Ivan Moscati

Bocconi University
Department of Economics

Paola Tubaro

University of Greenwich, Department of
International Business and Economics

RANDOM BEHAVIOR AND THE AS-IF DEFENSE OF RATIONAL CHOICE THEORY IN DEMAND ANALYSIS

This draft: May 2010

Abstract

Rational choice theory (RCT) models decision makers as utility maximizers, and is often defended via the as-if argument popularized in economics by Milton Friedman (1953). An alternative model is random-choice, which assumes that decision makers pick up an element from a given set according to a uniform distribution on the set. In a classic contribution by Gary Becker (1962) as well as in a series of recent experiments on consumer demand, random choice has played a significant role in discussions about the validity of RCT. In this paper we investigate how and why the random-choice model has been used in theoretical and experimental analysis of individual demand based on RCT. Moreover, we examine the standard as-if defense of RCT from the viewpoint of the random-choice model and argue that this defense is a weak one in a twofold sense.

Keywords: rational choice theory; random-choice theory; as-if defense of scientific theories; experimental economics; demand analysis.

JEL codes: B410, C140, D110, D120.

Correspondence may be addressed to Ivan Moscati at ivan.moscati@unibocconi.it or Paola Tubaro at p.tubaro@greenwich.ac.uk. We are grateful to Richard Bradley, Marco Dardi, John Davis, Franz Dietrich, Francesco Guala, Jean-Sébastien Lenfant, Aldo Montesano, Jack Vromen and Thomas Wells for helpful discussions and suggestions on previous drafts of the paper. We also thank the participants in seminars at the London School of Economics, Université Paris Ouest Nanterre La Défense, Università degli Studi di Milano, Erasmus University Rotterdam, as well as in sessions at the 2009 ASSA meeting, the 2009 STOREP conference and the 2009 HES meeting for their useful comments. Moscati also thanks the Centre for Philosophy of Natural and Social Science at LSE for its hospitality during part of the work on the paper, and is glad to acknowledge the financial support of MIUR PRIN 2007. Any errors are ours.

1. Introduction

Rational Choice Theory (RCT) designates the theory according to which individuals choose among available options to maximize their utility. More extensively, RCT assumes a decision-maker to be rational if her preferences are complete and transitive, and if she chooses her most preferred option. Preferences are not restricted with respect to their orientation, and can be selfish, altruistic, or determined by social norms. If continuous, the preferences of a rational decision-maker can be represented by a utility function, and her choices can be viewed as the result of utility maximization. This is the basic version of RCT that applies to individual demand analysis. In the paper we refer mainly to this basic version of RCT, but our arguments concern also more sophisticated versions of the theory.

The random choice model is one that is apparently very different from RCT. In it, a decision-maker chooses at random when he picks up an alternative in his choice set according to some probabilistic distribution on the set, typically a uniform distribution where each alternative in the choice set has an equal probability of being selected. In particular, the literature we discuss in the paper considers random consumers who pick up a bundle from among those that exhaust their whole income, i.e., a bundle on their budget hyperplane, so that each bundle on the hyperplane has an equal probability of being chosen.

Random choice was introduced into economics long ago, playing a significant role in discussions about the validity of RCT. For instance, Alchian (1950) depicted the economic system as a selection mechanism in which the survival of a firm could be the outcome not only of rational choices motivated by profit maximization but also of random behavior. Becker (1962) pointed out that some implications of RCT for demand behavior are, on average, still valid even if consumers choose in a random way. More recently, Gode and Sunder (1993) showed that the prices and quantities obtained in double auction markets populated by artificial random traders converge towards the equilibrium price and quantity predicted by the standard supply-and-demand model; this result suggests that market efficiency may be largely independent of the rationality and motivation of the agents. Finally, laboratory experiments such as those performed by Cox (1997), Sippel (1997), Mattei (2000), Harbaugh, Krause and Berry (2001), Andreoni and Miller (2002), and Février and Visser (2004), used the random-choice model as an alternative hypothesis in testing the extent to which individual demand behavior satisfies RCT.

The first contribution of the present paper is to study the use of random choice in theoretical and experimental research on individual demand from a methodological perspective.[1] Specifically, we analyze the use of random choice in recent laboratory experiments on individual demand and Becker's seminal 1962 article, which remained the basic reference for subsequent research. We spell out similarities and differences between the definitions, usages

and implications of random behavior in the approaches of Becker and later experimentalists. We also address the question of why demand analysis has focused on the extreme and quite implausible random-choice model rather than on some more realistic model of bounded rationality. Finally, we examine the various possible specifications of random behavior proposed in the literature.

As mentioned earlier, random choice was introduced into economics in discussions about the validity of RCT. Becker interpreted his results about the fact that RCT and the random-choice model have similar implications on demand behavior as a strengthening of RCT, in the sense that those results would show that RCT provides a compelling explanation of consumer behavior even if consumers choose at random. In making this point, Becker adopted the as-if methodological view popularized in economics by his Chicago teacher Friedman in "The methodology of positive economics" (1953). Becker (1962, p. 4) wrote: "Households can be said to behave not only 'as if' they were rational but also 'as if' they were irrational: the major piece of empirical evidence justifying the first statement can equally well justify the second."

Along these lines, the second contribution of the paper is to look at RCT and the as-if defense of it from the viewpoint of the random-choice model. This viewpoint allows us to contribute to the debate on the as-if methodology and to argue that the standard as-if defense of RCT is a weak one, in a twofold sense. First, it only states that RCT offers a possible explanation of demand behavior without showing that it provides the best explanation. Second, the defense concentrates on RCT's fit with observed behavior although this fit turns out to be poor. The random-choice angle also permits us to point out a methodological shortcoming in the series of experiments on consumer behavior mentioned above, namely that they are aimed at falsifying or validating RCT but do not investigate whether RCT explains human demand behavior better than other models of choice.

Some final specifications on scope and goals of the paper are in order. To begin with, it is important to stress that we claim neither that human consumers generally choose at random, nor that the random-choice model explains human demand behavior better than RCT or other models of boundedly rational choice. We focus on random choice for methodological and historical reasons rather than substantive ones. Although random choice does not constitute a plausible alternative to RCT, it offers a fruitful methodological perspective to investigate RCT and the as-if defence of it. Moreover, despite the historical importance of the random-choice model in demand analysis, to the best of our knowledge there is no detailed discussion of the way this model has been used in this area of economics. We provide this discussion, but we do not make any case for random choice.

Additionally, although this paper focuses on basic individual-demand RCT, it also bears upon the more sophisticated versions of RCT that apply to contexts involving uncertainty, strategic

interactions or intertemporal decisions, as well as upon the extensions that attempt to capture phenomena neglected by standard RCT. In fact, sophisticated versions and extensions of RCT also construe decision-making as the result of the maximization of some kind of utility function and are interpreted as as-if explanations of human choice behavior. Therefore, our critique of the standard as-if argument is also relevant for sophisticated and extended RCT.[2]

It may be argued, however, that nobody believes in basic individual-demand RCT anymore, so that we are merely beating a dead horse. However, basic RCT is still part of almost any basic or advanced textbook in microeconomics, and thus seems far from dead. More importantly, the horse we beat is not basic RCT but the as-if defense of it. The as-if argument is alive and well in microeconomics and, as already observed, is also widespread in defenses of sophisticated and extended forms of RCT.

The remainder of the paper is organized as follows. Section 2 analyzes the use of random choice, from Becker's article to experiments on individual demand that test RCT by using the random-choice model as an alternative hypothesis. Some parts of this section may seem rather technical, but they are necessary for understanding what follows. Sections 3 and 4 complete this analysis by addressing the question of why demand theory has focused on the random-choice model, and by discussing various possible specifications of random behavior proposed in the literature. Section 5 presents the first sense in which the standard as-if defense provides only weak epistemological support for RCT. Section 6 briefly reviews the results emerging from recent experimental research on individual demand, while in Section 7 we argue that these results indicate a second weakness of the as-if defense of RCT. Section 8 summarizes and concludes.

2. Random choice from Becker 1962 to demand experiments

As mentioned in the previous section, individual demand analysis has availed itself of random choice in discussions about the validity of RCT since at least Becker's 1962 article. However, the usages and implications of random choice have changed since then, and have, in particular, taken a remarkably different role in laboratory experiments. The following sub-sections illustrate the nature of these changes.

2.1 Becker

Under the additional assumption of locally non-satiated preferences, RCT implies that a rational consumer chooses a bundle on his budget hyperplane and, among other things, that his demand displays a negative substitution effect.[3] This means that when prices change and the rational consumer is compensated so that at the new prices he can just afford the bundle he chose at the old prices, his demand for the relatively dearer commodities will decrease.[4] A

negative substitution effect is a necessary condition for seeing the consumer's choices as if generated by utility maximization. The condition is not sufficient because substitution effects involve only compensated price changes and a consumer could violate RCT when uncompensated price changes occur.

Becker (1962) imagined a random consumer who chooses a bundle on his budget hyperplane according to a uniform distribution, and compared the implications of this random decision rule with those of RCT. Becker considered the two-commodity case so that in his paper the budget hyperplane comes down to a budget line, but this restriction involves no loss of generality. As a random consumer has an equal chance of choosing any bundle on the budget line, on average he is expected to pick up the bundle laying at the midpoint of the line. When the consumer has income I and faces prices p_x and p_y, the midpoint has coordinates ($I/2p_x$, $I/2p_y$), which means that the random agent is expected to consume quantity $I/2p_x$ of commodity x and quantity $I/2p_y$ of commodity y.

In Figure 1, the initial budget line is labeled as AB and e_0, the midpoint of AB, is the bundle the random consumer is expected to pick up. If commodity x becomes relatively dearer with respect to commodity y, the compensated budget line CD passing through e_0 is steeper than AB. As a result, the midpoint e_1 of the compensated budget line is to the left of e_0, which in turn means that the expected compensated demand of x has decreased. Therefore the random-choice model also implies, at least on average, that the substitution effect is negative.

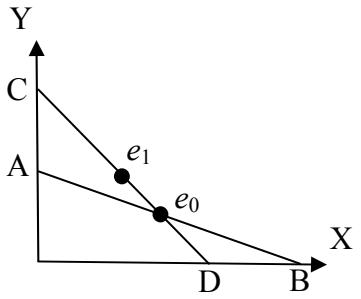

Figure 1: Random choice implies a negative expected substitution effect

Becker was chiefly interested in consumers' aggregate and non-compensated demand, i.e., in market demand, rather than in individual consumer behavior. RCT predicts that almost certainly the market demand curve is negatively sloped, i.e., that the market demand for a commodity and its price move in opposite directions.[5] Becker noted that the random-choice model also implies a negatively sloped market demand curve with almost certainty. As observed above, a random consumer is expected to consume quantity $I/2p_x$ of commodity x. Therefore, when p_x rises and income I does not change (this is an uncompensated price variation), his expected consumption of x will decrease. As the number of random consumers on the market gets higher, the average market demand gets closer to the individual expected

demand, so that it is almost certain that a market populated by a large number of random and uncorrelated consumers will display a negatively sloped demand curve.

2.2 The limitations of Becker's result

Becker's result about the convergence between rational and random consumers with respect to the sign of the substitution effect draws from his identification of *random choice* with the *expected outcome of random choice*. However the two are different: random choice means that every bundle on the budget line has an equal probability of being chosen, not that the midpoint bundle will always be chosen. In effect, if random choice is reduced to its expected outcome, a random consumer is indistinguishable from a rational consumer with a Cobb-Douglas utility function of the form $U(x, y)=x^{½}y^{½}$, as both pick up bundle $(I/2p_x, I/2p_y)$. Therefore, not only will an "expected-outcome random consumer" display a negative substitution effect, but he will also satisfy any necessary and sufficient condition characterizing RCT.

By contrast, when each and every random choice is considered, the convergence between rational and random consumers with respect to the sign of the substitution effect fades away. By looking at the compensated budget line CD in Figure 1, we notice that there is a probability α equal to the ratio between the length of segments e_0D and CD that the random consumer chooses a bundle on the right of e_0. Therefore, the random-choice model states in effect that there is a probability α that the consumer displays a positive substitution effect, i.e., that he violates RCT.

Contrary to Becker, later experimental studies take into account each and every choice made by random agents, and not only the expected or average outcome of their choices. Becker referred to the expected outcome of random choice because his main interest lay in the negative slope of aggregate demand rather than in the properties of individual demand, and the expected outcome is sufficient to obtain a negatively sloped market demand curve. The focus of experimental studies, by contrast, is primarily on individual demand and the rationality issues related to it: random behavior becomes relevant to these topics only if all choices made by random consumers are considered.

Notice, also, that Becker compared the behavioral implications of RCT with those of the random-choice model, but did not test the two theories against the demand behavior of human subjects as recorded on markets, e.g., by using panel data about household expenditure, or in laboratory experiments. In effect, when Becker published his 1962 article experimental research was still only a niche within economics.[6] More systematic efforts to test RCT in experiments on human demand behavior began in the 1970s, but most of the research was performed only after the mid 1990s.

2.3 Demand experiments

In experiments on individual demand, each human participant is typically asked to choose the preferred bundle among those affordable to him under different budget/price situations, that is, with different incomes and for different commodity prices. The experimenter records subjects' choices and usually checks whether they satisfy the Generalized Axiom of Revealed Preference (GARP).[7] GARP characterizes RCT in the sense that the choices of a subject can be seen as if generated by the maximization of a locally non-satiated utility function if and only if they satisfy GARP. Therefore, while a negative substitution effect is only a necessary condition to see choices as if generated by RCT, GARP is a necessary and sufficient condition.[8]

Figures 2a-2e below give a rough geometrical intuition about which choices satisfy GARP and which violate it in the two-commodity case.[9] In all Figures two budget/price situations are considered, the first identified by budget line AB and the second by budget line CD. It is assumed that the subject chooses bundle e_0 in situation AB while she picks up e_1 in situation CD. The choices represented in Figures 2a, 2b and 2c satisfy GARP, while those in Figure 2d violate it. Figure 2e represents the case when the two budget/price situations coincide but the subject chooses e_0 in AB and e_1 in CD. This pattern of choice can be interpreted as a manifestation of the subject's indifference between e_0 and e_1 and does not violate GARP.

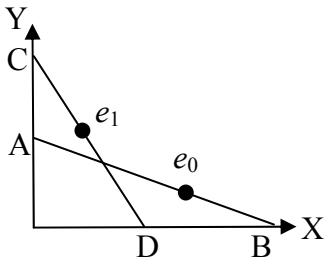

Figure 2a: GARP satisfied

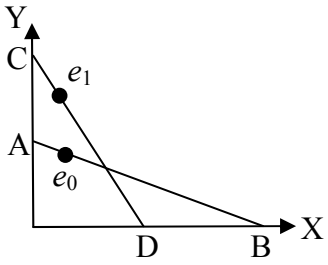

Figure 2b: GARP satisfied

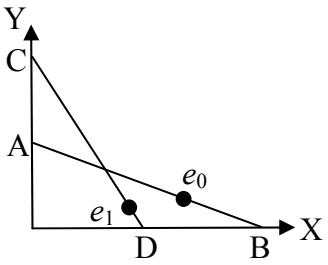

Figure 2c: GARP satisfied

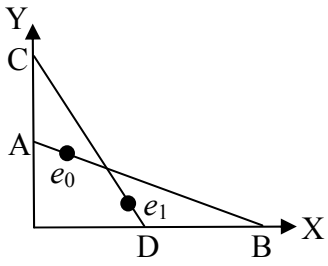

Figure 2d: GARP violated

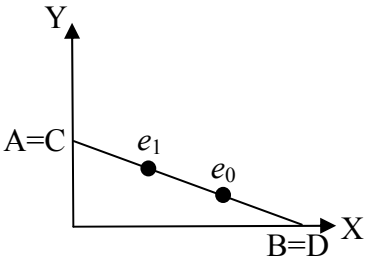

Figure 2e: GARP satisfied

There are, however, two problems with GARP as a test for RCT. First, in order to count a violation of GARP recorded in an experiment as a violation of RCT one needs to assume a form of separability between the demand choices the subjects made within and outside the

experiment. Experimental studies typically circumvent the problem by assuming this form of separability more or less implicitly.[10]

Second, even if all choices recorded in the experiment pass the GARP test, this finding may provide little support for RCT. Indeed, human subjects may satisfy GARP simply because in the budget/price situations they face it is extremely difficult or even impossible to violate it. For instance, it is easy to see that GARP violations become unlikely when the budget hyperplanes intersect near the axes, and impossible when they intersect on the axes or do not intersect at all.

In statistical terms, this is the problem of the power of a test, that is, the probability of a test to reject the so-called null hypothesis (in our case RCT) when the hypothesis is false. To assess the power of GARP as a test for RCT it is necessary to formulate an alternative hypothesis about the decision rule that could have generated human choices. The random-choice model enters the scene as the alternative hypothesis usually employed in experiments to assess the power of GARP as a test for RCT. The idea is that the lower the probability of GARP violations under random behavior, the lower the power of GARP. In this case even if RCT is false and human subjects choose at random, they would rarely violate GARP.[11]

In order to measure GARP's power a further obstacle has to be overcome, though. When consumers choose among more than two commodities and in more than two budget/price situations, calculating the a priori probability of GARP violations under random behavior turns out to be extremely difficult. To circumvent the obstacle, experimenters use computational techniques. Using Monte Carlo methods they create a large population of artificial random agents who face the same budget/price situations the human participants in the experiment were presented with. In each situation each random agent chooses a bundle on its budget hyperplane according to a uniform distribution. Its choices over the entire set of budget/price situations may or may not violate GARP. The *percentage of random agents* that violate GARP is adopted as a proxy for the a priori probability of GARP violations under random behavior, and hence as a measure of the power of GARP test. If a small proportion of random agents violate GARP, then the fact that human subjects also rarely violate it provides little support for RCT, since the rarity of human violations seems due to the objective difficulty of violating GARP under the budget/price combinations of the experiment rather than to the rational behavior of the participants.[12]

It is important to stress that in checking for GARP violations all choices made by random agents are taken into account, and not simply the expected or average outcome of their choices as in Becker's article. As observed in Section 2.2, when each and every random choice is considered, it turns out that random agents frequently violate GARP, so that Becker's suggestion of a possible convergence between random and rational individual behavior ceases

to be appropriate.

The GARP violations' relative lack of severity presents a further issue. In all experiments a number of human choices violate GARP, and this would imply that experimental data falsifies RCT. However, in many cases GARP violations are rare, in the sense that nearly all subjects satisfy GARP, and not severe, in the sense that violations would disappear by slightly relaxing the budget constraints. In these cases it may be doubted whether it is indeed appropriate to reject RCT.

Building on the work of Afriat (1967, 1972), Varian (1991) proposed a weakening of GARP called GARP(*e*), which takes into account the severity of violations through a parameter *e* called the Afriat Efficiency Index. The Afriat Efficiency Index measures the extent to which the budget constraints should be relaxed in order to let GARP violations disappear, and can be interpreted as indicating the proportion of income the subject wastes by not choosing rationally. The Index has a maximum value of one (in which case the subject satisfies GARP and no relaxation of the budget constraints is needed), decreases as violations become more severe, and has a minimum value of zero. GARP(*e*) with an Afriat Efficiency Index of 0.95, that is, GARP(0.95), can be associated to an acceptable 5% waste of income and is the one usually considered in the literature. So, for instance, if 99% of subjects satisfy GARP(0.95), they may be seen as "almost rational", in the sense that a small relaxation in the budget constraints would render the near totality of them consistent with RCT.

Like GARP, also the GARP(*e*) test presents a power problem. If human subjects satisfy GARP(*e*), this may not draw from their almost rational behavior, but from the fact that in the budget/price situations they face violations are extremely difficult. As in the case of GARP, random behavior is hence used to evaluate the power of GARP(*e*) as test for "almost-RCT".

On the whole, the use of the random-choice model as an alternative hypothesis to assess the power of GARP as a test for RCT, the reference to the percentage of random violators to measure this power, and the count of each and every choice made by random agents are all major differences between Becker's and later experimentalists' random choice.

3. Why random choice?

The above discussion raises the question of why, among the infinitely many possible alternatives to RCT, consumer demand analysis has focused on random choice rather than some model of bounded rationality.[13] It could be argued that random choice is as unrealistic as RCT, since assuming that each alternative has an equal probability of being selected also calls for unlikely cognitive skills. Furthermore, random behavior can be regarded as a special case of rational behavior when the decision-maker's preferences are represented by a flat utility function. Since in this case all alternatives maximize his utility, the agent is rational in picking

one at random; but if the random-choice model is just a special case of RCT, comparing the two would illuminate very little. These criticisms notwithstanding, random choice has attractive features that make it a natural focal point for research on RCT.

First, if one excludes the trivial case of a flat-utility decision-maker, random choice appears to contrast rational choice in a way that boundedly rational choice does not. In spite of all their differences, RCT and the various theories of bounded rationality share an intuitive idea of rationality according to which decision makers attempt to use their scarce resources as best as possible to attain their ends. While utility maximizers manage to employ their resources in an optimal way, boundedly rational agents adopt a heuristic that generally results in a sub-optimal use of resources. In random choice, on the contrary, the intentional and instrumental aspects of rationality are missing, for random decision makers do not seem to have intentions and ends, and in any case they do not use their resources as instruments to attain their ends. As far as rational choice entails maximal efficiency in the use of resources – while efficiency is not even pursued in random choice, and can easily be at a minimum – random choice appears as the opposite of rational choice. Hence if an economic phenomenon can be explained as the result of either a random or a rational decision, it seems to be due to factors that are unrelated to rationality. This makes the random-choice model an important reference point for research about RCT.

Second, even if RCT and the random-choice hypothesis can be viewed as opposite in the sense illustrated above, in effect both models frame choice as picking up an alternative from an exogenously given set. This differs markedly from the traditional practice of bounded-rationality theorists, who often modeled the choice set as endogenously determined by the choice process itself, in the belief that the discovery of available alternatives is an essential part of the process. The fact that RCT and the random-choice model frame choice in a similar way makes the comparison between them easier.

Third, since no one of the disparate theories of bounded rationality seems to have a definite pre-eminence over the others, it is not obvious which one should be compared with RCT. What random choice provides instead is a clear-cut benchmark to be contrasted with RCT. As we illustrate in the next Section, there are in fact some degrees of freedom even in defining random choice. However, for reasons that will be detailed below, Becker's original model of randomness has dominated the literature, rendering immaterial the concerns of arbitrariness related to the definition of random behavior.

Fourth, and most importantly, too often the behavioral implications of bounded-rationality models remain unspecified, and this makes the comparison between them and RCT difficult. In fact, most models of bounded rationality are introduced on the basis of plausible psychological insights, but they lack a choice-theoretic characterization analogous to the

GARP characterization of RCT. As a consequence, it is difficult to decide whether observed human behavior validates or falsifies these models, or to investigate whether they fit observed choices better than RCT.[14]

The lack of a choice-theoretic characterization of most models of bounded-rationality and the attractive features of the random-choice model discussed above partly account for the historical fact that consumer demand analysis has focused on random choice.

Things are changing, however. Masatlioglu and Ok (2005), Manzini and Mariotti (2007, 2009, 2010a) and others have recently put forward a new generation of bounded-rationality models that are easily comparable with RCT. Like RCT, these models frame choice as selecting an alternative from an exogenously given set. Moreover, in terms of behavioral implications they are characterized by axioms weaker than but analogous to GARP. Manzini and Mariotti (2010b) also performed an experiment on intertemporal decision that compares RCT with one of the new models of bounded-rationality. They argued that this model, called "Categorize Then Choose", fits their experimental data better than RCT.[15] By providing models that are directly comparable with RCT and are also more plausible than the random-choice model, this new strand of research on bounded rationality might make the attraction of consumer demand analysis for random choice decline or even disappear. We shall return to this point in the Conclusions section.

4. Types of random choice

As mentioned above, the definition of random choice allows for some degree of freedom. Generally speaking, random-choice models may vary in the identification of the set from which random agents pick an alternative and the specification of the probability distribution with which they do so.[16]

The choice set is the set of bundles on the budget hyperplane in Becker's model of random behavior. Chant (1963) put forward a different version of random behavior in which the decision-maker chooses among goods rather than bundles, in such a way that any unit of the available goods has an equal chance of being selected, independently of its price. Besides Becker's version of random behavior, Chant's has been sometimes used in experiments on individual demand to assess the power of GARP, namely by Février and Visser (2004). This article also illustrates the main problem with Chant's random agents, namely that they violate GARP more frequently than Becker's random agents, so that bringing them into play tends to overrate the power of GARP as a test for RCT.

Even if one is confined to random choice among bundles, the random consumer may be thought of as picking a bundle from among those he can afford, that is, not only from among the bundles *on* his budget hyperplane, but also from among those *below* it. In effect, restricting

the choice set to the budget hyperplane is not a trivial move, as it imposes a certain amount of instrumental rationality on random agents: if they select a bundle on the hyperplane they are at least avoiding any waste in income. Although this restriction seems to conflict with one of the main reasons why demand analysis has focused on random choice, namely that the latter appears as the opposite of rational choice, it can be justified from a technical viewpoint. Becker and the experimenters tested RCT under the additional assumption of local-non-satiation, which implies that decision-makers choose a bundle on their budget hyperplane. Requiring random agents to do the same makes the comparison between RCT (the null hypothesis) and the random-choice model (the alternative hypothesis) straightforward.

Regarding the probability law governing the choices of random agents, Bronars (1987), Cox (1997) and others have proposed various versions of random choice in which the probability law mimics the empirical distribution of observed human choices. We discuss briefly the "bootstrapping" random-behavior model which has been used in addition to Becker's random choice in two of the experiments reviewed in Section 6, namely those of Harbaugh, Krause and Berry (2001) and Andreoni and Miller (2002). Imagine an experiment where ten human subjects are confronted with a given budget/price situation. Note that only some bundles among the infinitely many on the budget hyperplane will be chosen by at least one subject, and assume in particular that the subjects' choices concentrate on bundles e_1, e_2, and e_3. For instance one subject selects e_1, five subjects pick up e_2, and the remaining four choose e_3. In the bootstrapping random-choice model, the choice set is restricted to $\{e_1, e_2, e_3\}$, and random agents pick up e_1, e_2, and e_3 with a probability of 0.1, 0.5 and 0.4, respectively.

Bootstrapping aims at an alternative hypothesis that measures the power of GARP more effectively than Becker's random choice. However, by conflating the choices of different individuals, it creates a sort of representative random agent, whose nature and significance are far from clear. In addition, Andreoni and Harbaugh (2008) have pointed out that bootstrapping random agents may violate GARP more or less frequently than Becker's random agents, which undermines the original motivation for the bootstrapping technique. Finally, if one looks for an alternative hypothesis to RCT that is more plausible and has more power than random behavior *à la* Becker, it would be more transparent to employ some explicit alternative model of bounded rationality rather than an opaque form of random behavior such as bootstrapping random choice.

In conclusion, Becker's original model seems to have a number of technical and methodological advantages over other versions of random behavior that explain why it has dominated the literature: it is simple, clear-cut, and keeps hidden assumptions at minimum.

5. Becker and the as-if defense of RCT

Sections 2-4 have investigated how and why the random-choice model has been used in theoretical and experimental analysis of individual demand based on RCT. This Section begins the second part of the paper, which examines the standard as-if defense of RCT from the viewpoint of the random-choice model.

Many have criticized RCT as being psychologically unrealistic. For instance, Simon and other scholars have contended that real decision-makers lack the cognitive capacities to solve the maximization problem hypothesized by RCT. As a defense against the criticism of unrealism, supporters of RCT have often adopted the as-if argument of Friedman (1953), which may be presented as follows: it is not the case that actual decision-makers consciously maximize their utility function when choosing, but it is the economist who rationalizes the decision-maker's choices as if they were generated by utility maximization. Therefore, the utility function and its maximization are in the economist's mind rather than in the decision-maker's, so that the psychology of the latter is not at issue. Insofar as the decision-maker's observed choices agree with those implicated by RCT, the theory is validated.

However, the as-if defense of RCT has been challenged too. Philosophers with sympathies towards scientific realism have disputed the conventionalist or instrumentalist views of scientific theories usually associated with the as-if argument.[17] Psychologists and behavioral economists have called attention to phenomena such as endowment effects and preference reversals that contrast rather than validate the empirical implications of RCT.[18] Our contribution to this debate is to argue that the standard as-if defense provides only a "weak" – in a first, specific sense – support for RCT. Notably, this defense appears weak even if one subscribes to a conventionalist view of scientific theories. Therefore, we need not to enter into the conventionalism *vs.* scientific realism controversy to make our point. In Section 7 we discuss a second sense in which the standard defense appears to be weak.

As mentioned in Section 1, Becker interpreted the convergence between RCT and the random-choice model with respect to their implications on demand behavior as a strengthening of RCT. According to him, this convergence shows that RCT provides a compelling as-if explanation of consumer behavior even if households choose at random. In Section 2.2, we noticed that in fact Becker's convergence results hold only for aggregate demand behavior, but this limitation is irrelevant to the point being made here.

We aim to call attention to the simple but often overlooked fact that the as-if argument only states that RCT offers *a* possible as-if explanation of (aggregate) demand behavior, but does not rule out that different models of choice, such as random choice, may provide an alternative explanation. An as-if-random explanation of demand would go as follows: actual consumers do not choose at random, but instead the economist rationalizes their choices as if they were

generated by a random process; as far as the consumer's observed choices agree with those implicated by the random-choice model, this model is validated.

Thus the very possibility of an as-if-random explanation of consumer demand suggested by Becker makes clear that the standard as-if defense provides only *weak* epistemological support for RCT: it shows only that RCT offers *a possible* explanation of demand behavior without ruling out other possible explanations. A *strong* as-if defense of RCT would require showing not only that RCT provides a possible explanation of consumer demand, but also that it offers the *best* available explanation of it.

The epistemological literature has pointed out that different and often diverging elements may be taken into account to determine which theory emerges as the best explanation. On the one hand, and quite unsurprisingly, the empirical virtues of a theory are important: the better a theory fits with statistical data and experimental findings the better it explains them. On the other hand, more formal virtues of a theory, such as its simplicity, tractability, or generality, are also relevant in determining its preeminence, even though they may be at variance with the theory's empirical virtues. In effect, supporters of RCT often mention its formal virtues as important strengths of the theory that could compensate for its empirical shortcomings and justify its selection as the best explanation at hand.

This is not the place to embark on a philosophical discussion of the best-explanation problem.[19] Here we are simply pointing out that Becker did not address the best-explanation issue and hence did not provide a strong as if-defense of RCT. Moreover, he did not seem to realize that the weak as-if argument alone is a double-edged sword that can be used to support not only RCT but also the random-choice model. More importantly, the way Becker used the as-if argument to defend RCT is also the way supporters of RCT tend to employ it: they claim that the theory provides a possible as-if explanation of observed demand behavior, but they rarely bother to show that RCT explains this behavior better than other competing models of choice. Besides, in a look-for-the-best-explanation approach RCT should be compared with credible competitors, such as some model of bounded rationality, rather than with the implausible benchmark of random choice.

In summary, we distinguish between two types of as-if defenses of RCT and, more generally, of a scientific theory. While the *weak* as-if defense states that a theory offers a *possible* explanation of the phenomenon at hand, the *strong* as-if defense claims that the theory provides the *best* explanation. We argue that the standard as-if defense of RCT is, in this specific sense, a weak one.

To make this point we do not need to dispute the conventionalist view of scientific theories associated with the as-if argument. Insofar as the conventionalist acknowledges that the same phenomenon can be explained by different theories, he necessarily faces the problem of

selecting one theory as the best as-if explanation at hand. The weak as-if argument does not address this selection problem, and thus even the conventionalist should find the argument wanting.

6. Review of demand experiments

In this Section we briefly review the results of six experiments on consumer demand in which Becker's model of random choice was used as an alternative hypothesis against RCT, while in Section 7 we discuss their import for our research question.

6.1 Cox (1997) tested GARP violations using data collected by Battalio, Kagel and others (1973) in an experiment conducted with 38 female patients at a mental hospital in Long Island. The patients were part of a token economy established in the hospital, that is, they could earn tokens by performing janitorial tasks and use them to purchase goods sold within the hospital. By varying the token prices of goods, the experimenters created 7 different budget/price situations. The patients could choose from among 16 goods that were grouped into 3 main categories.[20] Table 1 shows the percentage of GARP violations for patients and random agents, and the percentage of GARP(e) violations for e=0.95 for patients only.

Table 1 – Cox 1997

Group	Percentage of subjects violating GARP	Percentage of subjects violating GARP(0.95)
Patients	36.8%	13.1%
Random agents	48.2%	n.a.

A significant fraction of patients (almost 37 percent) violated GARP and hence RCT. The power of the GARP test is not particularly high, as only around 48 percent of random agents violated GARP. The fraction of human violators shrinks to around 13 percent when GARP(0.95) is considered, but the import of this information is not clear since what happens to random violators with GARP(0.95) is unknown.

6.2 Sippel (1997) tested RCT against the demand behavior of 42 students in law or economics at the University of Bonn, Germany. The students were offered 8 food and leisure goods in 10 different budget/price situations.[21] Sippel ran two experiments, involving 12 and 30 subjects, respectively. Table 2 presents their outcomes.

Table 2 – Sippel 1997

Experiment #	Group	Percentage of subjects violating GARP	Percentage of subjects violating GARP(0.95)
Experiment 1	Law/Eco students	41.7%	8.3%
	Random agents	61.3%	16.8%
Experiment 2	Law/Eco students	63.3%	10.0%
	Random agents	97.3%	12.8%

In Sippel's experiments 24 out of 42 students, or 57 percent of them, violated GARP. The number of GARP violators among random agents, and hence the power of the GARP test, is above 61 percent in both experiments. When GARP(0.95) is considered, the proportion of human violators and hence the power of GARP(0.95) shrinks noticeably.

6.3 Mattei (2000) studied the demand behavior of 20 microeconomics students (experiment 1), 100 business students (experiment 2), and 320 readers of a consumer affairs magazine (experiment 3). In all three experiments the subjects were faced with 8 goods and 20 different budget/price situations.[22] Table 3 shows the findings.

Table 3 – Mattei 2000

Experiment #	Group	Percentage of subjects violating GARP	Percentage of subjects violating GARP(0.95)
Experiment 1	Micro students	25%	0%
	Random agents	99.4%	43.2%
Experiment 2	Business students	44%	4%
	Random agents	98.9%	43.1%
Experiment 3	Magazine readers	32%	2%
	Random agents	98.9%	42.8%

Mattei also recorded a significant proportion of GARP violators among human subjects, namely between 25 and 44 percent. The power of GARP test is excellent, for in each experiment more than 98 percent or random agents violated GARP.

6.4 Harbaugh, Krause and Berry (2001) tested whether children choose rationally. They studied the demand choices of 31 second-grade students aged about 7 years, 42 sixth-grade students aged about 11 and, for comparison, the demand choices of 55 college undergraduates aged about 21. In 11 different budget/price situations children and undergraduates were presented with choice sets including between 3 and 7 bundles. Each bundle contained only 2 commodities, namely potato chips bags and boxes of fruit juice. Table 4 shows the percentage of subjects violating GARP as well as the Afriat Efficiency Index indicating how much the budget constraints should be relaxed to eliminate GARP violations for the three age groups

and for random agents.

Table 4 – Harbaugh, Krause and Berry 2001

Group	Percentage of subjects violating GARP	Afriat Efficiency Index
Second graders, age 7	74%	0.93
Sixth graders, age 11	38%	0.96
Undergraduates, age 21	35%	0.94
Random agents	> 98%	0.648

Harbaugh, Krause and Berry found a significant portion of GARP violators in all three age groups. From age 7 to 11 the number of violators significantly decreases, while from age 11 to 21 it remains more or less at the same level. In no age group were GARP violations particularly severe since they disappear by mildly relaxing the budget constrains.

6.5 In the experiment performed by Andreoni and Miller (2002), 142 students in economics at the University of Wisconsin and Iowa State University had to divide a given number of tokens between themselves and another subject. The tokens were transformed into money, but possibly at different exchange rates. For instance, each token a subject kept for himself became $0.10 while each token he passed to another subject became $0.30, or vice versa. Therefore, an individual had in effect to allocate a given token income between two goods, "money for himself" and "money for another subject", whereby the relative price of the two money-goods could be larger or smaller than one. By modifying the number of tokens to be divided, and the exchange rates of tokens into money, Andreoni and Miller presented the subjects with 8 different budget/price situations, and tested whether their preferences for giving were consistent with RCT. Table 5 shows the findings of the experiment.

Table 5 – Andreoni and Miller 2002

Group	Percentage of subjects violating GARP	Percentage of subjects violating GARP(0.95)
Economics students	9.1%	2.1%
Random agents	78.1%	n.a.

Among the six experiments reviewed, this is the one that records the lowest proportion of human subjects violating GARP. Moreover, the power of the GARP test is satisfactory as more than 78 percent of random agents violated GARP.

6.6 Février and Visser (2004) tested RCT against the demand behavior of 120 individuals from Dijon, France. The subjects were offered 6 different varieties of orange juice in 5 different budget/price situations. Table 6 presents the results of the experiment.

Table 6 – Février and Visser 2004

Group	Percentage of subjects violating GARP	Percentage of subjects violating GARP(0.95)
Dijon consumers	29%	15%
Random agents	22%	5%

In this case GARP violators among random agents turn out to be less numerous than among human subjects. If one considers GARP(0.95), the number of human violators decreases but that of random violators shrinks even more.

7. Demand experiments and the as-if defense of RCT

As a preliminary remark, it is important to notice that in the following discussion we do not dispute the validity of the experimental findings reviewed above. It is always possible to contend that the results obtained in the laboratory rely on the particular experimental design chosen, and that they would disappear under different designs. For instance, one may maintain that the six experiments should have been replicated to check whether GARP violations reduced with repetitions; if this were the case, GARP violators would not be irrational but only require some time to get used to choices in the lab. It is also always possible to bring into play the so-called Duhem-Quine problem and contend that the experimental results are not significant because some auxiliary assumption necessary to test the theory was not fulfilled.[23] Thus, one may claim that the separability assumption between choices made within and outside the experiment was not met, or that the subjects' preference changed during the experiment, so that the GARP violations recorded in it could not be interpreted as violations of RCT. Although these and possibly many other criticisms have a point, we think that they indicate the need for further experimental research rather than invalidating the results obtained. Moreover, the six experimental studies reviewed above have been published in major economics journals, and this suggests that the economics profession acknowledges their findings as provisionally sound. In the following discussion, we do the same.

In the first place the six experiments show that GARP violators are numerous. Violators range from a minimum of 9.1 percent to a maximum of 74 percent, while in most experiments they are around 30-40 percent. This outcome stands in contrast to RCT, which implies no GARP violations, and makes apparent the second sense in which the standard as-if defense of RCT appears to be weak. As mentioned in Section 5, the as-if argument focuses on the empirical implications of a theory and states that RCT is validated as far as observed choices agree with those it predicts. Since the experimental evidence suggests that this is not the case, the as-if argument would entail that RCT is disproved rather than validated by demand behavior. In other words, given the fact that in most experiments GARP violators are around 30-40

percent, the as-if argument would entail a criticism of RCT rather than a defense of it.

This is a second type of weakness of the as-if argument that adds to the one discussed in Section 5. There we pointed out that the as-if defense of RCT is weak in the sense that it only states that the theory offers a possible explanation of demand behavior without showing that it provides the best explanation. Here we notice that the as-if defense is a weak one because it centers on RCT's fit with observed behavior although this fit is poor.

At best, the as-if argument could be used to defend an almost-RCT claiming that *most* individuals choose in a *nearly* rational way. In effect when GARP(0.95) is considered, human violators range from a minimum of 0 percent to a maximum of 15 percent and in most experiments are below 10 percent. However with GARP(0.95) the proportion of random violators also shrinks significantly, so that the power of GARP(0.95) as test for almost-RCT is wanting.

A second outcome of the experiments is that the number of GARP violators, and thus the consistency of RCT with observed choices, varies strongly from experiment to experiment. In some experiments, such as the Andreoni-Miller one, RCT fits observed choices quite well, whereas in others, such as Sippel's, the theory is conspicuously at odds with experimental evidence.[24]

The variance of RCT's fit may depend on various factors, such as the amount of time available to the subjects to decide, the finite or nearly infinite number of bundles in their choice sets, or the order in which the different choice sets are presented to them. Yet, no general indication can be drawn from the experiments. The only tendency that seems traceable is that the more goods and budget/price situations the subjects face, the more they violate GARP. There are exceptions, though: in Mattei's experiment the subjects were confronted with 8 goods in 20 budget/price situations while in Sippel's they faced 8 goods in 10 situations, but Mattei recorded fewer violators than Sippel.

In summary, there is no satisfactory meta-theory stating in which situations RCT works and in which it doesn't. This conclusion is relevant for the best-explanation issue. A commonly accepted distinction between RCT and bounded-rationality models is that the former is a single theory with an ambition to have general explanatory power, while the latter constitute a whole range of different decision rules that are tailor-made to cater for particular contexts but are not always transferable to other sets of circumstances. If the formal virtues of theories contribute to determining which one emerges as the best explanation, RCT's generality, in contrast with the particularity of bounded-rationality models, seems to give the former an advantage over the latter. However, the experimental finding that RCT satisfactorily fits certain patterns of choice but not others, and hence appears to have at best local rather than general explanatory power, weakens RCT's pretense of generality and thus its claim to

provide the best explanation of human demand behavior.

Third, in Cox's study, in Sippel's experiment 1, and in the test performed by Février and Visser, the proportion of GARP violators among human subjects is significantly different from the proportion predicted by RCT, i.e. zero, and is instead quite similar to the proportion of random GARP violators. One may be tempted to infer that at least in certain situations the random-choice model fits human behavior better than RCT. This inference is, however, incorrect. The experimental evidence collected in those three experiments only says that human subjects and random agents violate GARP (and hence RCT) in a similar proportion, not that the two groups behave in a similar way. Indeed, the choice patterns of humans and random agents could be highly diverse and the two groups may violate GARP in very different ways. If this is the case, the random-choice model would fit human demand behavior as poorly as RCT.

From a statistical viewpoint, this may be seen by noting that the GARP test is constructed to check the null hypothesis that RCT fits human demand behavior, not to check the alternative hypothesis that the random-choice model does. To investigate this latter issue would require an explicit statistical test in which the null hypothesis is that the recorded choices of each human subject come from a uniform distribution on the subject's choice sets, whereas the alternative hypothesis is that they do not.

In effect, one may imagine that at least in certain circumstances, e.g. when the available alternatives are very numerous or it is difficult to evaluate and compare them, subjects might actually choose at random. For instance, in the Février-Visser experiment subjects had to choose among almost identical commodities (six varieties of orange juice), and this awkward choice situation might have induced them to pick up alternatives in a random way. However, the six experimental studies did not investigate whether the random-choice model might sometimes fit human demand behavior well, or even better than RCT. Random choice remained an implausible benchmark used to measure the power of the GARP test for RCT, and never rose to the role of credible alternative to RCT.[25]

More generally, the experimental studies were not designed to compare RCT's goodness-of-fit to human demand behavior with the goodness-of-fit of some other competing model of choice, may this be the implausible random-choice model or some more credible model of bounded rationality. This appears as a methodological shortcoming of the six experiments which is related to the widespread understanding of the as-if defense of RCT in its first weak sense. As far as the only question at stake is whether RCT provides a possible explanation of demand behavior, the experiments will aim at falsifying or validating RCT and will posit RCT as the unique null hypothesis under examination. It is only if one wants to make a strong case for RCT and to show that it offers the best explanation of consumer behavior, that the goodness-

of-fit of different competing choice models will be investigated.

8. Summary and conclusions

In this paper we have investigated how and why the random-choice model has been used in theoretical and experimental analysis of individual demand based on RCT. We have also examined the standard as-if defense of RCT from the viewpoint of the random-choice model and argued that this defense is a weak one in a twofold sense.

We have pinpointed the differences between random behavior in Becker and later experimental research. While Becker was interested in RCT's implications for aggregate demand and focused on the expected or average outcome of random choice, experimentalists investigated RCT's implications for individual demand and took into account each and every choice made by random agents. When this is done, it is apparent that random agents frequently violate RCT, so that Becker's convergence results between rational and random behavior fade away. In effect, the experimentalists used the random-choice model only as an alternative hypothesis to assess the power of GARP as a test for the null hypothesis RCT; random choice never became a credible null hypothesis to test.

We have argued that focus on random choice in demand analysis is due to the fact that it contrasts rational choice more radically than any form of boundedly-rational choice. This circumstance renders the random-choice model an important benchmark for research about RCT, in the sense that if observed behavior can be explained as the result of either random or rational decisions, it seems to be due to factors unrelated to rationality. Another important factor is that traditional bounded-rationality models often do not have clear-cut behavioral implications and so it is difficult to test them.

We have noted that the very possibility of an as-if-random explanation of demand suggested by Becker makes clear that the as-if defense is a weak one: it states only that RCT offers a possible explanation of demand behavior but not that RCT provides the best available explanation of it. Insofar as conventionalists face the problem of selecting the best explanation of demand behavior, this weakness concerns them too.

Based on the evidence provided by six experiments that test RCT by using random choice as an alternative hypothesis, we have pointed out a second weakness of the as-if defense of RCT. Since the as-if defense concentrates on RCT's fit with observed behavior, and the experiments show that this fit is often poor, the as-if argument turns out to be a weak one. The very variance of RCT's fit with observed choices suggests that it has at best local rather than general explanatory power. This renders RCT more similar to the various theories of bounded rationality than is usually assumed, and undermines a common argument in favor of RCT, namely that its generality would compensate for its empirical shortcomings and justify its

selection as the best explanation at hand.

Finally, we have called attention to the fact that none of the experiments on individual demand considered compared RCT's goodness-of-fit to human behavior with the goodness-of-fit of some competing, and possibly more credible, model of choice. We have argued that this is a methodological shortcoming in the experimental research on individual demand, and that it is related to the as-if defense of RCT: if one looks only for a possible as-if explanation of individual demand instead of searching for the best explanation, he will not engage in comparative testing between competing models of choice.

The search for the best explanation is more common in other areas of experimental economics. For instance, already in the 1970s Fiorina and Plott (1978) used experiments to investigate which among different competing models of committee decisions performed best. Plott (1986, p. S302) expressed the view that for those who study experimental markets "the rejection of a theory […] is not an especially challenging research objective"; rather for them "the question becomes which of several competing models is the most accurate, fully realizing that the best model might still be 'poor.'". More recently, Bateman, Kahneman and others (2005) experimentally tested competing theories of loss aversion, while in the study of behavior under uncertainty Halevy (2007) used an experiment to compare alternative theories of ambiguity aversion. In the paper on intertemporal decision by Manzini and Mariotti (2010b) already mentioned in Section 3, the authors tested RCT against their "Categorize Then Choose" model and concluded that the later fits experimental data better than RCT.

The difference between the reject-or-validate-RCT approach typical of experimental research in individual demand and the search-for-the-best-explanation approach of other areas of experimental economics is due to various factors. While in demand analysis the primacy of RCT has always been so firm that it might have appeared as the only alternative to test, in the analysis of committee decisions, loss aversion, ambiguity aversion or intertemporal decisions a bulk of competing models have been put forward. This proliferation of models has probably stimulated experimental research aimed at comparing them. Moreover, the fact that the behavioral implications of most bounded-rationality models were obscure made their comparison with RCT difficult and thus probably contributed to maintaining an exclusive attention on RCT. Finally the focus of demand experiments on the rejection or validation of RCT can be also traced back to the conviction that showing that RCT provides a possible as-if explanation of consumer behavior already constitutes a robust defense of it.

These factors, however, seem to be vanishing. The new models of bounded-rationality often have a precise choice-theoretic characterization that facilitates comparison with RCT. In the last two decades behavioral economics has called into question the supremacy of RCT more effectively than previous criticisms, often dating back to the late-nineteenth century. Lastly, in

this paper we hope to have exposed the epistemological weakness of the standard as-if defense of RCT. Thus we anticipate that future experimental research in demand analysis will abandon the reject-or-validate-RCT approach, and adopt the epistemologically sounder search-for-the-best-explanation method.

This version: May 30, 2010

Word count: 9360

Notes

[1] Tubaro (2009) discusses instead random choice in market experiments like the one performed by Gode and Sunder (1993).

[2] For a presentation of RCT in contexts involving uncertainty, strategy and time see Mas-Colell, Whinston and Green (1995), chapters 6, 8-9, and 20. An example of extended RCT that incorporates experiences and social forces into the theory is provided by Becker (1996).

[3] A consumer has locally non-satiated preferences when for any consumption bundle x there exists another bundle y arbitrarily close to x which is strictly preferred to x by the consumer.

[4] This kind of compensation is called Slutsky's compensation. With Hicksian compensation, instead, the consumer is compensated so that her utility level is kept constant when prices change. Both kinds of compensations induce negative substitution effects, but Slutsky's are those used in empirical studies because they can be determined even without knowing the consumer's utility function.

[5] This is the so called "law of market demand." Exceptions to the law are represented by Giffen goods, which however are rare for individual demand, and extremely implausible for market demand. For an analysis of Giffen goods in individual and market demand see Battalio, Kagel and Kogut (1991).

[6] For a history of the early experiments on demand behavior, see Moscati (2007).

[7] An earlier strand of experimental research employed pigeons and rats as experimental subjects, and checked whether their demand displayed a negative substitution effect. This research is summarized in Kagel, Battalio and Green (1995).

[8] On GARP, see Varian (1982). GARP is a modification of the Weak Axiom of Revealed Preference (WARP) proposed by Samuelson (1938), and the Strong Axiom of Revealed Preference (SARP) proposed by Houthakker (1950). WARP allows for cyclical choices, which are instead excluded by RCT, and excludes indifference curves with straight segments which are compatible with RCT. SARP rules out cyclical choices but still excludes straight indifference curves. GARP rules out cycles and allows for straight indifference curves, thereby providing a complete behavioral characterization of RCT.

[9] The intuition is rough since cyclical choices, which are ruled out by GARP, may materialize only when at least three commodities and three budget/price situations are involved.

[10] On the separability assumption, see Varian (1983, 1988).

[11] For a detailed discussion of the power of the GARP test, see Bronars (1987) and Andreoni and Harbaugh (2008).

[12] In addition to the percentage of random agents that violate GARP, one could also employ the percentage of GARP violations as a measure of GARP's power; indeed this measure is computed by some experimenters. The main problem with it is that there are different ways to count GARP violations. For instance, choices like those in Figure 2d count as one violation in some experiments, and are regarded as two violations in others. We focus on the percentage of random agents that violate GARP because this measure is univocally determined, was calculated in all experiments reviewed in Section 6, and no significant new insight is gained by combining it with the percentage of GARP violations.

[13] We broadly refer to bounded rationality to indicate any decision model diverging from RCT. Accordingly, bounded rationality encompasses here also decision models that may be labeled as "behavioral".

[14] On the importance of characterizing models of bounded rationality in terms of their behavioral implications see also Spiegler (2008).

[15] In "Categorize Then Choose" (CTC) the agent decides in two stages. In the first one she categorizes the alternatives in broad classes and focuses on one class; then in the second stage she chooses an alternative from that class. For instance, a CTC agent categorizes restaurants by type of cuisine and focuses on Mexican restaurants; then she chooses the preferred Mexican restaurant. CTC agents may violate RCT.

[16] In applied demand analysis, McFadden (1974) and others have introduced an extension of RCT called Random Utility Model (RUM). As in RCT, in RUM choice is the result of utility maximization but the utility

function V is the sum of a standard RCT utility U and a random component ε: $V = U + \varepsilon$. The random component is meant to accommodate for measurement errors on the part of the observer. The randomness present in RUM is thus very different from the one we discuss here.

[17] See for instance Hausman (1992).

[18] See for instance Tversky and Thaler (1990), and Kahneman, Knetsch and Thaler (1991).

[19] For a discussion, see among others Thagard (1978) and Lipton (2004).

[20] The goods were: cigarettes, coffee, two types of candy, cookies, soda, milk, meal deal with a cigarette (category one); private dormitory room, private locker, grounds pass to leave the ward for a fixed period of time (category two); repeated use of the ground pass, clothes, weekly dance, breakfast, different rights such as right to use cash for packages from home (category three).

[21] The goods were: Coca-Cola, orange juice, coffee, licorice, snacks, music video clips, computer games, magazines.

[22] In experiment 1 the goods were: milk chocolate, salted peanuts, biscuits, text markers, ball-point pens, plastic folders, writing pads, post-it. In experiment 2, milk chocolate, biscuits, orange juice, iced tea, writing pads, plastic folders, diskettes, post-it. In experiment 3, milk chocolate, biscuits, orange juice, iced tea, post-it, audiocassettes, ball-point pens, batteries.

[23] On the Duhem-Quine problem, see Hands (2001, chapter 3).

[24] D. Satz and J. Ferejohn (1994) discussed the variance of RCT's explanatory power and its dependency on the environment the theory is applied to. In particular, they suggested that rational-choice explanations are most plausible "under conditions of scarcity, where human choice is severely constrained" while "in environments without strong constraints, agents will not generally behave as the theory predicts." (p. 81). However, this argument hardly helps to understand RCT's violations and the variance of RCT's explanatory power across the six experiments, where subjects faced almost the same choice situation, and the Satz-Ferejohn scarcity condition for rational choice was satisfied, for the subjects' choices were always budget-constrained.

[25] Actually, Cox (1997, p. 1076) suggested a rough test to compare RCT and the random-choice model, and argued that the individual's choices in his experiments were more consistent with RCT than with random choice. However, not even Cox examined the issue in detail.

References

Afriat, S. (1967), 'The Construction of a Utility Function From Expenditure Data', *International Economic Review*, 8, 67-77.

Afriat, S. (1972), 'Efficiency Estimates of Productions Functions', *International Economic Review*, 13, 568-598.

Alchian, A.A. (1950), 'Uncertainty, Evolution, and Economic Theory', *Journal of Political Economy*, 58, 211-221.

Andreoni, J., and Harbaugh, W. (2008), 'Power Indices for Revealed Preference Test', mimeo, University of California, San Diego.

Andreoni, J., and Miller J. (2002), 'Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism', *Econometrica*, 70, 737-753.

Bateman, I., Kahneman, D. and others (2005), 'Testing Competing Models of Loss Aversion: An Adversarial Collaboration', *Journal of Public Economics*, 89, 1561-1580.

Battalio R.C., Kagel, J.H., and Kogut, C.A. (1991), 'Experimental Confirmation of the Existence of a Giffen Good', *American Economic Review*, 81, 961-970

Battalio, R.C., Kagel, J.H., and others (1973), 'A Test of Consumer Demand Theory Using Observations of Individual Consumer Purchases', *Western Economic Journal*, 11, 411-428.

Becker, G.S. (1962), 'Irrational Behavior and Economic Theory', *Journal of Political Economy*, 70, 1-13.

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

Bronars, S.G. (1987), 'The Power of Nonparametric Tests of Preference Maximization', *Econometrica*, 55, 693-698.

Chant, J.F. (1963), 'Irrational Behavior and Economic Theory: A Comment', *Journal of Political Economy*, 71, 505-510.

Cox, J.C. (1997), 'On Testing the Utility Hypothesis', *Economic Journal*, 107, 1054-1078.

Février, P., and Visser, M. (2004), 'A Study of Consumer Behavior Using Laboratory Data',

Experimental Economics, 7, 93-114.

Fiorina, M.P., and Plott, C.R. (1978), 'Committee Decisions under Majority Rule: An Experimental Study', *American Political Science Review*, 72, 575-598.

Friedman, M. (1953), 'The Methodology of Positive Economics', in M. Friedman, *Essays in Positive Economics*, Chicago: University of Chicago Press, pp. 3-43.

Gode, D.K., and Sunder, S. (1993), 'Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality', *Journal of Political Economy*, 101, 119-137.

Halevy, Y. (2007), 'Ellsberg Revisited: An Experimental Study", *Econometrica*, 75, 503-536.

Hands, D.W. (2001), *Reflection without Rules*. Cambridge: Cambridge University Press.

Harbaugh, W., Krause, K., and Berry, T. (2001), 'GARP for Kids: On the Development of Rational Choice Behavior' *American Economics Review*, 91, 1539-1545.

Hausman, D. (1992), *Essays on Philosophy and Economic Methodology*, Cambridge: Cambridge University Press.

Houthakker, H.S. (1950), 'Revealed Preference and Utility Function', *Economica* [NS], 17, 159-174.

Kagel, J.H., Battalio, R.C., and Green, L. (1995), *Economic Choice Theory: An Experimental Analysis of Animal Behavior*, New York: Cambridge University Press.

Kahneman, D., Knetsch, J.L., and Thaler R.H. (1991), 'Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias', *Journal of Economic Perspectives*, 5, 193-206.

Lipton P. (2004), *Inference to the Best Explanation*, London: Routledge.

Manzini, P., and Mariotti, M. (2007), 'Sequentially Rationalizable Choice', *American Economic Review*, 97, 1824-1839.

Manzini, P., and Mariotti, M. (2009), 'Consumer Choice and Revealed Bounded Rationality', *Economic Theory*, 41, 379-392.

Manzini, P., and Mariotti, M. (2010a), 'Categorize Then Choose: Boundedly Rational Choice and Welfare', mimeo, University of St. Andrews.

Manzini, P., and Mariotti, M. (2010b), 'Revealed Preferences and Boundedly Rational Choice Procedures: an Experiment', mimeo, University of St. Andrews.

Masatlioglu, Y., and Ok, E.A. (2005), 'Rational Choice with Status Quo Bias', *Journal of Economic Theory*, 121, 1-29.

Mas-Colell, A., Whinston M.D., and Green, J.R (1995), *Microeconomic Theory*, New York: Oxford University Press.

Mattei, A. (2000), 'Full-Scale Real Tests of Consumer Behavior Using Experimental Data', *Journal of Economic Behavior and Organization*, 43, 487-497.

McFadden, D. (1974), 'Conditional Logit Analysis of Qualitative Choice Behavior', in P. Zarembka (ed.), *Frontiers of Econometrics*, New York: Academic Press, pp. 105-142.

Moscati, I. (2007), 'Early Experiments in Consumer Demand Theory: 1930-1970', *History of Political Economy*, 39, 359-401.

Plott, C.R. (1986), 'Rational Choice in Experimental Markets', *Journal of Business*, 59, S301-S327.

Samuelson, P. (1938), 'A Note on the Pure Theory of Consumer's Behaviour', *Economica* [NS], 5, 61-71.

Satz, D., and Ferejohn, J. (1994), 'Rational Choice and Social Theory', *Journal of Philosophy*, 91, 71-87.

Sippel, R. (1997), 'An Experiment on the Pure Theory of Consumer's Behavior', *Economic Journal*, 107, 1431-1444.

Spiegler, R. (2008), 'On Two Points of View Regarding Revealed Preferences and Behavioral

Economics', in A. Caplin and A. Schotters (eds), *The Foundations of Positive and Normative Economics*, New York: Oxford University Press, pp. 95-115.

Thagard, P.R. (1978), 'The Best Explanation: Criteria for Theory Choice', *Journal of Philosophy*, 75, 76-92.

Tubaro, P. (2009), 'Is Individual Rationality Essential to Market Price Formation? The Contribution of Zero-Intelligence Agent Trading Models', *Journal of Economic Methodology,* 16, 1-19.

Tversky, A., and Thaler, R.H. (1990), 'Anomalies: Preference Reversals', *Journal of Economic Perspectives*, 4, 201-211.

Varian, H.R. (1982), 'The Nonparametric Approach to Demand Analysis', *Econometrica*, 50, 945-73.

Varian, H.R. (1983), 'Non-Parametric Tests of Consumer Behaviour', *Review of Economic Studies*, 50, 99-110.

Varian, H.R. (1988), 'Revealed Preference with a Subset of Goods', *Journal of Economic Theory,* 46, 179-185.

Varian, H.R. (1991), 'Goodness of Fit for Revealed Preference Tests', mimeo, University of Michigan.