Rationality, Behavioural Economics and Amartya Sen

S. R. Osmani[*]
Ulster University, UK

I. Introduction

A rapidly growing body of research in behavioural economics has emphatically demonstrated that actual behaviour of people in economic and related arenas do not often conform to the patterns that are assumed and/or predicted by standard economic models. Such non-conformist behaviour goes by the name of 'anomalies', a term made famous by a feature of the same name written for many years by Richard Thaler, the doyen of behavioural economists, in the *Journal of Economic Perspectives*. Over the years, so may 'anomalies' have accumulated that they can no longer be brushed aside as minor aberrations. Taken together, these anomalies amount to a serious challenge to the models of decision-making standardly employed in economics and related disciplines.

There remains some confusion, however, over exactly what this challenge amounts to. At an empirical level, it is clear that the challenge relates to the predictive power of some specific models of economic behaviour – the prime targets being the expected utility model as a description of behaviour under uncertainty and the model of exponential discounting as a description of how people make inter-temporal choice. But at a more foundational level, behavioural economics also claims to challenge the 'rational choice theory' underlying such models (e.g., Jolls *et al.* 1998; Rabin 1998; Thaler 2015). It is at this level that, I believe, a good deal of confusion prevails. Since rational choice, in some sense, underpins all economic models, the anomalies related to these models are also often claimed, by extension, to imply deviations from rational behaviour. It is this claim, in particular, that I wish to examine in this paper.[1]

For this purpose, I use a rather idiosyncratic methodology – I examine the claims of behavioural economics in the light of Amartya Sen's critique of rational choice theory. Over the years, Sen has examined the assumptions and content of rational choice theory through a series of outstanding contributions, some of which (such as his *Rational Fools* paper of 1977) have become part of the folklore of economic as well as philosophical literature. In fact, it would, I believe, be fair to say that, although the ideas of rationality and rational behaviour have been searchingly explored by many eminent scholars, no one has done more than Sen to subject the rational choice theory to a critical scrutiny. It would, therefore, make eminent sense to use Sen's ideas to examine the validity of the claims about deviations from rational

[1] It should be clarified that there are a number of other debates surrounding behavioural economics that I do not go into here. These include: the empirical robustness of the anomalies; how important these anomalies really are for the decisions economic agents have to make in the real word ; and the lively debate on the policy implications that focusses on what some behavioural economists call 'libertarian paternalism'.

behaviour made by behavioural economics. I first present a brief but self-contained account of (my understanding of) Sen's critique of rational choice theory in Section II. Then, in Section III, I critically examine a number of major anomalies identified by behavioural economics and try to assess, through the lens of Sen's analysis, how far they actually challenge the precepts of rational choice theory underlying standard economic models.[2] To anticipate my conclusions, I find that (a) many of the anomalies do not pose any challenge at all to even the narrowest version of the rational choice model, (b) some do challenge the narrow versions but not the broader ones, and (c) a few genuinely call for extending rational choice foundations of economic theory more radically – especially, in ways that Sen has been advocating for a long time. Some concluding remarks are offered in Section IV.

II. Sen on Rationality and Rational Choice

Sen takes a very broad view of rationality. Thus, in his Introduction to the book *Rationality and Freedom*, he states "Rationality is interpreted here, broadly, as the discipline of subjecting one's choices – of actions as well as objectives, values and priorities – to reasoned scrutiny." (Sen 2002: 4). His essential point is that the standard rational choice model is much too narrow, in the sense that it doesn't make enough demand on the so-called 'rational' human being to apply her reason – specifically, to subject her choices and actions to reasoned scrutiny. This narrowness is partly an act of omission – some of the considerations that reasoned scrutiny ought to demand are not allowed in the model; but partly it is also an act of commission, in the sense that some of the conditions that the model tends to impose as 'requirements' of rationality actually inhibit the exercise of reason.

There are some variations in the way rational choice is interpreted in mainstream economic theory, but a couple of features have been claimed to be essential for behaviour to be rational – namely, (a) internal consistency of choice, and (b) maximization of a person's objective function.[3] Sen has insightfully critiqued both these features. But the broad view of rationality he adopts – namely, the exercise of reasoned scrutiny – can accommodate both consistency and maximization, appropriately defined. His critique is rather directed to the unduly restrictive ways in which these two notions are specified in standard theory. I shall organize my account of Sen's critique of 'rational choice theory' along three lines: (a) the nature of consistency conditions, (b) contents of the objective function, and (c) the nature of the exercise of maximization.

The Nature of Consistency Conditions

It is useful to begin by noting that in a rational choice model one can take either preference or choice as the primitive concept. In the classical models of utility maximization, preference is taken as the primitive concept, and it is invested with a specific content – for example, we could prefer outcomes that yield greater pleasure or satisfaction (going under the

---

[2] At least on one occasion, Sen (2002:26-23) himself undertook a brief analysis of this kind, focussing mainly on the influential paper by Jolls *et al.* (1998). My analysis builds upon his work, examining a larger number of anomalies and at greater length, while also deviating slightly from his analysis on some occasions.

[3] Depending on the exact specification of the model, the two conditions may be applied either singly or in combination.

name of utility), as in the hedonistic version of utilitarianism. Choice is thus a derived variable in this model; consistency of choice in this case would be judged in relation to the external reference of maximising an underlying preference ordering.

In revealed preference theory, in contrast, choice is taken to be the primitive concept, and preference ordering is taken simply as a formal representation of choice behaviour. Consistency is then defined as a property of choice itself, without any reference to any underlying motivation (unlike in the classical models of utility maximization). For example, the 'weak axiom of revealed preference' requires that if I have chosen x when an alternative y was available, then on a different occasion I cannot choose y when x is also available; if I do I would be deemed inconsistent and irrational.

The problem with defining rationality in terms of internal consistency lies in this very disregard for the motivations underlying choice. To give an example made famous by Sen, if I am given a choice between 'picking up the last apple on the table' (x) where there are other guests around and 'not having any apple at all' (y), I might choose y based on my objective function which attaches importance not just to my consumption of apple but also to my desire not to appear greedy. However, if there were to be more than one apple on the table, I might have happily picked up one for myself (i.e., chosen x instead of y), while being perfectly consistent with my objective function. I would then be violating the weak axiom but would not be irrational in any fundamental sense.

The crucial feature of this example is that the two apparently contradictory choices that have been made relate to two different 'menus'. If we disregard the menus, the actual choices (x on one occasion and y on another) may appear inconsistent; but as soon as we bring the menus within the purview of our evaluation and consider our choices in the light of our motivations, the apparent inconsistency immediately disappears. Sen describes this phenomenon as 'menu-dependence' of choice.[4] Many different motivations may give rise to menu-dependence of choice.[5]

Several implications of menu-dependence of choice are worth emphasizing in the present context. First, the so-called 'conditions of internal consistency' rule out menu-dependence by not allowing consistency to be judged in relation to the external reference of one's objective function. In so doing, they inhibit the exercise of reason. Second, 'choice reversal' should not necessarily be interpreted as 'preference reversal' unless choices are equated with preference by definition as in revealed preference theory. Third, which is a corollary of the preceding two points, when a person is observed to have made a 'choice reversal' she may simply have behaved in a menu-dependent way, which, as we have seen, is fundamentally rational, when preference is accorded an independent content. Fourth, to say that people violate internal consistency by behaving in a menu-dependent way is not the same thing as saying that they are not 'maximizers'; internal consistency is neither necessary nor

---

[4] The underlying idea, though not the phrase 'menu-dependence', can be found as early as in Sen's (1973) classic engagement with revealed preference theory. Formalization of the concept of 'menu-dependence' (in the context of individual decision-making) and explorations of its varied motivational sources as well as implications for behaviour can be found in Sen (1993, 1997).

[5] See Sen (1993) for illustrations of various sources of menu-dependence when choices do not involve uncertainty. Elsewhere, Sen (1985a) shows that menu-dependence can also arise in multiple ways in the presence of uncertainty, and thereby violate the 'axiom of strong independence' – the internal consistency condition that underpins the expected utility approach.

sufficient for maximization of an objective function to be possible and meaningful. As Sen explains, "The ordering of the alternatives on the basis of which he or she is maximizing varies with the menu, but this does not deny that for each menu there is a clear and cogent ordering – the basis of the maximizing decisions. So the conditions he is violating, which are standardly presumed to be necessary conditions for maximization, need not be taken to be so for a broader interpretation of maximization."[6] (Sen 1993: 501)

Content of the Objective Function

The rational 'economic man' is supposed to behave in a way that 'maximizes' his 'utility'. Much of economic theory invests the concept of utility with a very specific content – namely, self-interest, which itself may be defined more or less narrowly. To understand the precise nature of Sen's critique of this aspect of rational choice theory, it is useful to distinguish, following him, three different ways in which consideration of self-interest could enter into a person's decision-making process.

> *(A) Self-centred welfare*: A person's welfare depends only on her own consumption and other features of richness of her life.
>
> *(B) Self-welfare goal*: A person's only goal is to maximize her own welfare.
>
> *(C) Self-goal choice*: A person's choices are based entirely on the pursuit of her own goal.

There are three questions implicit in this classification: (a) what does my welfare depend upon, (b) what are my goals, and (c) what does my action (choice) depend upon? In the dominant version of rational choice model, all three questions are answered in a 'purely 'selfish' manner. In this model, the so-called rational person takes her welfare to depend solely on her own consumption and nobody else's (self-centred welfare); she defines her goal as maximizing only her own welfare and nobody else's (self-welfare goal); and she chooses her action so as to maximize her own goal to the exclusion of all other considerations (self-goal choice). This three-dimensional selfishness is the hallmark of the dominant version of rational choice theory.

If rationality is deemed to consist in reasoned scrutiny of one's values and actions, there is no reason to suppose that all three questions must be answered exclusively from the selfish perspective. Reasoned scrutiny might lead to some relaxation of the purely selfish perspective in any of the three dimensions – and indeed in any combination of the three dimensions – without denying the importance of self-interest altogether. Of course, it is possible that I might decide to pursue pure self-interest in all three dimensions, in line with the standard model of rational choice, even after careful scrutiny. This possibility cannot be ruled out, but the problem with the standard model is that this is not a mere possibility but an inexorable certainty, by assumption. People are simply assumed to pursue pure self-interest in all three dimensions.

In more recent times, some parts of economic theory have begun to broaden the scope of rational choice theory by abandoning the assumption of self-centred welfare. This

---

[6] The reference to 'broader interpretation' in this sentence relates to the distinction between culmination outcomes and comprehensive outcomes discussed below under 'the nature of maximization'.

broadening has occurred principally through the pioneering works of Gary Becker (1976, 1996), who has systematically incorporated the role of altruism in particular and other-regarding concerns in general in a rational person's utility calculus. However, even after this broadening, the assumption of pure self-interest is still retained, by definition, in the other two dimensions. A rational person is still assumed to accept (uncritically) that her goal is to maximize her own welfare, even if it depends partly on other people's consumption as well (self-welfare goal) and that her choices will be guided purely by her own goal and nothing else (self-goal choice). This is precisely where Sen takes issue with the role of 'self-interest' in the current state of rational choice theory.

He is especially critical of the assumption of self-goal choice. He finds no great difficulty in the assumption that a rational person may have the goal of pursuing her own welfare, but he considers the insistence that her choices must be guided *solely* by this goal to be deeply problematic. This insistence, he argues, renders the contents of a rational person's objective function unduly narrow – because, if reasoned scrutiny of one's actions were allowed full play, a person might well want to base her choices on a broader set of considerations, which will very likely include her self-welfare goal, but would also include other considerations.

Sen uses the concept of 'commitment' to refer to all these other-regarding considerations that might act as limitations on the pursuit of self-goal choice. Commitment is distinguished from the concept of 'sympathy', which is also an other-regarding concern but the two concerns work in very different ways in affecting behaviour. Sen defines sympathy to mean that one person's welfare is affected by other people's welfare – its role is thus to broaden the concept of welfare as perceived by a person (as in Becker's works). But sympathy does nothing to deflect a person from pursuing single-mindedly the goal of maximizing her welfare (albeit after taking into account the impact of other people's welfare on her own); self-goal choice still reigns supreme. Only commitment can impose limitations on the pursuit of self-goal.[7] Rationality, as 'reasoned scrutiny', demands that while deciding on some choice of action, a person will weigh up the importance of her welfare against the importance of the commitments she values and take her decision accordingly. But, by insisting on self-choice goal, the current models of rational choice rule out the independent role of commitment and thus do not allow full exercise of reasoned scrutiny. This, according to Sen, is one of the most fundamental limitations of rational choice theory as currently employed in economics and related disciplines. The reach of rationality ought to be broadened to allow a role of commitments in addition to the role of personal welfare in the process of decision-making.

Commitments can of course be indirectly accommodated in the current models of rational choice by allowing the possibility that a person's ability or failure to keep her commitment may impinge on her welfare by affecting her sense of satisfaction when she succeeds in keeping her commitments and dissatisfaction when she fails. To that extent, commitment can be recognised without compromising self-goal choice. Sen's argument, however, is that this would be an inadequate recognition of the importance of commitment because in the real world people are often moved by commitments to take actions because they value them in ways that go beyond the mere satisfaction they derive from them. They do it

---

[7] The importance of commitment, as distinct from sympathy, in rational decision-making was first highlighted in Sen (1977) and developed further in Sen (1985b). For a more recent contribution, see Sen (2005).

because they just feel they ought to do it. To put it in philosophical jargon, people may attach deontological value to commitments that is quite independent of the consequences they may have, including the consequences for their own welfare; that's why people often make great personal sacrifices for the sake of commitments. In order to give commitment its full due, we must, therefore, move beyond the dictates of self-goal choice.

The Nature of Maximization

The idea that a rational person would choose her actions in a way that maximizes her objective function is a central premise of rational choice theory. Sen agrees that there is an inescapable logic in the idea that a rational person would want to enhance her objectives as much as possible, given the constraints she faces, provided the objectives are not tied too narrowly to the conception of 'self-interest' as discussed above. However, in addition to this proviso, Sen also argues that the notion of maximization itself needs to be broadened in several directions, if we are to give reasoned scrutiny its full due as the hallmark of rational behaviour.

First, while maximizing, a rational person is expected to scrutinize not only the final outcomes of her actions but also the process through which those outcomes are achieved. The essential point here is that a person's preference ordering over alternative outcomes may depend not just on the nature of the outcomes but also on various features of the very act of choice itself. As a consequence, rationality demands that maximization cannot proceed without paying due attention to the process of choice.

In order to accommodate the significance of process in the maximization exercise, it is important to make a distinction between what Sen calls 'culmination outcomes' and 'comprehensive outcomes'. Culmination outcomes are simply the standard outcomes over which a person is supposed to have preferences (for example, commodity bundles in consumer theory); standard maximization would proceed on the basis of these preferences. Comprehensive outcome, by contrast, is defined as the combination of a culmination outcome and the process through which it was achieved. Giving importance to the choice process demands that a rational person scrutinizes her preference rankings over the comprehensive outcomes, instead of culmination outcomes, and maximizes accordingly.

Second, rationality does not simply consist in maximizing on the basis of 'given' preferences, it also involves scrutinizing the preferences themselves. In the process of this scrutinization, one would have to ask: do I really like the preferences that I happen to have or would I want to have a different set of preferences? Sen argues that a rational person would apply reasoned scrutiny in order to evaluate her alterative preference rankings, and in the process try to figure out her 'preference over preferences' – or a meta-ranking, as he calls it. Maximization on the basis of this meta-ranking would be much more of a rational exercise than maximization on the basis of 'given' preferences.

Finally, maximization should not be construed as an exercise that will necessarily lead to the choice of the 'best' option – defined as an option that is *at least as good as* any other option. We may instead end up choosing options that are *not worse than any other*. The relevant distinction here is between maximization and optimization. An alternative is said to be 'optimal' if it is at least as good as every other alternative. In contrast, an alternative is said to be 'maximal' if there is no better alternative. If one is looking for the 'best' option, one would

be engaging in an exercise of optimization, and this is what is normally done in economic models of rational behaviour.

But optimization is a relatively limited exercise; it is in fact, a special case of maximization. The two exercises coincide only when the preference rankings are said to be 'complete', in the sense that the elements of every pair of options can be ranked vis-à-vis each other. But there is no logical necessity why preferences must be complete in this sense, because incompleteness may be intrinsic in the nature of the problem. The real world throws up many instances where it is so hard to choose between alternative options that one must simply concede that one cannot rank them vis-à-vis each other.[8] In other words, it may not be possible to resolve incompleteness of ranking even after we have applied as much reasoned scrutiny as possible. Rationality must allow for incompleteness where it exists because of the inherent difficulty of resolving what might be unresolvable conflicts of values.

Sen, therefore, advocates the adoption of a broad framework of maximization that admits of incomplete rankings and urges us not to be beholden to the special case of optimization (based on complete rankings) as applied in standard models of rational behaviour. Even though this framework may not always be able to identify the 'optimal' ('best') option, it need not be altogether ineffective either, because it might still be able to identify the 'maximal' option i.e., an option for which there is no better alternative; and in many instances that could be a good enough basis for making rational choices.[9]

III. Behavioural Economics and Its Claims of Irrational Behaviour

Having laid out Sen's views on the notion of rationality in rational choice theory, I would now examine, in the light of Sen's analysis, the claims made by behavioural economists that the 'anomalies' they have identified challenge the foundations of the very notion of rational choice. I shall do so by first spelling out what the anomalies actually tell us about human behaviour and then scrutinise whether and to what extent they challenge the presumption of rational behaviour in economic models.

In examining the anomalies, however, I have no ambition to be comprehensive.[10] Instead, I shall pick and choose a few of the more important ones, but my sample will cover the major types of anomalies. The typology is itself a rather contentious issue, however, as there is no consensus among behavioural economists on the best way to organize the anomalies.[11] The most well-known typology – popularised by Thaler – classifies them according to three types of boundedness – namely, bounded rationality, bounded will-power and bounded self-interest. Although this classification makes for a nice symmetry, it is slightly misleading because while all or most of the anomalies are supposed to challenge the precept of 'rational choice' in one way or the other, only some of them are branded here as cases of 'bounded rationality'. Other authors have used somewhat different typologies – e.g., Rabin

---

[8] See, for example, the discussion of 'hard choices' in Levi's (1986) book of the same name.

[9] Advocacy for the need to recognize the existence of unresolvable incompleteness of rankings pervades the whole gamut of Sen's writings, spanning both economics and philosophy. The most elaborate articulation of the argument can be found in Sen (1997). See also Sen (2017).

[10] Excellent surveys of the anomalies can be found in, *inter alia*, (Rabin 1998, 2002), Mullainathan and Thaler (2001), Camerer (2006), DellaVigna (2009), Barberis (2017), and the two best-sellers written by the pioneers of behavioural economics – Kahneman (2011) and Thaler (2015).

[11]

(1998, 2003) and DellaVigna (2009) – which do not involve this oddity, but they have problems of their own. Instead of trying to adjudicate among the contending typologies, however, I have decided to choose my sample of anomalies in such a way that each major category of each typology gets represented in my discussion. My list consists of: (a) errors in judgement, (b) reference-based preferences, (c) time-inconsistent preferences, (d) the framing effect, and (e) other-regarding behaviour or social preferences.

Errors in Judgements: Biased Forecasts

One of the assumptions behind the economists' models of behaviour under uncertainty is that people make their assessment of probabilities of uncertain events according to the laws of probability – in particular, in accordance with the logic of what is known as the Bayes' theorem. In a series of psychological experiments, researchers have found, however, that people habitually deviate from the predictions based on Bayes theorem while forming their probabilistic judgements. By way of explaining these deviations, Tversky and Kahneman (1974: 1124) noted that people tend to rely on "heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations." In following these 'heuristics', or rules of thumb, however, people make systematic errors in judgements, leading to biased forecasts (biased, that is, relative to the predicted probabilities according to the Bayes theorem). This is the anomaly of 'errors in judgement' or 'biased forecasts'.

A widely used heuristic is known as the *availability heuristic*. A person is said to employ "the availability heuristic whenever he estimates frequency or probability by the ease with which instances or associations could be brought to mind." (Tversky and Kahneman (1973: 208). This process of forming probability judgements violates the precepts of Bayes' theorem which says that our assessment of probabilities should combine information on representativeness with that of base rates (i.e., the percentage of population falling into various groups). In practice, however, people form their judgement on the basis of whatever they are immediately familiar with i.e., information that is immediately available to them. Since what they are familiar with may bear little relation to the true base rate, this practice often leads to systematic biases.

One manifestation of such bias has been dubbed, somewhat ironically, as the *Law of Small Numbers* (in contrast to the much-revered *Law of Large Numbers*, which is one of the bedrocks of statistical theory). It states that people exaggerate how closely a small sample will resemble the parent population from which the sample is drawn. Yet another manifestation of bias is called the *anchoring effect*. For example, in forming numerical estimates of uncertain quantities, people attach too much weight on the initial value (the anchor), making too little adjustments that might be warranted by changed circumstances. The upshot of all this is that research in behavioural economics has made it difficult to sustain the standard assumption that probability weights are formed in accordance with Bayes theorem.

But does any of this call into question the assumption of rational behaviour? Some behavioural economists discuss the issue in a language that would suggest that it does. For example, Rabin (1998: 24) states that "researchers have documented many systematic departures from rationality in judgment under uncertainty." What seems to be implied here is

that Bayes' theorem is the gold standard of rationality when it comes to forming subjective probabilities.

However, if irrationality is to be judged in more substantive ways, rather than by the force of definition, one must ask: why does the departure from Bayes theorem occur in the first place? Why do people employ 'heuristics'? The answer, as can be seen from the quote from Tversky and Kahneman (1974) given above, is that, in view of the complexities involved in estimating probabilities, people take short cuts and follow various rules of thumb. This is clearly a case of what Simon (1955) called bounded rationality. Behavioural economists recognise it as such, but the problem arises from a common tendency to interpret bounded rationality itself as irrational. That this interpretation is highly questionable can be seen from the logic of bounded rationality, which is based on the recognition that human cognitive abilities are not infinite, our computational skills are limited and our memories are flawed. Recognising these limitations, we have to consider the trade-off between the costs of decision making and the cost of errors that would follow from imperfect decisions. Errors would obviously arise as a consequence of this trade-off, which is itself perfectly rational.

In the context of probabilistic assessment, the same arguments imply that the use of heuristics and the resulting errors and biases are an inevitable consequence of a reasonable trade-off with the costs of overcoming the cognitive and other limitations. As a result, one might even call them 'rational errors'. It would, therefore, be a mistake to characterise the use of probability heuristics as departure from rationality, when rationality is interpreted broadly in line with Sen as reasoned scrutiny of one's options.

Reference-based Preferences

There is a cluster of anomalies that have a common origin in a special feature of people's preference pattern identified by behavioural economists – namely, that preferences appear to vary parametrically depending on some reference point, which is usually the *status quo*. One consequence of such preferences is that people seem to care more about *changes* in the levels of income, wealth, consumption, etc. than about their *levels*. In standard theory, when we speak of utility function $U(x)$, the argument x is taken to be the level of some object (such as, income) over which we are supposed to have preferences. But it is now being suggested that preferences should instead be defined over changes in x from the current level, which in the present case happens to be the reference point.

A related anomaly is known as *loss aversion*. People value modest losses roughly twice as much as equal-sized gains (Tversky and Kahneman 1991). As Rabin (1997) has demonstrated, no sensible concave utility function under conventional expected utility theory can be consistent with this degree of loss aversion. There's the anomaly. On the other hand, the same degree of loss aversion can be explained much more easily by a S-shaped utility curve, with the kink at the *status quo*, which was made famous by Kahneman and Tversky (1979) in their classic exposition of prospect theory as an alternative to the expected utility theory.

A corollary of loss aversion is yet another anomaly called the *endowment effect* (Thaler 1980, Kahneman *et al.* 1990).[12] The essential idea is, as Rabin (1998: 14) put it, "once a person comes to possess a good, she immediately values it more than before she possessed it." An important implication is that the compensation a person would demand in order to part with a good one already possess is higher than what she would be willing to pay in order to acquire it. This is contrary to the prediction of standard economic theory of utility maximization.

Finally, there is the phenomenon called *diminishing sensitivity* which says that the marginal effects on utility of changes close to one's reference level are greater than for changes further away; and this applies to both upside and downside of the reference level. This is contrary to the implications of diminishing marginal utility embodied in a concave utility function, according to which marginal effects should diminish only on the upside and increase on the downside. But, once again, this is perfectly consistent with a S-shaped utility function. An interesting implication of this anomaly is that a person who is risk-averse at higher levels of income may become, paradoxically, a risk-lover at lower levels of income!

All these manifestations of reference-based preferences make a compelling case for reconsidering the usual formulation of utility maximization models.[13] Specifically, a couple of modifications are needed. First, one needs to employ a S-shaped utility function à la Kahneman-Tversky instead of an 'everywhere concave function' that is common in economic models; and secondly, the utility function needs be redefined to allow for parametric variation of preferences in accordance with the reference level.[14] But neither of these changes call into question the fundamental tenets of rational choice theory underlying these models, which claim that economic agents try to maximize their self-interest. For all the novel implications these anomalies have for a richer understanding of human behaviour, there is no challenge in all this to the foundations of rational choice theory.

There is, however, one sense in which these modifications would represent a radical departure from the traditional formulations (as distinct from foundations) of rational choice model. The relevant issue here is Sen's distinction between culmination outcomes and comprehensive outcomes discussed in the preceding section. In traditional formulations of maximizing behaviour, preferences are defined over culmination outcomes – the final objects of value. This will not do, if we are to take note of reference-based preferences. The domain of utility function must now be defined as comprehensive outcomes where the culmination outcomes are tagged with the reference level.[15]

---

[12] Thaler, however, came upon this anomaly independently; it was only later that he came to realize the connection with loss aversion and other elements of prospect theory. For a highly entertaining account of the history of endowment effect, and many other anomalies, see Thaler's most recent book with the tantalizing title of *Misbehaving* (Thaler, 2015).

[13] Most of the anomalies discussed in this cluster, with the exception of endowment effect, were originally identified in the context of choice under uncertainty, but both conceptually and empirically they apply to behaviour in general, with or without uncertainty.

[14] Formally, $U(x)$ needs to be replaced with $U(x|r)$, where x stands for the objects over which preferences are defined and r is the reference level of those objects.

[15] Sen introduced the idea of comprehensive outcomes in order to accommodate process-dependence of preferences, but the same logic applies to reference-dependence of preferences as well.

Time-inconsistent Preferences

One of the reasons why behavioral economists are so critical of rational choice theory is that they do not think ordinary humans have the kind of coherent and stable preferences the theory assumes. As a clear manifestation of incoherent preferences, they point to the existence of what is known as time-inconsistent preferences and take it as a sign of irrational behaviour.[16]

Inconsistency arises, according to behavioural economics, because the degree of impatience, as measured by the relative weight people attach to present over future consumption, appears to vary over the time horizon of planning. If, because of impatience, I value 'one dollar today' x per cent more than 'one dollar tomorrow', then consistency would seem to demand that I should also value 'one dollar 100 days from now' exactly x per cent more than 'one dollar 101 days from now'. In other words, I should be consistent over time in my degree of impatience. This is indeed what is assumed in the typical economic model of inter-temporal choice, known as the constant exponential discounting model.

But this is not how people actually behave, according to behavioural economists. They argue that, contrary to what the exponential discounting model predicts, psychological experiments show that impatience is at its peak 'now' i.e., at the time we take decision about inter-temporal choice; but we feel less impatient as we think about future choices from the vantage point of 'now'. Thus, although I am always impatient, I am more impatient today than what I think (today) I would be at a future date. This is called, quite appropriately, present-biased preference.

This is also what behavioural economists call time-inconsistent preference[17], and then use it as an argument against the validity of rational choice theory (as applied to inter-temporal choice) on the grounds of inconsistent behaviour. In order to assess that argument, it is first necessary to clarify what exactly is the sense in which present-biased preference can be viewed as inconsistent. It seems to me that the argument of inconsistency is made in two quite distinct senses, and we need to assess them separately.

The first sense of inconsistency is akin to the violation of an internal consistency condition as discussed in section II. In the present case, the consistency condition comes in the guise of constant exponential discounting that is used in the standard models of inter-temporal choice. By definition, constant exponential discounting entails that the 'discount factor' is constant for any given interval of time, at whatever point in time that interval happens to occur. Since the discount factor is a measure of impatience, the model of exponential discounting must imply that the degree of impatience remains constant over the time-horizon of planning. And this constancy of the degree of impatience can be seen as a consistency condition. If this notion of consistency is accepted, then any behaviour that fails to exhibit constant degree of impatience must, by definition, be seen as inconsistent.

---

[16] For example, Thaler and Sunstein (2003: 176) mention the finding that "Furthermore, in the context of intertemporal choice, people exhibit dynamic inconsistency, valuing present consumption much more than future consumption" as an illustration of their general claim that "Research by psychologists and economists over the past three decades has *raised questions about the rationality* of the judgments and decisions that individuals make." (*italics* added)

[17] It should be noted that this is very different from the phenomenon of time inconsistency discussed in macroeconomics.

In assessing this claim, the important question to ask is, why should a constant degree of impatience be seen as a requirement of consistency? If this condition has to have any legitimacy at all, it has to come from the legitimacy of constant exponential discounting, because that is where the condition originates from. But what is the legitimacy of exponential discounting itself? None whatsoever. As Rabin (2002: 670, fn10) rightly observes, "… those who first developed exponential discounting in our profession never claimed it was good assumption. It was proposed as an unrealistic, psychologically false, convenient simplification." It is singularly odd that a condition that is derived from such an unrealistic simplifying assumption should be treated as a standard of consistency, and then a pattern of behaviour that is found to be at variance with that standard be declared as inconsistent and irrational.

The second sense of inconsistency has a little more substance. In an influential paper, Thaler (1981) reported a couple of findings from experimental studies that are especially relevant in this context. First, people often lack self-control in many spheres of decision-making – in the sense that they do things they wished they hadn't done. Second, people who are aware that they lack self-control often respond to this awareness by tying themselves to various commitment devices which are meant to ensure that they will do what they would have liked to do. Thaler later came to call it the problem of 'bounded will-power'. Present-biased preference can be interpreted as an example of such bounded will-power, and in this sense it does imply some kind of inconsistency.

Recall that present-biased preference implies that people are more impatient about deferment of consumption 'now' (i.e., at the time of decision-making) than about deferment at a future date. So, when a decision to save is to be made, people are tempted to defer it for the future ('let's enjoy now; we'll save later'). But when the future comes, the same problem recurs; as a result, the decision to save does not materialize. There is evidently a kind of inconsistency here, because people do want to save – albeit in the future – but end up not doing so.

What is important to recognise, however, is that time-inconsistent preference in the sense just described does not necessarily amount to inconsistent or incoherent preferences of the kind that would render inter-temporal utility maximization a meaningless exercise, as it is sometimes claimed. There are different ways of approaching this issue, depending on exactly how present-biased preference is modelled. The most popular modelling approach is to redefine the inter-intertemporal utility function by replacing exponential discounting with some version of 'hyperbolic' discounting, which allows for the fact that from the vantage point of decision-making the degree of impatience declines systematically over the time-horizon of planning (e.g., Laibson 1997). Given this type of utility function, the predictions of a maximization exercise will neatly conform to the behaviour that is consistent with present-biased preference. There is no need to throw out the utility maximization model itself.

An alternative strategy of modelling was proposed by Thaler and Shefrin (1981). They used the metaphor of 'multiple selves' to model the problem of self-control in an overlapping multiple-period setting. One of the 'selves' is a 'doer', who for each period of time maximizes the utility of consumption in that period alone. Thus, for the same person, there is a series of 'doers', one for each period of time. The other 'self' is a "planner" who takes a longer-term view and cares about the welfare of all the 'doers' who will emerge over time. This leads to an

internal conflict between the 'doer' and the 'planner': the 'planner' tries to prevent each period's 'doer' from consuming as much as possible immediately so as to leave something for the future 'doers'.

A moment's reflection shows that this metaphor neatly captures in an entertaining way what Sen describes as meta-ranking i.e., preference over preferences. Sen's view of rationality allows for the possibility that people may have possible alternative preference orderings – for example, one they currently have and another they would like to have. Mere existence of such conflicting preferences is not a sign of irrationality; it's just the way preferences are. What would be irrational is to accept the existing preference blindly without any questioning. By the same token, rationality would demand that people subject the alternative preferences to reasoned scrutiny and arrive at a meta-ranking. This is precisely what, in my view, the doer-planner model tries to do. As a formal matter, there is no difficulty in describing this process as a rational exercise in utility maximization, where utility represents preference ordering in terms of a meta-ranking. It is of course possible that decisions coming out of this process may not be always be complete (in the sense that the mental tussle over 'to save or not to save' may not always be fully resolved) because the meta-ranking may itself be incomplete. Such incompleteness, Sen argues, is quite common when people face 'hard choices', but this does not invalidate the maximization exercise so long as it is interpreted, properly, as the search for 'maximality' rather than 'optimality'. Once again, there is no need to throw away the utility-maximization framework itself, but of course the framework will need to be reformulated in appropriate ways.

Framing, Context-effects and Preference Reversal

The issues to be discussed here do not relate to any specific anomaly; rather they relate to the behavioural implications of some of the other anomalies discussed above. In short, the argument is that people's choices can change, and may even be reversed, depending on the on the manner in which the choices are presented or 'framed'. The resulting change, and possible reversal, of choices is interpreted by most behavioural economists as change and reversal of preferences, which in turn is taken as evidence of unstable preferences. From there, it is only a small step to argue that with unstable preferences the very foundation of the model of utility maximization falls apart.[18]

I shall first deal with the special case of framing, known as the context effect, and then move on to the general case of framing, as the special case is much more straightforward to handle and the lessons from this case can be usefully transferred to the general case. Consider the following example, which is a slight variation of an example given by Simonson and Tversky (1992). A consumer is evaluating two personal computers; one has 60 GB memory and costs $1200 ($x$) and the other has 50 GB memory and costs $1000 ($y$). The choice between x and y would then depend on whether the consumer is willing to pay $200 more for an

---

[18] For example, Sunstein and Thaler (2003: 2-3): "Our emphasis is on the fact that in many domains, people lack clear, stable, or well-ordered preferences. What they choose is a product of framing effects, starting points, and default rules, leaving the very meaning of the term "preferences" unclear." Also, Rabin (2002: 662) "…there may not be stable, well-defined, time-invariant, and "hedonically correct" preferences U(x|s) such that behavior is best described by assuming that people maximize $\Sigma_{s \in S}\, p(s)U(x|s)$."

additional 10 GB of memory, which comes to $20 per GB. The consumer is not sure whether the cost of the extra memory in *x* is worth it's while, and is, therefore, thinking of deferring the choice (*z,* which stands for the option of deferment). Now assume an alternative scenario in which in addition to the pair of computers she was trying to choose from, the shop offers the choice of another pair of computers, also with unequal memories, for which the cost of additional memory is more than $20 per GB. The consumer now promptly decides to buy *x*. In the first scenario where the choice set was limited to (*x, y, z*), the consumer had chosen *z*. In the second scenario, when an extra pair of computers has been added to the choice set, the choice becomes *x*. Thus, the difference in context (the expansion of the choice set) has resulted in a difference in choice, which is being interpreted as unstable preference.

In the light of our discussion presented in the preceding section, one can easily see that this example demonstrates what Sen calls menu-dependence of choice. As we have also seen, menu-dependence does not imply unstable preferences in the sense that people jump around between different preference orderings in a random or unpredictable manner. It implies instead that preference orderings vary parametrically with the nature of menus on offer and these parametric variations are rational in the sense that they occur upon reasoned scrutiny of one's options; i.e., there are good reasons why they occur.[19] In this particular case, the reason is what Sen termed as epistemic value of menus, i.e., useful information that new menus might bring to the attention of the decision-maker which may have a bearing on her choice. The new information that made the difference in the present case was that when the new pair of computers was added to the menu the consumer realized that the cost of extra memory she was earlier fretting about was not too high after all.

As we noted in section II, menu-dependence of this kind violates internal consistency conditions such as the weak axiom of revealed preference. The authors recognise this, but the problem is that they also claim it to violate the model of utility maximization in general because they consider the consistency conditions to be necessary for utility maximization. Thus, referring to utility maximization as value maximization (VM), they write: "A major implication of VM is that the preference between alternatives is independent of the context, as defined by the set of alternatives under consideration" (Simonson and Tversky 1992: 281). But, this argument is flawed because there is no necessary connection between utility maximization and internal consistency conditions unless one falls into the definitional straightjacket of revealed preference theory. In the more general case, as Sen showed, we can abandon these consistency conditions and still retain the utility maximization framework. A change in formulation is needed, however, insofar as preferences now need to be defined over comprehensive outcomes instead of culmination outcomes, where culmination outcomes refer to the final objects of value and comprehensive outcomes refer (in this case) to final objects of value as tagged with alternative menus.

We now turn to the more general case of 'framing', which says that choice behaviour is found to depend not just on the menus on offer but also the manner in which a particular

---

[19] While explaining 'context-dependent' choice, Shafir *et al*. (1993) not only recognise the role of reason, they also emphasize it, as can be seen from the title of their paper "Reason-based Choice".

menu is presented.[20] Behavioural economics offers a plethora of examples from both laboratory experiments and real life where different ways of framing the choice of options actually make a difference to which option gets chosen. For illustrative purposes, we shall analyse a particular example taken from Sunstein and Thaler (2003).

The focus here is on the issue of present-biased preference, which we have already discussed. One of the major claims of the framing literature is that people can be made to save more or less depending on how the options of saving are presented to them. In the context of employer-offered savings plans, for example, Sunstein and Thaler distinguish between two different modes of presentation embodied in two alternative plans. One plan requires employees to take an affirmative action electing to save a specified fraction of current income, if they wish to save, but not otherwise; and the other one is an automatic enrolment programme in which a certain fraction of employees' income will be transferred to a savings fund by default, unless the employees explicitly choose to opt out of the scheme. Experience from real life shows that the latter plan – in which savings occur by default – invariably elicits higher levels of savings. A difference in the 'frame' of saving has made a difference in the choice to save.

Sunstein and Thaler (2003: 3) go on to ask, rhetorically: "Can workers then be said to have well-defined preferences about how much to save?", implying clearly that they don't. The spectre of incoherent and unstable preferences haunts again. Actually, this is nothing other than the spectre of revealed preference theory, which equates choice with preference by definition and thus allows choice reversals to be interpreted as preference reversals. Once we come out of the revealed preference trap, however, by giving the notion of preference an independent content, it becomes clear that the savings plan example is no more a reflection of unstable preferences than the case of context-dependent (i.e., menu-dependent) preference just discussed. Following Sen, this can be seen as a special instance of process-dependent preference, where preference orderings may be influenced by the process of choice.

There may be good reasons why the nature of the act of choice might influence the preference ordering. For the savings plan example, Sunstein and Thaler (2003: 21-22) themselves offer a powerful explanation: "In many settings, any starting point will carry some *informational content* and hence will affect choice. Hence if a default rule affects behavior, it might well be because it is taken to *carry information about* how sensible people usually organize their affairs… With respect to savings, the designated default plan apparently carries a certain legitimacy for many employees, perhaps because it seems to have resulted from some conscious thought about what makes most sense for most people" (*italics* added). What is being claimed here is that the default plan favours the choice to save because it carries useful information about how 'sensible' people behave and about how experts may have applied conscious thought in order to devise a plan that 'makes most sense'. This is precisely Sen's 'epistemic reason' why the process of choice may influence the preference ordering.

Once it is recognised that process-dependence of preferences can arise out of reasoned scrutiny (making good use of implicit information, in this case), there is no reason to suspect

---

[20] Evidently, the term context-dependent choice would have applied nicely to the general phenomenon of framing too, but somehow in the literature it has come to be associated specifically with menu-dependence. More recently, Thaler (2015) has been using the more elegant term 'choice architecture' in place of framing.

that preferences are unstable. Nor is there a reason to suspect that the utility maximization model does not work here. As in the case of context-dependent (i.e., menu-dependent) preferences discussed earlier, all that needs to be done is to recognise that the maximization exercise must now be carried out over the domain of comprehensive outcomes (taking the process of choice into account in addition to objects of choice) instead of culmination outcomes (which only includes the objects of choice).

Other-regarding Behaviour

Behavioural economics has come down heavily on the fundamental premise of standard economic models that self-interest is what drives all economic activities. A growing body of experimental research makes it abundantly clear that people's motivations, even in purely economic matters, often depart systematically and significantly from pure self-interest.[21] "Subjects contribute to public goods more than can be explained by pure self-interest; they often share money when they could readily grab it for themselves: and they often sacrifice money to retaliate against unfair treatment." (Rabin 1998: 17). Evidence from market outcomes also points in the same direction; in particular, experimental studies of labour markets demonstrate that behavior seldom converges to the self-interested outcome (e.g., Fehr *et al.* 1998). In the behavioural economics literature, this 'anomaly' goes under various appellations such as 'bounded self-interest', 'social preferences', and 'other-regarding behaviour'.

As noted in section II, economic models have already incorporated other-regarding behaviour to some extent by allowing altruistic behaviour, thanks largely to the works of Gary Becker. Behavioural economists insist, however, that the range of other-regarding behaviour they observe in their studies go well beyond altruism. For example, Rabin (1998: 17) states: "Simple altruism may parsimoniously capture important phenomena in many contexts. But there is a mass of experimental evidence that indicates it is often a very wrong description of social preferences."

Altruism, as modelled by economists, simply means that in forming judgements about one's own welfare people put a positive weight on other people's welfare as well, i.e., they incorporate other people's welfare, to some extent, into their own welfare. In contrast, the concept the behavioural economists emphasize the most is that of fairness. As Jolls *et al.* (1988: 1479) explain: "In many market and bargaining settings (as opposed to nonmarket settings such as bequest decisions), people care about being treated fairly and want to treat others fairly if those others are themselves behaving fairly. As a result of these concerns, the agents in a behavioral economic model are both nicer and (when they are not treated fairly) more spiteful than the agents postulated by neoclassical theory."

Two distinct features of the concept of fairness emphasized by behavioural economists are worth noting. First, as is evident from the quotation above, fairness is not simply a matter being nice or noble; how people want to treat others depends very much on how others are treating them. Thus, a critical feature of fairness is its reciprocal nature. Rabin (1998: 21) calls

---

[21] One of the earliest and most well-known experiments of this type is known as the 'ultimatum game'. For a fascinating account of the history and significance of this game, see chapter 15 of Thaler (2015). See also Thaler (1988).

it 'reciprocal altruism' and explains it thus: "The same people who are altruistic toward deserving people are often indifferent to the plight of undeserving people, and motivated to hurt those whom they believe to have misbehaved." He in fact goes one step further and interprets the existing evidence to demonstrate that the nature of reciprocity depends not just on the actions of others but also on the intentions and motivations underlying those actions. "When motivated by reciprocal altruism, for instance, people differentiate between those who take a generous action by choice and those who are forced to do so." (p.22)

Secondly, behavioural economists argue that whatever the moral or metaphysical connotation of the concept of fairness, when economic actors think of fairness they employ a distinctly empirical yardstick. Most people are supposed to have some notion of 'reference transaction' in mind – a transaction that defines the benchmark for the parties' interactions; and any deviation from the benchmark is perceived to be unfair. The point is illustrated clearly by Jolls *et al.* (1998: 1496) as follows: "When the interactions are between bargainers dividing a sum of money to which neither is more entitled than the other (and this is common knowledge), the "reference transaction" is something like an equal split; substantial departures are viewed as unfair...If parties are bargaining over the division of money and both have reason to view one side as more entitled than the other, then the "reference transaction" is a split that favors the more-entitled party. And if the parties are a consumer and a firm in the market, the "reference transaction" is a transaction on the usual terms for the item in question."

It is quite clear that simple altruism cannot capture the nuanced concept of fairness that behavioural economists have observed. The question I would like to explore, however, is: how much of a challenge does this pose to the kind of rational choice models employed in economics? One way of answering this question is to recall Sen's three-fold distinction of the possible roles of self-interest in decision-making – namely, self-centred welfare, self-welfare goal and self-goal choice. The narrowest model of rational choice theory combines all three roles of self-interest; Becker's treatment of altruism abandons the first role, by allowing that one's welfare may be influenced by other's welfare, but retains the other two. How does the behavioural economists' notion of fairness deal with the self-interest?

The answer would depend on exactly how the concern for fairness is perceived to influence a person's decision-making process. Consider two possibilities. First, the concern for fairness could be imagined to influence a person's welfare, in the same way that altruism is supposed to do, by giving satisfaction to a person for behaving fairly. In that case, we would be moving away from self-centred welfare, but if we insist that people behave according to the dictates of maximisation of own welfare, then we shall be keeping faith in both self-welfare goal and self-goal choice. The second possibility is that, regardless of whether or not we allow fairness-driven behaviour to influence our welfare, we might want to attach independent weight to the concern for fairness as a constraint on our behaviour. In this case, we shall be abandoning self-goal choice.[22]

How much of a radical challenge the emphasis on fairness poses to rational choice theory depends on which of the two scenarios better describes human behaviour. The issue of central importance here is the distinction Sen makes between sympathy and commitment. If

---

[22]What happens to the other two roles of self-interest would then depend on what assumption we made about the effect of fairness-driven behaviour on our welfare.

fairness concern is incorporated into own welfare, it's a case of sympathy[23]; and if it is perceived as a constraint on self-goal pursuit, it's a case of commitment. And as we saw in section II, only commitment can cause a radical departure from standard models of rational choice behaviour. So, the question arises: from the perspective of behavioural economics, what does the concern for fairness mean for people's behaviour – sympathy or commitment?

As it happens, it's not easy to answer this question because, while behavioural economists rightly differentiate between altruism and fairness, they seldom spell out which of the two ways of treating the concern for fairness they believe to be the right way of understanding human behaviour. Perhaps, because of this ambiguity, Sen (2002: 30) was led to the view, while commenting on Jolls *et al* (1998), that behavioural economics packs "everything into a broader view of self-interest itself. People are willing to "treat others fairly" to the extent that they enjoy that, or are otherwise made better off as a result of it, not because they see this as a commitment to which they should adhere whether or not they enjoy it – or otherwise benefit from it."

But the same ambiguity also means that one can also take the other view – that behavioural economists are really thinking of commitment when they see the influence of fairness. At least, that seems to be the view taken by Rabin (1998: 17), who, in the context of fair allocation of a given object of value, proposes to model fairness into choice theory in the following way.[24]

> "Very roughly, imagine that person 1's utility function takes the form
> $$U_1 \equiv (1 - r)\Pi_1 + r.W_1(\Pi_1, \Pi_2)$$
> where W_1 is person 1's view of the proper allocation, and Π_1 is … her self-interested payoff, and $0 < r < 1$ measures the weight this person puts on self-interest versus proper allocation. To understand the implications of fairness and justice for behavior, we need to know both the nature of the W_1 function, and the level of r."

In this formulation, the decision-maker's view of what counts as a fair allocation rule (W_1) acts as a constraint on the pursuit of own welfare goal (Π_1), which is consistent with the commitment interpretation rather than with the sympathy interpretation.[25]

But modelling is one thing, and how people actually take fairness into account is quite another. On the latter, as far as I am aware, there is no clear-cut evidence. The experimental studies in behavioural economics clearly demonstrate that people's concern for fairness cannot be interpreted as simple altruism, but whether it is to be interpreted as sympathy or commitment is not altogether clear, although I think the weight of evidence points more in the direction of commitment. Perhaps, more discriminating experiments are needed to answer this question satisfactorily. Meanwhile, it can be accepted that, of all the anomalies discussed by behavioural

---

[23] It may be noted that since sympathy also covers antipathy (seen as sympathy with a negative sign, so to speak), there is no problem in accommodating under sympathy the issue of reciprocal fairness, which allows that people may want to punish others for perceived unfair behaviour.

[24] For a fuller account of his views about how fairness is to be incorporated into economic models, see Rabin (1993).

[25] In fact, this interpretation becomes even more apparent when Rabin's equation is interpreted as a Lagrangian, with the primitive utility function being Π_1, which is maximised subject to the constraint of the fair allocation rule W_1, and $r/(1-r)$ being the Lagrange multiplier.

economists, the one relating to fairness-driven behaviour has the greatest potential to call for the most radical departure yet from the standard formulations of rational choice theory.

It needs to be emphasized, however, that even if commitment turns out to be the right interpretation, so that models of human behaviour will finally be required to abandon the precept self-goal choice, it would not follow that the precept of utility maximization will also have to be abandoned. Formally, one can still say that people will maximize the utility function such as the one proposed by Rabin, where, in the terminology of Sen, people's preferences are defined over comprehensive outcomes rather than culmination outcomes. The comprehensive outcomes in this case would tag the culmination outcome (Π_1) with considerations of fairness of processes as parametrised by r and W_1.

IV. Concluding Remarks

Behavioural economics has made outstanding contributions over the last few decades towards enriching our understanding of human behaviour. Studies carried out by behavioural economists, using mainly laboratory-based psychological research but increasingly also empirical data from the real world, clearly demonstrate that human behaviour is much more complicated and much more nuanced than what is typically assumed in standard models of economic behaviour. So as to emphasize the starkness of the contrast between models and reality, Richard Thaler and Cass Sunstein have constructed the interesting duality between the 'Econ' and the 'Human', where the former stands for how people are supposed to behave according to economic models and the latter for how people actually behave (Thaler and Sunstein 2008). While the contrast between the two is real enough, there are reasons to question the way many behavioural economists interpret the contrast. At a very general level, they tend to project the Econ as rational and the Human, by implication, as irrational (not always, but in many important ways).[26] My objective in this paper has been to examine the validity of this particular interpretation of the contrast between the Econ and the Human.

Since, however, rationality is a rather slippery concept, amenable to many different interpretations, my first task was to sort out in a systematic way what exactly the demand for rationality entails. In undertaking this task, I have followed the pathbreaking works of Amartya Sen, who has done more than any other scholar to clarify the notion of rationality, especially as it is applied in the rational choice models used in economics and related disciplines.

When the claims of behavioural economics are seen through the lens of Sen's analysis, a number of conclusions emerge. First, it turns out that many of the contrasts (or anomalies, as they are called) that behavioural economists interpret as departures from rational behaviour are not really so. Those departures arise from some auxiliary assumptions (for example, concavity of the utility function and probability judgements based on Bayes theorem) that can be dispensed with while keeping intact even the narrowest versions of the rational choice model. Second, there are some types of 'Human' behaviour that seem to be more substantively at variance with the precepts of rational choice theory as captured by the image of the 'Econ', but upon closer inspection some of those behaviours too turn out to be rational in a broad sense.

---

[26] Thus, referring to the growing body of research in behavioural economics, Thaler and Sunstein (2008: 8) remark: "That research has raised serious questions about the rationality of many judgements and decisions that people make."

However, recognition of those 'anomalies' (captured, for example, though 'framing' effects and 'menu-dependence' of choice) does call for significant reformulations of the rational choice models, and Sen's work suggests innovative ways in which such reformulations can be accomplished (for example, by distinguishing between culmination outcomes and comprehensive outcomes). Third, there are some aspects of 'Human' behaviour (for example, a concern for fairness) that call for really radical departures from the rational choice models; and here again Sen's analysis has a lot to offer by way of suggesting how to proceed (for example, by distinguishing between sympathy and commitment as motivation for other-regarding behaviour). In short, behavioural economics does have great potential to bring about significant changes in the models of rational choice used in economics and related disciplines, but in order to do so effectively it needs to be more discerning in the use of the concept of rationality. Amartya Sen's analysis of rational behaviour can be of enormous help in regard.

***

References

Akerlof, G. (1982). "Labor Contracts as Partial Gift Exchange", *Quarterly Journal of Economics*, 97(4): 543-69.

Barberis, N. (2018). "Richard Thaler and the Rise of Behavioral Economics." (Mimeo.) Yale University, New Haven.

Becker, G. (1976). *The Economic Approach to Human Behaviour*. Chicago: Chicago University Press.

Becker, G. (1996). *Accounting for Tastes*. Cambridge, Mass.: Harvard University Press.

Camerer, C. (2006). "Behavioral Economics." In R. Blundell, W. K. Newey and T. Persson (eds) *Advances in Economics and Econometrics: Theory and Applications*; Volume 2. Cambridge and New York: Cambridge University Press.

DellaVigna, S. (2009). "Psychology and Economics: Evidence from the Field", *Journal of Economic Literature*, 47(2): 315–372

Fehr, E., Kirchler, E., Weichbold, A. and Gachter, S. (1998). "When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets", *Journal of Labour Markets*, 16(2): 324-51.

Harsanyi, J. (1955). "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy* 63: 315.

Jolls, C., Sunstein, C., and Thaler, R. (1998). "A Behavioral Approach to Law and Economics". *Stanford Law Review*, 50: 1471-1550.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.

Kahneman, D., and Tversky, A. (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–91.

Kahneman, D., Knetsch, J. L. and Thaler, R. H. (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy*, 98(6): 1325–48.

Laibson, D. (1997). "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443–77.

Levi, I. (1986): *Hard Choices: Decision Making under Unresolved Conflict*. Cambridge: Cambridge University Press.

Mullainathan, S. and Thaler, R. H. (2001). "Behavioral Economics." In N. J. Smelser and Paul B. Baltes (eds.) *International Encyclopedia of the Social and Behavioral Sciences*, Volume 20. Oxford and New York: Oxford University Press.

Rabin, M. (1993) "Incorporating Fairness in Game Theory and Economics", *American Economic Review*, (83:5): 1281-1301.

Rabin, M. (1998). "Psychology and Economics." *Journal of Economic Literature*, 36(1): 11–46.

Rabin, M. (2000). "Risk Aversion and Expected-Utility Theory: A Calibration Theorem", *Econometrica*, 68(5): 1281-92.

Rabin, M. (2002). "A Perspective on Psychology and Economics", *European Economic Review*, 46(4–5): 657–85.

Sen, A. (1973): "Behaviour and the Concept of Preference," *Economica*, 40, 241-2.

Sen, A. (1974). "Choice, Orderings and Morality," in S. Körner, (ed.), *Practical Reason*. New Haven: Yale University Press.

Sen, A. (1977). "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory," *Philosophy and Public Affairs*, 6, 317-44.

Sen, A. (1982) "Introduction", in *Choice, Welfare and Measurement*. Cambridge, Mass: Harvard University Press.

Sen, A. (1985a). "Rationality and Uncertainty," *Theory and Decision*, 18, 109-127.

Sen, A. (1985b). "Goals, Commitment and Identity," *Journal of Law, Economics and Organization,* 1: 341–55

Sen, A. (1993): "Internal Consistency of Choice," *Econometrica*, 61, 495-521

Sen, A. (1997): "Maximization and the Act of Choice", *Econometrica*, 65(4): 745-79.

Sen, A. (2002). "Introduction", in *Rationality and Freedom*. Harvard University Press: Cambridge, Mass.

Sen, A. (2005). "Why Exactly is Commitment Important for Rationality", *Economics and Philosophy, 21: 5-13.*

Sen, A. (2017). "Reason and Justice: The Optimal and the Maximal", *Philosophy*, 92: 5-19.

Shafir, E., Simonson, I. and Tversky, A. (1993). "Reason-Based Choice," *Cognition*, (Special Issue: Reasoning and Decision Making), 49(l-2): 11-36.

Simon, H. A. (1955). "A Behavioral Model of Rational Choice", *Quarterly Journal of Economics*, 69(1): 99–118

Simonson, I. and Tversky, A. (1992). "Choice in Context: Tradeoff Contrast and Extremeness Aversion." *Marketing Research*, 29(3): 281-95.

Strotz, R. H. (1955). "Myopia and Inconsistency in Dynamic Utility Maximization", *Review of Economic Studies*, 33(3): 165-80

Sunstein C. and Thaler, R. (2003). "Libertarian paternalism is not an Oxymoron", *University of Chicago Law Review,* 70:1159–1202.

Thaler, R. (1980). "Toward a Positive Theory of Consumer Choice", *Journal of Economic Behavior and Organization,* 1: 39-60.

Thaler, R. (1988). "Anomalies: The Ultimatum Game." *Journal of Economic Perspectives*, 2(4): 195-206.

Thaler, R. (2015). *Misbehaving*. London: Allen Lane.

Thaler, R. and Shefrin, H. (1981). "An Economic Theory of Self-Control", *Journal of Political Economy,* 89: 392-406

Thaler, R. and Sunstein, C. (2003). "Libertarian Paternalism", *American Economic Review*, 93(2): 175-79.

Thaler, R. and Sunstein, C. (2008). *Nudge*. Yale University Press: New Haven

Tversky, A. and Kahneman, D. (1973). "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology*, 5(2): 207-32.

Tversky, A. and Kahneman, D. (1974). "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185(4157): 1124–31.

Tversky, A. and Kahneman, D. (1991). "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 106(4): 1039-61.

*****