**Ivan Moscati**

Bocconi University, Department of Economics and London School of Economics, CPNSS

**Paola Tubaro**

Centre Maurice Halbwachs-Centre National de la Recherche Scientifique

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

This draft: July 2009

**Abstract**

Rational choice theory (RCT) models decision makers as utility maximizers and is often defended via an as-if argument. According to this argument, although real individuals do not consciously maximize their utility function, their choices can be explained as if they were generated by utility maximization. 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 this paper we examine a series of experiments that compare RCT and the random-choice model as alternative explanations of consumer demand, and investigate how these experiments contribute to clarifying the actual scope of RCT and the shortcomings of the standard as-if defense of it.

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

**JEL codes**: B410, C140, D110, D120.

Correspondence may be addressed to Ivan Moscati at ivan.moscati@unibocconi.it or Paola Tubaro at paola.tubaro@ens.fr. We are grateful to Marco Bonetti, Richard Bradley, Marco Dardi, Franz Dietrich, Jean-Sébastien Lenfant, Aldo Montesano and Piero Veronese for helpful discussions and suggestions on previous drafts of the paper. We also thank the participants in seminars at the London School of Economics, the Choice Group, at Université Paris Ouest Nanterre La Défense, 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

In mainstream economics, 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. 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 consumer demand theory and forms the subject matter of the present paper.[1]

Many have criticized RCT as psychologically unrealistic. For instance, Herbert Simon and other scholars have contended that actual 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 popularized in economics by Milton Friedman (1953). According to this argument, it is not the case that actual decision makers consciously maximize their utility function when choosing. Instead, 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. As far 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 also been challenged. Philosophers have disputed the conventionalist or instrumentalist views of scientific theories usually associated with the as-if argument.[2] Psychologists and behavioral economists have provided robust experimental evidence suggesting that RCT is not validated by observed choices.[3] The present paper addresses the as-if defense of RCT in its consumer-demand version from a further angle, namely that involving random choice.

A decision maker chooses at random when he picks up an alternative in his choice set according to a uniform distribution on the set, where each alternative in the set has an equal probability of being selected. In particular, we will consider consumers who pick up at random a bundle from among those that exhaust their whole income, i.e., a bundle on their budget line or, when more than two goods are available, budget hyperplane. Random decision makers can be easily constructed by computer, and are often employed in agent-based computational economics.[4] One may expect rational and random decision makers to behave very differently. However, Gary Becker (1962) pointed out that some implications of RCT

[1] For a presentation of consumer-demand RCT, see Mas-Colell, Whinston and Green 1995, chapters 1-3.
[2] See among others Nagel 1963, Caldwell 1980, Musgave 1981, Hausman 1992.
[3] See for instance Tversky and Thaler 1990, and Kahneman, Knetsch and Thaler 1991.
[4] For an introduction to agent-based computational economics, see Tesfatsion 2006.

for demand behavior are, on average, still valid even if consumers choose in a random way. Becker's article has inspired a series of recent experiments that employ the random-choice model as an alternative hypothesis in testing the extent to which consumer demand behavior satisfies RCT. In the present paper we review six of the most relevant among these experiments and investigate how they contribute to clarifying the actual scope of RCT as well as the shortcomings of the standard as-if defense of it.[5]

The experiments show that consumers often violate RCT and hence suggest that it is difficult to see human demand behavior as if generated by RCT. At best one might defend an "almost RCT", claiming that *most* individuals choose as if they were *nearly* rational. Moreover, the number of violations varies strongly from experiment to experiment and this hints that RCT's explanatory power depends very much on context. The problem is that there is no meta-theory stating in which contexts RCT works and in which ones it does not.

With respect to the as-if argument, the experiments reinforce Becker's article in bringing to light the simple but often overlooked fact that there are two types of as-if defenses 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 of it in terms of fit with experimental data, simplicity or other relevant qualities. The very possibility of an as-if-random explanation of consumer demand makes clear that the standard as-if defense provides only weak epistemological support to RCT, and does not rule out that different models of choice, such as random-choice, may provide an alternative explanation.

The problem is not the weak defense *per se*, for it is well understood that the same phenomenon can be explained by two or more different theories. One may even claim that since theory and the real world are necessarily dissimilar, any scientific theory constitutes an as-if rational reconstruction of the world. But it is precisely the presence of multiple as-if explications that limits the epistemological force of the weak as-if defense and calls for selecting one theory as the best explication at hand.

Becker did not raise this problem in his article, though, and the subsequent experiments were also oriented by the weak as-if defense of RCT. Indeed, by design they either falsify or validate RCT and do not investigate whether RCT explains human demand behavior better than other models of choice and, in particular, better than the random-choice model. As a consequence of the weak-defense attitude, the explanatory power of the random-choice model has never really been investigated in consumer demand analysis.

[5] Tubaro 2009 discusses instead rational and random choice in market experiments.

The gap between the weak and the strong as-if defense of RCT is particularly relevant when policy issues are at stake, for instance if policy makers aim to influence consumers' decisions about smoking, drinking, or buying healthy food. The circumstance that RCT offers a possible explanation of demand behavior (weak defense) does not sufficiently argue for a RCT policy, e.g. an incentive-based policy, as the most effective one. As far as the random-choice model provides another possible as-if explanation of consumer behavior, by the same argument one could also claim that incentives would prove ineffectual and a different policy should be adopted, e.g. a paternalistic policy imposing severe restrictions on the consumers' choice set. Further and different as-if explanations of consumer behavior would suggest further and different policies. In order to argue that a RCT-based policy is the most effective one, one would need to make a strong case for RCT, that is, to show that RCT offers the best as-if explanation of consumer behavior among the various explanations at hand.

Although this paper focuses on basic consumer-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 of RCT that attempt to capture phenomena neglected by standard RCT. In fact, sophisticated versions and extensions of RCT also construct 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, the difference between the weak and the strong as-if argument is relevant also for sophisticated and extended RCT.[6]

The remainder of the paper is organized as follows. Section 2 illustrates why random choice has become a focal point for research on RCT. In Section 3 we present Becker's approach to random behavior and discuss his main results and claims. In Section 4 we review the six experiments that test RCT by using the random-choice model as an alternative hypothesis, while in Section 5 we discuss their main findings and the import of the latter for our research questions. Section 6 sums up and concludes.

## 2. Random, boundedly rational, and rational choice

As RCT allows for an infinite number of theoretically admissible alternatives, one wonders why much experimental research focuses on random choice rather than, for instance, some model of bounded rationality. In effect, it could be argued that random choice is as unrealis-

[6] 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.

tic 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 some 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 in fact 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 in random choice efficiency is not even pursued 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 established practice of bounded-rationality theorists who often model 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 circumstance that RCT and the random-choice model frame choice in a similar way makes the comparison between them easier.

Third, since, among the disparate theories of bounded rationality none seems to have a definite pre-eminence over the others, it is not obvious which one should be compared with RCT. Random choice provides instead a clear-cut benchmark to be contrasted with RCT.

One might object that not even random choice is that clear-cut, as the specification of the

choice set from which random agents pick up an alternative allows for some degree of freedom. For instance, John Chant (1963) put forward a 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. It can be shown that random agents choosing in this way (labeled by Chant as "impulsive goods deciders") rarely violate RCT.[7] As a result, they are of little help in testing RCT and in fact are not employed in the experimental studies discussed in Section 4. Even if one sticks to choice among bundles, the random consumer may be thought of as picking up a bundle among all those he can afford, that is, not only among the bundles *on* his budget hyperplane, but also among those *below* it. Although this option is feasible, all the experimental studies we consider restrict random choice to the budget hyperplane. This is because they test RCT under the additional assumption that consumers have locally non-satiated preferences, i.e., that for any consumption bundle $x$ there exists another bundle $y$ arbitrarily close to $x$ which is strictly preferred to $x$ by the consumer. Since local non-satiation implies that a rational consumer chooses a bundle on her budget hyperplane, comparison between rational and random choice requires the random consumer to pick up a bundle on his budget hyperplane too. In sum, despite some degrees of freedom, defining the choice set of a random agent seems less arbitrary than opting for one specific model of bounded rationality. Moreover, all experimental studies adopt the budget hyperplane as the choice set and this consensus makes arbitrariness concerns less pressing.[8]

Finally, random choice has a practical advantage over bounded rationality, namely that a computer program can easily generate random agents, whereas constructing artificial agents that follow boundedly rational decision rules can be much trickier.

## 3. Random and rational behavior in Becker 1962

### *3.1 Becker's results*

A rational consumer with locally non-satiated preferences chooses a bundle on his budget hyperplane, and among other things his demand displays a negative substitution effect. This means that when prices change and the rational consumer is compensated so that at the new prices she can just afford the bundle she chose at the old prices, her demand for the relatively dearer commodities will decrease.[9] A negative substitution effect is a necessary con-

[7] On this point see Bronars 1987.

[8] For further discussion on the different possible types of random behavior, see Harbaugh, Krause, and Berry 2001, and Andreoni and Harbaugh 2008.

[9] Becker considers Slutsky's compensations. With Hicksian compensations, instead, the consumer is compensated so that her utility level is kept constant when prices change. Both kinds of compensations induce

dition 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) imagines a consumer who chooses a bundle on his budget hyperplane in a random way, and compares the implications of this random decision rule with those of RCT. Becker considers the two-commodity case so that in his paper the budget hyperplane comes down to the more familiar 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, 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 $y$. This may also be seen by observing that in selecting a bundle on the budget line the consumer is implicitly choosing what fraction of his income he is going to spend on each commodity. Since he chooses following a uniform distribution, he is expected to spend an equal fraction of his income on each good and so to purchase $I/2p_x$ of $x$ and $I/2p_y$ of $y$.[10]

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.

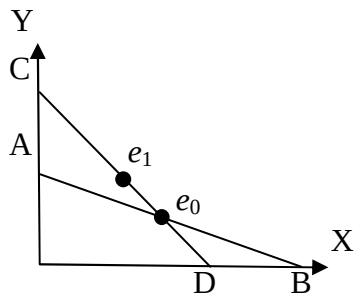


*Figure 1*: Random choice implies a negative expected substitution effect

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.

[10] Since a consumer who chooses randomly what fraction of his income to spend on each commodity and a consumer who chooses randomly a bundle on his budget line behave in the same way, Chant (1963) labeled both as "impulsive money deciders."

When economists study markets rather than individual behavior, they are chiefly interested in consumers' aggregate and non-compensated demand, that is, in market demand. 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. 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.[11] Becker notes that the random-choice model also implies the law of market demand 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.

Becker does not see the convergence between RCT and the random-choice model, with respect to the sign of the substitution effect and the slope of market demand, as a possible problem for the as-if defense of RCT. According to him, this convergence shows in fact that RCT and the random-choice model provide two alternative, but to a certain extent equally plausible as-if explanations of consumer behavior: "Households can be said to behave not only 'as if' they were rational but also 'as if' they were irrational." (Becker 1962, p. 4) Those results can even strengthen the as-if defense of RCT, in the sense that they show that RCT provides a compelling as-if explanation of household behavior even if consumers choose at random.

### *3.2 Discussion of Becker's results*

Becker's article raises a number of questions and comments. To begin with, the convergence between rational and random consumers with respect to the sign of the substitution effect draws from Becker's 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

[11] For an analysis of Giffen goods in individual and market demand see Battalio, Kagel and Kogut 1991.

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 $e_0$D 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.

An important difference between Becker’s article and the experimental studies on consumer demand we discuss in Section 4 is that the latter take into account each and every choice made by random agents, and not only the expected or average outcome of their choices. Becker refers to the expected outcome because his main interest lies in the negative slope of aggregate demand rather than in the properties of individual demand. As noticed in Section 3.1, the expected outcome of random choice is sufficient to obtain a negatively sloped market demand curve. The focus of the experiments discussed in Section 4 is instead on individual demand and the rationality issues related to it. As we will see, random behavior becomes relevant to these topics only if all choices made by random consumers are considered.

Becker’s conclusion that RCT and the random-choice model provide two interchangeable as-if explanations thus holds only for aggregate demand behavior. Even in this more circumscribed sense, Becker’s point is important for the purposes of this paper because it shows that the as-if argument alone provides weak epistemological support for RCT: it only states that RCT offers a possible explanation of aggregate demand behavior and does not rule out that random choice may provide an alternative explanation. An as-if-random explanation of aggregate demand would go as follows: actual consumers do not choose at random, but it is the economist who 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.

A strong case for RCT would require showing not only that RCT provides a possible explanation of consumer demand, but that it offers the best explanation. 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 explanatory power, even though they may be at variance with its empirical virtues. This is not the place to embark in a philosophical discussion of the best-explanation problem.[12] Here we only point out that Becker's as-if defense of RCT does not take into consideration the best-explanation issue: it makes the weak case that RCT offers a possible explanation of demand behavior, not the strong case that RCT provides the best explanation of it.

The difference between the weak and the strong as-if defense of RCT proves important not only from an epistemological point of view, but also when policy issues are at stake. Policies inspired by RCT, e.g. an incentive-based policy, seem often to be supported by an inference of the following kind: since RCT provides an as-if explanation of consumer behavior, *then* a RCT policy is an effective one. However, the weak as-if defense of RCT does not provide a sufficient footing for such an inference. By the same argument one could argue that since the random-choice model offers another possible as-if explanation of consumer behavior, *then* incentives would be ineffectual. In order to argue that a RCT-based policy is effective or even the most effective one, one would need to make a strong case for RCT, that is, to show that RCT offers the best as-if explanation of consumer behavior among the various explanations at hand.

## 4. Random and rational behavior in experiments on individual demand

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 in laboratory experiments. In effect, when Becker published his 1962 article experimental research was still a niche within economics.[13] 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. In particular, we focus here on those experimental tests of RCT that consider random choice as the alternative decision rule that could have generated the data recorded in the experiments.

These experiments share a similar design. Each human subject participating in the experiment is 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 checks whether they satisfy the Generalized

[12] For a discussion, see among others Thagard 1978 and Lipton 2004.

[13] For a history on the early experiments on demand behavior see Moscati 2007.

Axiom of Revealed Preference (GARP). 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.[14]

Figures 2a-2e below give a rough geometrical intuition about which choices satisfy GARP and which instead violate it in the two-commodity case.[15] 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.

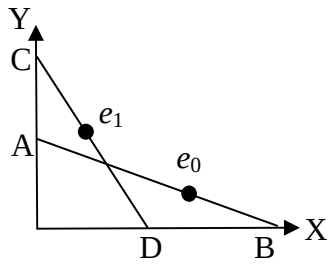


*Figure 2a*: GARP satisfied

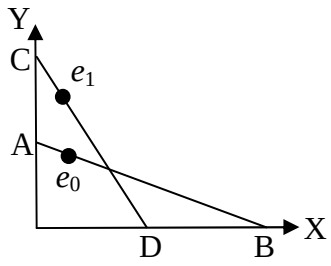


*Figure 2b*: GARP satisfied

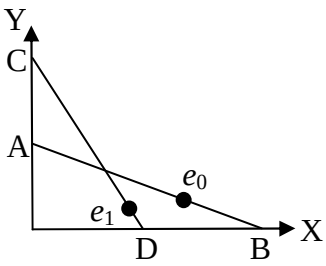


*Figure 2c*: GARP satisfied

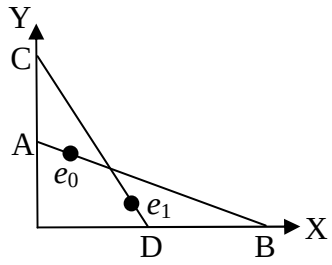


*Figure 2d*: GARP violated

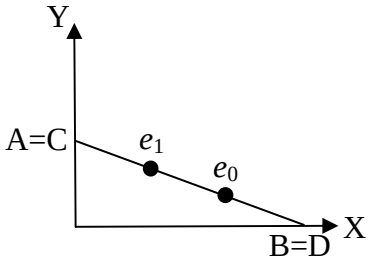


*Figure 2e*: GARP satisfied

[14] 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.

[15] 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.

There is, however, a problem with GARP as a test for RCT: even if all recorded choices pass the 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 the budget hyperplanes 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.[16]

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. By 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.[17]

---

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

[17] 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 of the 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 prefer to focus on the percentage of random agents that violate GARP be-

It is important to stress again 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 3.2, 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 behavior ceases to be appropriate.

The GARP violations' relative lack of severity presents a further issue. It turns out that 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 one may doubt whether it is indeed appropriate to reject RCT. Building on the work of Sidney Afriat (1967, 1972), Hal 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".

At least six experimental works have tested RCT using the random-choice model as an alternative hypothesis to assess the power of GARP and GARP(*e*) tests. Below we briefly review these works, while in Section 5 we discuss their import for our research questions.

---

cause this measure is univocally determined, was calculated in all experiments, and no significant new insight is gained by combining it with the percentage of GARP violations.

**4.1** James Cox (1997) tested GARP violations using data collected by Battalio, Kagel, and others (1973) in an experiment conducted with 38 female patients at Central Islip State Mental Hospital, 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.[18] Cox also included in the choice sets labor supplies and token assets. Table 1 shows the percentage of GARP violations for patients at Central Islip 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. |

From this table appears that a significant fraction of patients (almost 37 percent) violate GARP and hence RCT. Although GARP violators are more numerous among random agents (around 48 percent), their proportion seems comparable to that of human violators. 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 we do not know what happens to random violators with GARP(0.95).

**4.2** Reinhard 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.[19] Sippel ran two experiments with different budget/price situations, involving 12 and 30 subjects, respectively. Table 2 shows the percentage of students and random agents violating GARP and GARP(0.95) in the two experiments.

[18] 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).

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

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% |

The proportion of GARP violators among Bonn students recorded by Sippel is larger than the one recorded by Cox among patients at Central Islip. In Sippel's two experiments 24 out of 42 students, or 57 percent of them, violated GARP. In both experiments violators among random agents outnumber those among students, and in experiment 2 nearly every random agent violates GARP. When GARP(0.95) is considered, the proportion of human violators shrinks noticeably, but so does the proportion of random violators. Therefore, the power of GARP(0.95) is negligible.

**4.3** Aurelio 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). The participants in experiment 3 received and fulfilled a questionnaire at home, and then returned it by post. In all three experiments the subjects were faced with 8 goods and 20 different budget/price situations.[20] Table 3 shows the percentage of human subjects and random agents violating GARP and GARP(0.95) in the three experiments.

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 records a significant proportion of GARP violators among human subjects (be-

[20] 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.

tween 25 and 44 percent), but in his experiments random violators are clearly more numerous.

**4.4** William Harbaugh, Kate Krause and Timothy 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. The choice presented to task participants in this experiment was easier than the task faced by participants in the previous three experimental studies. In 11 different budget/price situations children and undergraduates were presented with bundles containing only 2 commodities, potato chips bags and boxes of fruit juice. Moreover, the subjects did not have to choose one of the infinite bundles satisfying the budget constraint, but were instead presented with finite choice sets including between 3 and 7 bundles that already laid on the budget line. 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. GARP violations by random agents are noticeably more frequent and more severe than violations by human subjects, even for 7-year-old children.

**4.5** In the experiment performed by James Andreoni and John 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 a subject kept for himself and those he passed to another subject were transformed into money, but possibly at different exchange rates. For instance, each token kept became $0.10 while each token

passed became $0.30, or vice versa. Therefore, a subject 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 percentage of GARP violations for students and random agents, and the percentage of GARP(0.95) violations for students only.

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 violations by random agents are noticeably more frequent than violations by human subjects.

**4.6** Philippe Février and Michael 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 shows the percentage of Dijon consumers and random agents violating GARP and GARP(0.95).

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 experiment 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.

## 5. Discussion of experiments on individual demand

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 are an artifact of the experimental procedure itself. For instance, in the case of the last experiment one may argue that the subjects violated GARP because they had to choose among almost identical commodities, and this confounded them or even induced them to choose indeed at random. One may also maintain that the 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. 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 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. Therefore, experimental evidence appears to disprove the standard version of RCT implying no GARP violations. As a consequence, it seems difficult to claim that RCT provides a good explanation of choices recorded in the experiments, even if one adopts a weak as-if interpretation of RCT. At best an almost-RCT, claiming that *most* individuals choose in a *nearly* rational way, could be defended. 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.

Another main outcome of the experiments is that the number of GARP violators varies strongly from experiment to experiment. This suggests that the explanatory power of RCT, even in its more permissive almost-RCT version, is highly context dependent. Such context dependency is a problem since we do not have any meta-theory stating in which contexts RCT works and in which ones it doesn't. The variance in RCT's explanatory power 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. In any case, 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. However even this tendency is shallow. For instance, in Mattei's experiment the subjects are confronted with 8 goods in 20 budget/price situations while in Sippel's they face 8 goods in 10 situations, but

Mattei generally recorded fewer violators than Sippel.

A consequence of RCT's context-dependency is that its image as a general theory of choice becomes less persuasive. A commonly accepted distinction between RCT and bounded-rationality theories is that the former is one single theory with ambition to have universal explanatory power, while the latter constitute a whole range of different decision rules tailor-made to cater particular contexts but not always transferable to other sets of circumstances. The experimental finding that RCT satisfactorily explains certain patterns of choice but not others, and hence appears to have at best local rather than general explanatory power, makes it relatively more similar to the various theories of bounded rationality than is usually assumed.

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. For instance, in the Février-Visser experiment 29% percent of human subjects and 22% of random agents violate GARP. One may hence be tempted to infer that at least in certain situations the random-choice model explains 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. For instance, humans may violate GARP because they stick to previous choices and respond too weakly to price changes, not because they choose randomly. If this is the case, the random-choice model would explain 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 provides a good explanation of human demand behavior, not to check the alternative hypothesis that the random-choice model may provide such an explanation. To investigate this latter issue requires 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. It turns out that this is a tricky problem of nonparametric statistical inference for which no standard test exists. Some of the difficulties draw from the fact that the sample space is different for each budget/price situation the subjects face, and that there is only one observation for each sample space, namely the choice made in the budget/price situation at

issue. Nonparametric tests to check whether a sample comes from a uniform distribution have been developed in medical statistics to assess whether a disease is uniformly distributed among the population of a given region.[21] By adapting these medical tests to the environment studied by consumer theory, it seems in principle possible to construct a test to check whether human subjects choose at random, at least in certain circumstances. One may in fact imagine that in awkward choice situations, e.g. when the available alternatives are numerous or it is difficult to evaluate and compare them, a nonparametric statistical test could validate the random-choice model. More generally, it would be interesting to investigate whether there are specific classes of situations for which the random-choice model explains consumer behavior better than RCT.

Unfortunately, as far as we know the issues concerning the construction and use of a specific test for the random-choice hypothesis have not been investigated in the literature. In our opinion this state of affairs depends on the widespread understanding of the as-if defense of RCT in its weak sense. As far as the only question at stake is weather RCT provides a possible explanation of demand behavior, the tests will aim at falsifying or validating RCT, that is, the tests will be constructed by positing RCT as the only null hypothesis. It is only if one wants to make a strong case for RCT and show that it offers the best explanation of consumer behavior, that tests with different null hypotheses will be investigated.[22]

Finally, the experiments suggest that policy makers should be careful in using RCT as a basis for consumer demand policies, for instance if they endeavor to induce people to drink less alcohol, quit smoking or consume healthier food. In the first place the experimental findings support at best almost-RCT, and it is not even clear in which contexts almost-RCT works and in which ones it does not. Moreover, the experimental support for RCT is weak, in the specific sense that the experiments do not show that RCT provides the best explanation of demand behavior but only a possible explanation. Given these findings, one wonders whether policies not based on RCT may sometimes perform better than RCT-based policies. For instance, Richard Thaler and Cass Sunstein in their recent book *Nudge: Improving decisions about health, wealth, and happiness* (Thaler and Sunstein 2008) show how policies which exploit the cognitive limitations of individuals and the framing effects on human choice can be much more effective in influencing consumer demand behavior than RCT policies based on incentives.

With respect to random behavior, if further research suggested that individuals tend to

[21] On this kind of statistical inference, see for instance Ripley 1988 as well as Bonetti and Pagano 2005.

[22] In the study of behavior under uncertainty and intertemporal decisions, some experimental tests have indeed compared the explanatory power of different theories. See for instance Halevy 2007 or Manzini and Mariotti 2009.

choose at random when they face too many alternatives, a policy that modifies the prices of the alternatives without restricting their number would scarcely affect consumer behavior. If instead random behavior were induced by the difficulty in understanding and evaluating the alternatives at hand, to be effective a policy should be aimed in the first place at making the alternatives more transparent and easier to compare. In general, if the random-choice model proved useful in explaining human choices at least in certain situations, this circumstance would be relevant in designing effective consumer demand policies.

## 6. Conclusions

This paper has discussed RCT and the as-if defense of it through a methodological analysis of a series of recent experiments that employ the random-choice model as an alternative hypothesis in testing whether consumer demand behavior validates RCT.

A first, remarkable result of these experiments is that consumers frequently violate RCT, so that the claim that human choices can be seen as if generated by utility maximization loses ground. At best the laboratory findings support the weaker version of RCT we have called "almost-RCT". Comparison of the different studies also shows that the number of violations varies widely across experiments, which in turn suggests that the explanatory power of RCT depends on context and is therefore less general than previously believed. Given the current lack of a meta-theory identifying the contexts and conditions in which RCT performs poorly as opposed to those in which it does well, we anticipate further research on the cognitive and environmental factors that may play a role in this respect.

In relation to Becker's path-breaking article of 1962, the experiments confirm the limited applicability of his claim that RCT and random choice may lead to equivalent outcomes. Indeed Becker's convergence results hold only if one is interested in aggregate rather than individual demand and concentrates on the expected or average outcome of random choice. In contrast, the experimentalists were interested in individual rather than aggregate consumer demand so that they counted each and every choice; when this is done it is apparent that random agents frequently violate RCT. Generally speaking, RCT and random choice appear today as two divergent models of choice whose behavioral implications differ considerably.

In some experiments the proportion of human subjects and random agents who violate RCT is rather close. Though striking at first glance, this result does not imply that in these experimental situations the random-choice model explains human demand behavior better than RCT, for human subjects and random agents may violate RCT in very different ways.

The literature under review does not explore whether random choice could provide a good explanation of human demand behavior, because it only uses random choice to check the power of the test for RCT: if random agents pass the test, the fact that human subjects also pass it can be attributed to some extrinsic factor rather than human rationality. In order to study the explanatory power of the random-choice model one would need instead to construct a specific test that posits it as the null hypothesis under scrutiny. The methodological and theoretical issues concerning the construction of such a test have not been investigated in the existing literature and are open to future research.

Finally, the experiments show that the prevailing as-if defense of RCT has remained a weak one, and that in this sense there has been little progress since Becker's 1962 article. The standard as-if defense only states that RCT offers a possible as-if explanation of demand behavior, and does not attempt to make the strong case that RCT provides the best explanation among the available ones. The latter point is particularly relevant when policy issues are at stake. If RCT and the random-choice model offer two possible as-if explanations of demand behavior, it is not clear what kind of policy would prove more effective in influencing consumer demand. Because the experimental research reviewed here has never tested the explanatory power of random choice and never compared it to RCT, at present this question remains unsolved.

**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.

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.

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., 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.

Bonetti, M., and Pagano M. (2005), 'The Interpoint Distance Distribution as a Descriptor of Point Patterns, with an Application to Spatial Disease Clustering', *Statistics in Medicine*, 24, 753-773.

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

Caldwell, B.J. (1980), 'A Critique of Friedman's Methodological Instrumentalism', *Southern Economic Journal*, 47, 366-374.

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.

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

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

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.

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. (2009), 'Categorize Then Choose: Boundedly Rational Choice and Welfare', mimeo, Queen Mary University, May 2009.

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.

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

Musgrave, A. (1981), ' "Unreal Assumptions" in Economic Theory: the F-Twist Untwisted', *Kyklos*, 34, 377-387.

Nagel, E. (1963), 'Assumptions in Economic Theory', *American Economic Review* (Papers and Proceedings), 53, 211-219.

Ripley, B.D. (1988), *Statistical Inference for Spatial Processes*, Cambridge: Cambridge University Press.

Samuelson, P. (1938), 'A Note on the Pure Theory of Consumer's Behaviour', *Economica* [NS], 5, 353-4.

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

Tesfatsion, L. (2006), 'Agent-Based Computational Economics: A Constructive Approach to Economic Theory', in L. Tesfatsion and K.L. Judd (eds), *Handbook of Computational Economics* (Vol. 2), Amsterdam: Elsevier, pp. 831-880.

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

Thaler, R.H., and Sunstein, C.R (2008), *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale, Yale University Press.

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