B_1, B_2]) = \{B_1\} \text{ (or } \{B_2\})
\]

This last choice seems to contradict the above preference ordering. In fact, the fact that I take a banana is due to the act of choice which includes the fact I am following a social norm according to which one should not take a specific fruit when it is the remaining one. Thus, I am committed to follow a social norm which leads me to make a choice that contradicts my personal preferences. The same reasoning applies in the case of the prisoner’s dilemma. Given the players’ preference ordering constitutive of this game, any player has a preference for the outcomes where he defects. Thus, the fact of choosing to cooperate in the prisoner’s dilemma contradicts the preference ordering of the players. Still, one reason the players cooperate may be that cooperation is an unconditional moral imperative and that they are committed to follow it. If the same agent is sometimes cooperating and sometimes defecting in the prisoner’s dilemma, the recovery of the consistency of his choices necessitates to take into
account this commitment to an external factor that in some cases is modifying or displacing the agent’s goals (Sen 1985).

To characterize more fully commitment, Sen ((2004, p. 33-34) (1985, p.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 agent without any sympathy or antipathy for others satisfies the three criteria of privateness. A self-interested but partially other-regarding agent satisfies self-welfare goal and self-goal choice but not self-centered welfare, since some of his goal may be to enhance the welfare of other persons. Things are quite different in the case of commitment, as suggested above. Commitment provides an agent with a reason for action that can be totally independent from his welfare and/or from his goals. As Philipp 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. Commitment is thus a loose concept designed to capture a range of reason for action which cannot be meaningfully reduced to the satisfaction of some set of private preferences. Interestingly, the necessity to recognize this type of reason for actions has also been felt outside economics. The next section discusses the sensible example of John Searle’s theory of practical reason, which incidentally shares many features with Sen’s writings on rationality, commitment and preferences.

4. Searle, Practical Reason and Desire-Independent Reasons for Action

John Searle is a well-known and long standing contributor in numerous fields of philosophical thinking. Searle’s earlier writings have advanced many important ideas in the philosophy of language and in the philosophy of mind. More recently, Searle started to turn his attention to social sciences and to society. His celebrated book The Construction of Social Reality (Searle 1997) constitutes a first step toward the advancement of a complete theory of institutional facts which is still developing. In the meantime, building on his philosophy of mind and his philosophy of language, Searle has also advanced an original philosophical account of practical reason in his book Rationality in Action (Searle 2003). This account is conceived by

5 See also Sen (2004, p.35). The verification 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.

6 See in particular Searle’s last book, Making the Social World (Searle 2010).
Searle as an important step toward the edification of a complete philosophy of society. The leading thread of Searle’s account of practical reason is the critique of what he dubs the “Classical Model of Rationality”. This leads Searle to develop an interesting and novel theory of action and of rationality that is highly relevant for economics.

Searle highlights six assumptions that he thinks are constitutive of the classical model (Searle 2003, p.12-32):

1. Rational action is caused by beliefs and desires.
2. Rationality is about following specific rules.
3. Rationality is a separate cognitive faculty.
4. Weakness of will (also called Akrasia) is impossible unless there is something wrong with the psychological antecedents of action.
5. Practical reason starts with an agent’s primary ends (goals, desires, objectives, purposes); these ends cannot be the subject of rational deliberation.
6. Rationality supposes that primary ends form a consistent set.

Assumptions 1, 5 and 6 are particularly relevant regarding the use of the concept of rationality in the social sciences and especially in economics. The first assumption is surely the most crucial. In the classical model, desires and beliefs are at the same time sufficient efficient causes of action and reasons for action. To explain an action is thus to uncover the desires and the beliefs that caused it. Assumptions 5 and 6 are both related to the first one. They indicate that practical reason resumes into a causal relationship between some given, well-ordered, ends and an action. Consistency of ends is constitutive of rationality because inconsistent ends would lead to contradictions in action. Assumption 5 is reminiscent of the Humean view that “reason is the slave of passions”: ends cannot be said “rational” per se because they are given and cannot be subjected to rational deliberation. Only the rationality of means can be assessed given the set of ends pursued by the agent.

According to Searle, all six assumptions are dubious because they cannot provide a general theory of practical reason. Among the objections made by Searle, the major one is that action cannot be explained entirely by desires and beliefs. Searle offers two complementary arguments to this claim. First, desires and beliefs are not and cannot be causally sufficient to determine the action. This is what Searle calls the gap between the “causes” of action and the action proper. The gap is just a shorter name to what is traditionally called the “freedom of the will”. The very possibility of rationality and of rational action presupposes free will: “We cannot avoid the presupposition, because even a refusal to engage in rational decision making is only intelligible to us as a refusal if we take it as an exercise of freedom” (Searle 2003, p.14). The gap occurs at three distinct moments in the decision process (Searle (2003); see also Searle (2010, p.38-41)): first, there is a gap between the reasons for action (beliefs, desires and beliefs).

7 Other assumptions are also important but less central given the object of this paper. Assumption 2 presupposes the possibility to discover a “logic of practical reason” made of a set of rules of inference analogical to *modus ponens* or *modus tollens*. Assumption 3 is of a particular interest for psychology and particularly for evolutionary psychology. Searle’s account rejects the existence of a separate “module of rationality” of any kind in the brain. Assumption 4 relates to issues now tackled by behavioral economists. In the classical model, weakness of will is non-sense because psychological antecedents of action are considered as sufficient conditions for action. Action must follow from these antecedents. Searle argues that this assumption is wrong on the ground that there is a gap between the psychological causes of action and the action proper.
desires, and other kinds of reasons such as obligations and needs) and the decision. Second, the gap occurs between the decision and the action. Finally, the gap occurs for a third time between the initiation of the action and its continuation through time.

Because of the gap, any account of practical reason and of rational decision cannot reduce into a mere summation of the desires and of the beliefs entertained by the agent at any point. The second argument is even more important for the applied social scientist: contrary to the classical model, Searle argues for the existence of desire-independent reasons for action. This leads to a rejection of Hume’s view of the relationship between passion and reason at two different levels. Firstly, passions (or desires) are never causally sufficient reasons for action because other “motivators” can make a decisive impact upon the agent’s behavior. These other motivators are typically external to the agent’s motivational set, i.e. they are not epiphenomenal to a psychological element that is entirely located inside the brain of the agent. The point is that these external motivators come from the external world and are recognized by the agent as a rationally compelling reason for acting in some way or another. Borrowing one of Searle’s favorite examples, imagine I go into a bar and order a beer. After having drunk the beer, the waiter brings me the bill. According to Searle, I have then a compelling and rational reason to pay the bill even though my intrinsic desire would be to leave without releasing any cent from my wallet: “we understand that when you ordered the beer and drank it, if you are a sane and rational person, you were intentionally creating a desire-independent reason, a reason for doing something regardless of what was in your motivational set when the time came to do it” (Searle 2003, p.27-28). The bulk of Searle’s 2003 book is to provide an argument for the possibility and the inevitability of such a kind of desire-independent reason for action. Desire-independent reasons for action are generated by commitments which are themselves constitutive of speech acts. For instance, if I promise to you that x, then the mere speech act “I promise you that x” commits me to act such that x. This is so because such an utterance expresses a meaning intention imposing conditions of satisfaction on conditions of satisfaction: the first set of conditions of satisfaction is the utterance itself (to promise x to you I must use a linguistic device to promise); the conditions of satisfaction over this set are the truth conditions. To promise somethingrationally implies an intention with conditions of satisfaction. Hence, the promise creates a commitment which itself generates a desire-independent reason for action. Promises are a particular instance of a more general practical law that Searle calls the semantic categorical imperative: “When you make an assertion of the form a is F, rationality requires that you be able to will that everyone in a similar situation should assert that a is F” (Searle 2003, p.159). Notwithstanding the

---

8 An obvious counterargument is that the obligation to pay the bill is not really external. I may pay the bill because I desire to fulfill what I recognize to be an obligation. Searle notes this objection and replies that the desire to fulfill the obligation cannot exist without a prior rational recognition of the obligation. In other words, the recognition of the obligation precedes the formation of a desire and it is thus impossible to explain why I pay the bill without taking into account this external motivator.

9 See in particular Searle (2003, chap.6). Searle’s argument essentially builds on his philosophy of language. According to Searle, we create commitment through the use of a linguistic device: the utterance of sentences results in imposing conditions of satisfaction (truth conditions) on conditions of satisfaction (the utterance proper). The utterance thus becomes a status function, i.e. it attributes to itself a new status in a self-referential way.
difficulties undermining this account, it provides the ground for distinguishing a special set of reasons for action that cannot be merely reduced to first-order desires.

The Humean view is rejected by Searle at a second level. According to the Humean view, rational deliberation regarding the passions is impossible. Reason always operates on the behalf of given passions. According to Searle, the classical model of rationality materializes this postulate by assuming that prior engaging in any deliberation and action the agent is already equipped with a well-ordered set of preferences. A standard follow-up is then to argue that *de gustibus non est disputandum*, *i.e.* no rational argument can be made regarding the desirability of a preference ranking over another one. But in fact, as the previous paragraph has suggested, many desires are grounded on rational deliberation because they cannot exist without a prior recognition of an external reason to action. Moreover, in many cases, external and internal reasons will be in conflict. Practical reason is precisely about the way rational agents are able to overcome such kind of conflict to make decision and to act.

The similarity between Sen’s concept of commitment and Searle’s idea of desire-independent reason for action is striking. Both authors make the same argument that (rational) choice and action may not be reducible to preferences representing intrinsic desires or self-welfare goals. Moreover, both locate the origins of external reasons for action partly in institutional, moral and normative imperatives. More to the point, Searle’s critic of the Humean view regarding the sovereignty of passions over reason matches another dimension of Sen’s discussion of rationality and preferences. As I will show below, Sen’s concept of commitment is also deeply linked to his account of the incompleteness of preference orderings leading him to distinguish optimization from maximization. The very idea of commitment implies that the rational agents have the ability to recognize external and new reasons for action, thus helping to augment incomplete preference orderings. Interestingly, most commentators of Sen’s writings have put less emphasis is this particular point than on the overall relation between commitment and preferences. I discuss in the next section the way Sen’s concept of commitment has been related with the standard understanding of preference and rationality in economics, and I return then to what I argue is the core idea of both Sen’s and Searle’s theory of rationality: that preference orderings are fundamentally incomplete and evolve through public discussion and reasoning.

5. Rationality and Preferences in Decision Theory

According to Searle, the ultimate failure of the classical model of rationality lies in its endeavor to produce a reductive causal account of action by referring solely to desires (and beliefs). According to Sen, revealed preference theory fails because it assumes that choices perfectly reflect the agent’s preferences which themselves are a representation of his personal

---

10 Searle’s linguistic argument for the existence desire-independent reasons for action has been strongly criticized, e.g. Viskovatoff (2003) and Schmid (2003). The latter rightly suggests that Searle’s account would have been far more persuasive if, instead of the semantic categorical imperative, he had used the concept of collective intentionality that is pivotal in his theory of institutional facts. I have introduced Searle’s collective intentionality in a game-theoretic framework elsewhere (Hédoin 2013).

11 By the way, Sen (2004, p.25) mentions Searle’s book in a footnote regarding the limitation of the reduction of rationality to instrumental reasoning.
welfare. In both cases, the failure allegedly lies in the attempt to reduce choice and action to desires and preferences narrowly constructed. However, the validity of this critique has been questioned on the basis of an accurate reading of the true analytic content of the concept of preference in decision theory (e.g. (Hausman 2008); (Hausman 2011); (Petit 2008); (Ross 2005)). In this section, I argue that even though critics may be right that the preference concept may encompass commitment and desire-independent reasons for action, they fail to acknowledge that a significant part of Sen’s (and Searle’s) discussion is more about the incompleteness of preference orderings and that rationality is also about how individuals construct their preferences through reasoning and public deliberation.

There are several prevailing interpretations of the preference concept as used by economists. In a recent book on the subject, the philosopher Daniel Hausman (2011) argues convincingly that in the actual practice of applied economists, the concept of preference is used as a total comparative evaluation where everything which is judged relevant by the agent for making a choice is taken into account. As a result, preferences are not reducible to desires whatever the way we may define them. A crucial difference is this: while desires are inputs that enter into the deliberation process and that agents may balance against other motivations or reasons for action, preferences are the product of an overall comparative assessment taking into account all the reasons for action (Hausman 2011, p.6). Properly understood, the concept of preference is not only far more general and encompassing than the concept of desires (desires are only one of the possible sources for preferences). They do not have the same analytical status: desires are primary data that may help to explain one’s choice (he chooses to do... because he desires that...) while preferences are theoretical statements describing how an agent assesses and compares various alternatives. The main implication is then that nothing in principle makes impossible to include in one’s preferences ordering reasons for action other than desires.

What should we make of Sen’s argument that, as mere representations of one’s personal welfare, preferences cannot be identified to actual choices because other factors such as commitment enter the party? Strictly speaking, if Hausman is right that preferences in economics are total or overall ranking of everything that is deemed to be relevant for the agent, then Sen’s all argument is almost wrong by definition. Indeed, as well as concerns for others (sympathy or altruism) must be included in one’s preference set and thus must be reflected in one’s utility function, various kinds of commitment should enter as well in a person’s overall comparative evaluation of outcomes. Then, because preferences are full descriptions of everything that the agent considers to be relevant, it appears that the discrepancy between choices and preferences disappears.

The example of the basket of fruits of the preceding sections helps to make this point clearer. On a superficial examination, one may conclude on the inconsistency (and hence, the irrationality) of the person who chooses one apple (x or y) in the first case with two apples in the basket but who prefer to have nothing (z) in the second case with only one apple in the

---

12 Hausman uses Sen’s distinction between culmination outcomes from comprehensive outcomes discussed above (Hausman 2011, p.49-55). The former refers to the physical or material consequences of an action captured by the payoffs that appear at the terminal node of a dynamic game. The former includes everything that is relevant, including the path through which a particular node has been reached.
basket. Apparently, WARP is violated because this person prefers in one case $x$ (or $y$) over $z$ while in the other case she prefers $z$ over $x$. But it can be objected that we arrived at this conclusion only because we assumed that $x$ and $z$ are exactly the same alternatives in both cases. In fact, our conclusion relies on a confusion about the nature of the reference classes over which utility functions and preference sets are constructed (Ross 2005, p.134). The point is simply that the choice to take apple $x$ or $y$ in the first case (denote these two choices $x_1$ and $y_1$) are different choices than the choice to take the apple ($x_2$) in the second case. As soon as one realizes that $x_1 \neq x_2$, no inconsistency remains. It is perfectly consistent to prefer $x_1$ and $y_1$ to $z$ while to prefer $z$ to $x_2$ (provided of course that $x_1$ is preferred to $x_2$). As a result, we are lead to define a richer preference relation $R_i$ of an agent $i$ by making it conditional on the choice set $S$. Because preferences are total comparative evaluations of outcomes, they should include everything that is relevant from the person’s point of view, including the menu of alternatives the chooser is presented with. Accordingly, in this example, the context in which the choice is made (whether there are one or two apples in the basket) is fully relevant to understand the choices made. Thus, it should enter into the person’s preference ordering.\footnote{Note that this not a violation of another axiom that is sometimes stated explicitly, the axiom of context independence. Indeed, in this axiom “context” means “everything that is not relevant to the agent’s choice”. In the example of the text, what I call the context is an integral component of the choice function.}

This formal twist has been criticized on the basis that it makes revealed preference theory either tautological or irrefutable ((Basu 2000); (Hodgson 2012)). The fact is that it is virtually always possible to refine the agents’ preference orderings and choice functions such as the observed choices are made consistent. Because two decision problems are never fully identical, it is always possible to include an additional feature in the choice function to recover the consistency of choices condition. The same is true for strategic interactions: if you design a prisoner’s dilemma setup and the two persons you asked to play through cooperate, then according to revealed preference theory this observation is not a refutation of game theory. It only means that other considerations than monetary gains enter in the persons’ utility functions and preference orderings. Actually, these two persons are not playing a prisoners’ dilemma but another game where cooperation is rational by definition ((Binmore 1994); (Ross 2005)). As long as revealed preference theory is conceived as no more than a descriptive tool, its rejection for being “tautological” or “irrefutable” misses the point: it merely provides a formal language that permits insightful analytical derivations (Binmore 2009). Nevertheless, this interpretation still suffers from two insufficiencies.

First, in spite of contrary claims made by some game theorists, revealed preference theory does not allow to identify preferences with choices. This is so because additionally to preferences, beliefs contribute to determine choices. Indeed, the classical model of rationality is fundamentally consequentialist in the sense that choices are formed by the conjunction of preferences and beliefs (Hausman 2011).\footnote{Hausman (2011, p.44) distinguishes distal preferences from final preferences. Distal preferences correspond to the economists’ traditional understanding of a preference ordering as a ranking over every possible states of the world and hence over outcomes. Quite differently, final preferences are only defined over the object of choices, that is over actions. By definition, final preferences are identical with choices but this is not the case for distal preferences.} This becomes obvious as soon as an agent faces uncertainty regarding the consequences of her choices. In this case, to decide what to do, the
agent must not only considers what he prefers to be the case (his “distal” or “fundamental” preferences), he must also use the (possibly subjective) probabilities that a particular state of the world obtains. The choice will thus not reflect only this person’s preferences. This is even clearer in a game-theoretic situation where these times there is uncertainty regarding what others will do. The prisoner’s dilemma is not a good example because if the players’ fundamental preferences are such that the interaction has the structure of a prisoner’s dilemma, then everyone will defect. But for any game with no dominant strategy, beliefs about what others will do are relevant: the fact that I drive at the right side of the road does not reveal that I prefer to drive at the right side rather than at the left side; it is rather the consequence of the fact that I believe that everyone will be driving at the right side. To explain choices, economists have to derive them from both preferences and beliefs.

The second limit is more fundamental. The understanding of preferences as total comparative evaluations leads to a similar conclusion than Sen’s critique of the internal consistency conditions approach of rationality: consistency of choice depends on external factors such as the identity of the chooser, the content of the menu of available alternatives or the accounting of social or moral requirements. The claim that all these factors may be introduced in the preference relation attached to an agent to recover the identity between choice and preference is not in contradiction with Sen’s concept of commitment as a paradigm for counter-preferential choice. Sen (1997, p.173-5) fully acknowledges that it is still possible to refine and to enrich a preference ordering to recover (inter-menu) consistency of choices: formally, a ranking of menu-dependent preferences $R^s$ can generate a menu-independent choice function $C(S)$. In this sense, a choice function (expressing what Hausman calls “final preferences”) may encompass everything that matter for an agent who is choosing. This may explain the general skepticism Sen’s concept of commitment has generally met. The idea that an agent may act according to goals other than his seems irrelevant since formally it will almost be always possible to construct a choice function corresponding to a well-ordered preference ranking (Pettit 2008). As a result, it is sometimes argued that the commitment concept is redundant and that one and unique concept of preference will do (Hausman 2008). However, if preferences are total comparative evaluations and if the consequentialist interpretation of revealed preference theory is accepted, then it seems revealed preference theory has nothing more to offer than a mere re-description of choices in terms of preferences and beliefs. In this sense, an examination of the mechanisms through which preferences are formed is required to give revealed preference theory an explanatory power.

I will argue in the next section that, though mostly ignored by critiques of Sen’s account of preferences and commitment, the main insight of Sen’s discussion is that it shows the necessity to reflect on the way rational agents build and complete their preference ordering through public deliberation and reasoning. This echoes perfectly Searle’s attack on a key aspect of the Humean paradigm where reason is seen as the slave of passions. Interestingly, this suggests that while Sen’s critique of the rationality concept is generally understood as focusing on the axiom of transitivity of ordinal utility theory, in fact its main target may be the axiom of completeness. A proper understanding of commitment necessitates seeing it as a reconsideration of both axioms.
6. Commitment and the Incompleteness of Preference Orderings

Sen’s concept of commitment is not only related to his critique of the idea that rationality consists in meeting the demand of some internal consistency conditions. It also corresponds to a rejection of the approach identifying rationality with the maximization of a quantity expressing the agent’s self-interest. However, Sen makes it clear that he does not intend to reject the idea of maximization. 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 2004, p.32, p.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 2004, p.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 2004, p.36).

“Rationality of choice, in this view, is primarily a matter of basing our choices – explicitly or by implication – on reasoning that we can reflectively sustain if we subject them to critical scrutiny” (Sen 2009, p.180).

This idea of rationality as one’s ability to reason and to reflect on one’s choices is clearly reminiscent of Searle’s claim that the Humean view of rationality wrongly assumes that an agent enters into an interaction with an already complete and well-behaved preference ranking (Searle 2003, p.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 2003, p.253). Sen relates this issue to the idea of the incompleteness of preference orderings. But I suggest that this issue is also fundamentally linked to commitment. Therefore, commitment is not only a concept responding to the difficulties resulting from the association of rationality with the transitivity axiom; it is a also deep questioning of the completeness axiom.

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 strict preference xPy (or yPx), or of weak preference xRy (or yRx), or of indifference xIy. 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 2004, p.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 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

$$
\begin{align*}
-xRz \\
-zRx
\end{align*}
$$

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 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 $M(R, S)$:

$$
B(R, S) \subseteq M(R, S)
$$

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 $y^{15}$ 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 enters 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 2004, p.17); (Sen 2009, p.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.

Incompleteness issues are deeply related to Sen’s latest interest for public reasoning for 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 regarding 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

---

15 This means that neither $xRy$ nor $yRx$ is true.
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.\textsuperscript{16}

This ability to redefine and to complete a preference ranking is what both Sen and Searle refer to \textit{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); (2004)) formalizes this intuition. But this should also lead to a reconsideration of the commitment concept. As I have argued above, the very notion of commitment implies that the agent is correlating is behavior to something that is not reducible to what he intrinsically prefers. This implies in turn that the agent has the ability to \textit{recognize} a reason for action that is external to his intrinsic desires. This recognition is rational\textsuperscript{17} because it depends on the agents reasoning abilities. External reasons for action are almost surely never determined \textit{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 norms or a moral requirement must have, before choosing to cooperate, recognize 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. But 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 prefers 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 agent because of his commitment to follow a moral requirement. To relate commitment to incompleteness and not only to consistency issues thus help to make clearer a key distinction in Sen’s account of rational behavior and to demonstrate the centrality of reasoning abilities in a relevant account of rationality.

7. Conclusion

Amartya Sen’s writings on rationality, preference and consistency of choice have produced many insightful results contributing to enhance our understanding of the implications of the economic theory regarding the human behavior. However, his concept of commitment as

\textsuperscript{16} 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)).

\textsuperscript{17} Searle (2003) simply speaks of \textit{rational recognition}.
counter-preferential choice has never really caught among economists. The reason for this is that commitment is generally regarded as related to Sen’s discussion of conditions of internal consistency of choice. In this perspective, many economists tend to argue that the preference concept is sufficiently malleable to incorporate “external reasons for action”. Sen’s important claim that consistency of choice is conditional on taking into account external factors such as social norms or moral requirements is not wrong, but parsimony should lead to reject the use of an additional concept such as commitment.

The idea that a unique concept of preference may be sufficient to account for a variety of reasons for action is probably valid. However, the parallel I have made between Sen’s writings on commitment and Searle’s critique of the classical account of rationality shows that Sen is not only concern with the problem of consistency of choice. Even though this is most often ignored, the commitment concept may be associated to Sen’s discussion of another important axiom of utility theory, the axiom of completeness. This axiom is related to the Humean assumption that “reason is the slave of passions”: rational individuals have already a pre-established complete preference ranking before entering in an interaction or before facing a decision problem, and this ranking is not subject to rational scrutiny. Sen not only argues that rationality and maximization are compatible with incompleteness. He also suggests that rationality entails the ability to reflect on one’s preferences, in particular when the ordering is initially incomplete. The possibility for a rational agent to commit relies on this ability: commitment implies to recognize external reasons for action that were previously ignored. The commitment concept thus underlines the fact that one’s preference ordering may be never fully complete.

References


