Projektbereich B
Discussion Paper No. B–274

An Experiment on the Pure Theory of Consumer's Behaviour

by

Reinhard Sippel

June 1994, revised version August 1995

Author's address:
Wirtschaftstheoretische Abteilung II
Rheinische Friedrich-Wilhelms-Universität Bonn
Adenauerallee 24–42
D–53113 Bonn
Fon +49–228–73 92 45
Fax +49–228–73 79 40
E-mail: rsippel@track.econ2.uni-bonn.de

An Experiment on the Pure Theory of Consumer's Behaviour

by

Reinhard Sippel*

Abstract

By checking whether consumer demand satisfies the axioms of revealed preference, one can test the empirical validity of the neoclassical theory of consumer behaviour. However, applying the axioms to actual consumer purchase data is difficult, if not impossible, since serious problems of both a methodological and practical nature arise. This paper, after commenting on the few previous empirical studies, reports an experimental approach to revealed preference theory. Data were obtained through a controlled experiment that involved real consumption of the goods chosen. We find that our subjects often violated the axioms.

JEL-Classification: C 91, D 12

Keywords: consumer theory, revealed preference, consistency, experimental economics

1 Introduction

When Paul Samuelson formulated revealed preference theory (cf. Samuelson, 1938) he intended to provide empirically meaningful theorems on consumer behaviour. Somewhat surprisingly, rather than using the axioms as a basis for empirical tests of the validity of the theory, subsequent research has focused on the normative implications of the revealed preference approach eventually showing the equivalence of the strong axiom of revealed preference and the utility maximization hypothesis (Houthakker, 1950). It must be kept in mind, though, that the simple static

*I would like to thank Susie Hellmann, Mike Jerison, Abdolkarim Sadrieh, Reinhard Selten, Jürgen Wiesmann, numerous seminar participants and two anonymous referees for helpful comments on an earlier version of this paper. Part of the revision was undertaken at the European University Institute, Florence, and benefitted from its pleasant research atmosphere.
Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged.

neoclassical theory of consumer demand is assumed to hold only under ideal conditions. Since these conditions are hardly met in economic reality, little room is left for empirical investigation. One of the major problems is that the consistency requirements imposed on the consumer by the revealed preference axioms refer to changes in demand resulting from purely hypothetical changes in prices or income. His actual behaviour over time is not restricted by the axioms unless a change of taste can be ruled out. The few studies that have been performed using real consumption data suffer from this, as well as from other difficulties, and their results are far from clear-cut, so that, up to now, "(t)here is little empirical evidence concerning whether individual demands satisfy the revealed preference axioms" (Härdle *et al.*, 1991, p. 1529).

Experimental economics can provide a means for a more thorough analysis of the empirical side of the revealed preference approach since conditions very close to the theoretical requirements can be created and the data are collected under controlled conditions. However, so far no real laboratory experiment on this topic involving human subjects is found in the literature.[1] This is all the more surprising since, as Pollak (1990, p. 150) points out, "(a)n 'experimental' interpretation of revealed preference—one in which the economist-observer confronts the consumer with a price-expenditure situation and observes the resulting demand behavior—(...) is consistent with the positivist methodology" of Samuelson's seminal contributions to the theory.

The present paper tries to fill this gap. We report on a laboratory experiment where subjects were asked to choose goods they liked to consume facing different budget constraints. Subjects made the different choices almost simultaneously, knowing that they would actually receive one of the chosen bundles of goods. With regard to the asked-for decisions on real consumption items, the experiment is related to a number of studies on the determination of individuals' indifference curves (Thurstone, 1931; MacCrimmon and Toda, 1969; Knetsch, 1992). For example in the MacCrimmon and Toda (1969) experiment subjects were asked to indicate the number of pastries they were willing to give up for a given dollar amount. Forced consumption of the pastries proved effective in inducing the subjects to think carefully before stating their preferences. Actual consumption of the chosen goods is used as the incentive mechanism in the present experiment as well, the commodity space, however, goes far beyond that of the indifference curve studies. The random lottery incentive mechanism applied here to consumer choices on goods with a nontrivial value seems to be well suited for revealed preference tests and a definite improvement on the

[1] Note, however, the experiments by Kagel *et al.* (1975) and Battalio *et al.* (1981) using rats and pigeons as consumers. Changes in the animals' demand for food pellets and liquids after the imposition of considerable price changes are analysed using the revealed preference methodology. While there can be no doubt about the otherwise high degree of experimental control in these studies, preferences, unfortunately, are not controlled, hence again a change in taste over time is possible. These experiments, though interesting in their own way, will not be discussed further in this paper.

earlier studies.

The rest of the paper is organized as follows: In the next section, we give a brief account of revealed preference theory and methodological remarks on how to test it empirically before we shortly review the existing empirical work on this subject. Section 3 presents in detail the consumption experiment and discusses its main results. Section 4 concludes the paper.

2 Revealed Preference Theory and Empirical Applications

The traditional way of modelling the behaviour of a neoclassical consumer is to postulate the existence of well-behaved preferences and a corresponding utility function which is maximized subject to a budget constraint. As long as nothing is known about the actual form of the individual's utility function, this theory might seem void of any empirical content. However, Samuelson (1938) and Houthakker (1950) have shown that demand behaviour must satisfy certain consistency requirements if it is to be exhibited by a utility maximizing household. These requirements, known as the axioms of revealed preference, allow us to test the empirical validity of the theory. Since a presentation of the revealed preference approach to consumer theory can be found in almost any textbook on microeconomic theory, we will state here only the main results relevant for our investigation.

Let the consumer be endowed with a fixed budget B which he can use for the purchase of n different goods $x_1, \ldots, x_n$ with prices $p_1, \ldots, p_n$. If in a particular situation (p^0, B^0) he chooses the bundle of goods $x^0 = (x_1^0, \ldots, x_n^0)$, he reveals his preference for this bundle over all other bundles, which he could have purchased as well but did not. I.e. we say that a chosen bundle x^0 is revealed preferred to some other bundle x^1, written $x^0 \circledS x^1$, if and only if x^1 is not more expensive than x^0 at the prevailing prices p^0; formally:

$$x^0 \circledS x^1 \iff p^0 x^0 \geq p^0 x^1, \quad x^0 \neq x^1$$

If we assume that the consumer always chooses the best bundle he can get[2] (a utility maximizer necessarily does, of course), then, if $x^0 \circledS x^1$, he must never choose x^1 when x^0 is available, i.e.

$$x^0 \circledS x^1 \Rightarrow \neg(x^1 \circledS x^0), \tag{1}$$

[2]Without this additional assumption, a violation of (1) could not be called "inconsistent". As Sen (1993) has convincingly argued, choice behaviour is consistent or not only with respect to some external objective or motivation. If someone has "a desire to violate, deliberately, the standard conditions of consistent behavior to confuse the observer (or to perplex some decision theorists)" (Sen 1993, p. 502), he will be consistent in violating (1). In a well-designed experiment, however, such a motive will never arise.

where $\neg$ denotes negation. This consistency requirement, known as the Weak Axiom of Revealed Preference (WARP) and first formulated by Samuelson (1938), implies homogeneity of demand (only single-valued demand functions are permitted) and downward-sloping compensated demand curves. A consumer always satisfying (1) for arbitrary x^0 and x^1 is not, however, necessarily a utility maximizer. "Revealed preference cycles" such as $x^0 \otimes x^1$, $x^1 \otimes x^2$ and $x^2 \otimes x^0$ might occur. Houthakker (1950) showed that demanding acyclicity of the relation $\otimes$, i.e. ruling out cycles of arbitrary length, is necessary and sufficient for the demand to be generated by a utility maximizer:

$$x^0 \otimes x^1, x^1 \otimes x^2, \ldots, x^{k-1} \otimes x^k, \quad k \geq 1 \Rightarrow \neg(x^k \otimes x^0) \tag{2}$$

(2) is known as the Strong Axiom of Revealed Preference (SARP). A utility maximizing consumer must satisfy (2) for any $x^0, x^1, \ldots \in \mathbb{R}$, i.e. for an infinite number of bundles. Note, however, that in an empirical investigation of consumer behaviour, we only have a finite number of observations. Hence, a confirmation of (2) is impossible,[3] but, as long as the data obtained do not violate (2), we might even construct utility functions for the consumers generating these data.

This "nonparametric approach to demand analysis", to quote the title of Varian's (1982) influential paper, constitutes an alternative to the more traditional approach in consumption analysis, i.e. estimating parameters for specific functional forms. Apart from formulating a somewhat weaker version of SARP, which allows multi-valued demand functions, and showing it to be equivalent to the maximization of a piecewise linear utility function,[4] Varian (1982) presents algorithms that quickly determine whether consumption data are consistent with the axioms of revealed preference.[5]

However, before applying these algorithms to real demand data, some caution is in order. Since standard demand theory is essentially static, it is implicitly assumed that the consumer acts according to an unchanged scale of preferences when he chooses the different bundles of goods. This assumption is of course most questionable if time-series of consumption data covering months or even years are analysed,

---

[3]See Sen (1973).

[4]He calls this axiom the Generalized Axiom of Revealed Preference (GARP):

$$x^0 \otimes x^1, x^1 \otimes x^2, \ldots, x^{k-1} \otimes x^k, \quad k \geq 1 \Rightarrow p^k x^0 \geq p^k x^k$$

Matzkin and Richter (1991) show that, if the data do not violate SARP, then there is a strictly concave utility function that generates the data.

[5]Details of the procedure can be found in Varian's paper and will not be given here. Basically, the procedure is as follows. We first compute a $T \times T$-matrix M, where T is the number of observations per household, with elements $m_{ij} = 1$ if $x^i \otimes x^j$, i.e. if $p^i x^i \geq p^i x^j$, and $m_{ij} = 0$ otherwise. We then systematically look for i and j, $i \neq j$, such that $m_{ij} = m_{ji} = 1$, which is a violation of WARP. After accounting for "indirectly revealed" preferences by computing the transitive closure of $\otimes$, the same procedure detects violations of SARP. If we do not find any such i and j, the observations are consistent with the utility maximization hypothesis.

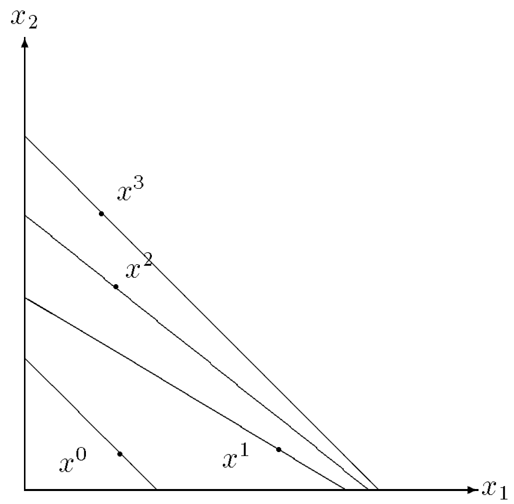

Fig. 1

as is in fact often the case. Any "inconsistency" discovered might then as well be attributed to a change in taste,[6] so that no definite conclusion can be drawn. Even if a change of preferences can be ruled out, as might be the case when observing demand behaviour over a rather short period of time, "inconsistencies" are likely to occur with the existence of durable goods and the possibility of storing or deep-freezing, since purchases of these goods are made only infrequently. Excluding these goods from the analysis amounts to assuming (weak) separability.[7]

While these difficulties might lead to erroneously rejecting the utility maximization hypothesis, the opposite case, i.e. accepting the null hypothesis when it is not true, is also quite likely when real consumption data are analysed. This is because the power of the test might be very low.[8] If the budget planes do not intersect, as in Fig. 1, no inconsistency will ever show up regardless of whether the consumer is maximizing utility, choosing randomly, or is using any other type of decision rule.

For the test to have sufficient power, it is therefore necessary that relative prices vary greatly while income remains more or less fixed. However, this is typically not

[6]Chalfant and Alston (1988) take this approach when analysing the demand shift from red meats to chicken. Using Varian's method, and finding violations of the weak axiom "negligible", they conclude that the demand shift is due to a change in relative prices rather than a change in taste.

[7]Patterson (1991), in a nonparametric analysis of UK data, rejects the weak separability assumption for different definitions of "durable".
Note that separability is also implicitly assumed when attention is restricted to observed purchases, i.e. to a subset of all commodities. See Polemarchakis (1983) for the severe implications of this observability problem.

[8]See Bronars (1987) and Russell (1992) for a discussion of the power of revealed preference tests.

true for real data, since relative prices vary only slightly over time but real income increases from year to year.[9] A major advantage of an experimental analysis is that prices and income are under control of the experimenter, so that the power of the revealed preference test can be made very high.

Despite the serious difficulties mentioned above, there is a limited number of empirical studies on consumer behaviour using the axioms of revealed preference. These will be briefly reviewed in the following paragraph.

Since we are interested in whether consumers actually behave as is postulated in neoclassical theory, a straightforward approach is to analyse real consumption data as are recorded in various expenditure surveys. Households participating in these surveys write down their weekly purchases of consumption goods. Koo (1963) and Koo and Hasenkamp (1972) use Michigan data, Mossin (1972) analyses Danish data, and the recent study by Mattei (1994) is based on a survey of Swiss household budgets. Although these studies differ in various respects such as length of the observation period, level of aggregation of the goods, and the number of households participating, their common feature is that consumers are found to behave inconsistently. However, as pointed out above, from a theoretical point of view there is no reason at all to expect consistency in time-series of consumption data if we exclude the trivial case of non-intersecting budget planes. In addition note that households do not always correctly report their purchases but tend to underreporting, and that information regarding the price vector faced by the consumers is often incomplete. Therefore, errors in the data might lead us to falsely reject the utility maximization hypothesis. Summing up, we can conclude that the analysis of panel data is of a rather limited use when we want to know whether consumers actually behave in a neoclassical manner.

The need for controlled conditions when obtaining data for revealed preference tests was clearly seen by Battalio *et al.* (1973) in their study of consumer demand in a token economy. They observed the demand behaviour of 38 patients in a psychiatric hospital who earned tokens for work performed and spent them for various consumption goods offered in the hospital's token store. Analysis of patients' responses to large, systematic price changes from week to week showed half of them behaving inconsistently. This, however, cannot be regarded as a rejection of the theory. As Battalio *et al.* (1973) point out, errors in the data might be the reason for all but two of the inconsistencies.[10] In addition, the experimental setup of the token economy in a hospital did not allow a sufficient control of the relevant parameters. The token store was not the only source of consumption goods since several patients could leave the hospital and buy goods with real currency while others received foodstuffs from

[9] See Varian (1982), p. 965, and Manser and McDonald (1988), p. 917.

[10] Errors in counting and classifying purchased goods lead to discrepancies between two independently obtained data sets. Since the "true" data are not known, it might well be that some of the subjects found to be consistent did in fact violate the axioms of revealed preference. Battalio *et al.* do not consider this possibility.

visiting relatives. Moreover, for therapeutic reasons, token income or expenditure were not controlled, a fact that could easily lead to non-intersecting expenditure planes. Battalio *et al.* (1973) do not give information on the number of consistent subjects where this "consistency" is the necessary result of non-intersecting budget planes. For all these reasons, this token economy experiment, while being a considerable improvement on the studies mentioned earlier, does not give a satisfactory answer to the question of whether consumers satisfy the axioms of revealed preference, either. In order to avoid the methodological difficulties of the earlier studies we carried out two laboratory experiments, which are documented in the following section.

3 The Experiment

In this section, we first describe the experimental setup and then present and discuss the main results. Two slightly different experiments, which will be referred to as Exp1 and Exp2, respectively, were run. Unless otherwise indicated, the following applies to both experiments.

3.1 The Experimental Setup

To obtain conditions close to the theoretical requirements and thus to provide a serious test of the theory, the consumption experiment was designed as follows.

Since we wanted to analyse individual choice behaviour, only one subject at a time came to the laboratory. There, he or she had to spend a fixed amount of time (one hour) during which nothing else was allowed but the consumption of goods previously purchased from the experimenter. The subject's main task was to choose goods he or she liked to consume in the laboratory as an alternative to just sitting around and doing nothing. The goods' prices, as well as the available budget, were denoted in an artificial unit of account and had no resemblance to actual DM-prices found in the real world. In fact, subjects did not actually pay for the goods but were asked to choose, according to their preferences, a bundle of goods they could afford given their budget and the prices. This, of course, is exactly the problem the neoclassical consumer faces when having to pick his best bundle out of the budget set.

Clearly, we need more than one demand vector per subject to address the question of consistency. One possibility would be to ask the subjects to come to the laboratory in regular intervals, e.g. every day or every week, and each time let them choose from a different budget set (with relative prices having changed). But this procedure would not allow a sufficient control of the subjects' preferences, which, as pointed out above, can be expected to change over time. Therefore, we decided to confront the subjects with 10 different budget situations almost simultaneously.

Subjects were asked to state their demand in each of the 10 situations knowing that every situation had the same (one in ten) chance of being selected, with the subjects actually receiving the goods chosen in this situation. While this random lottery incentive mechanism is a standard procedure in experiments on expected utility theory, it has never before been used in experiments on revealed preference theory. This is quite remarkable since the revealed preference axioms do not relate to actual consumer choices over time but to hypothetical choices "at the same time" (Samuelson, 1938, p. 7). The random lottery mechanism puts the subject in the required position of a consumer who has to make several choices almost simultaneously repeatedly being asked, "What would you choose if prices and income were like this?" The fact that one of the chosen bundles will actually be received gives economic significance to these otherwise purely hypothetical questions and induces the subject to state his true preferences.

Eight different goods were offered, listed in Table 1. They can be divided into two groups, one consisting of various food and drink items, and the other containing goods especially suited for spending the time. The goods were selected in order to allow an as fine division as possible. Another consideration was that every subject should find at least some of the goods desirable. Therefore, the goods offered covered a wide range of tastes.

Table 1:
Goods offered in the experiment

Good	Description	Range[a]
Videoclips	Watching a videotape of rock and pop music video clips	30–60 min.
Computer Game	Playing "Super Blast" (in Exp1) or "Pinball" (in Exp2)	27.5–60 min.
Magazines	Reading a selection of German newspapers and magazines	30–60 min.
Coca-Cola	Cold soft drink	400–2000 grams
Orange juice	Cold drink	750–2000 grams
Coffee	Prepared when demanded	600–2000 grams
Haribo	Popular German brand of candy, licorice etc.	400–2000 grams
Snacks	Pretzels, peanuts etc.	600–2000 grams

[a] In Exp1: Amount of the good available when the entire income was spent on this good only, in the situation where it was most expensive and cheapest, respectively. In Exp2, these amounts differed from subject to subject but were of the same order.

In both experiments, prices were chosen such that there was a large number of budget set intersections giving the revealed preference test a high power. Comparing two budget situations, typically four goods were cheaper in one situation while the other four goods were cheaper in the other situation. However, budget sets were different in Exp1 and Exp2. In Exp1, two of the 10 situations (1 and 7) were virtually identical, the only difference being that prices and income were 1.15 times higher in situation 7 than they were in situation 1. This was intended as a test of homogeneity of degree zero (implied by WARP). In order not to make the identity of situations 1 and 7 too obvious, the nominal budget available was different not only in these but in each of the 10 situations in Exp1, varying from 1500 to 3300. Exp2 did not test homogeneity of demand but implemented a Slutsky-type compensation for the price changes. After having chosen their preferred bundle of goods x^0 in the first situation (denoted S0 and also referred to as the "reference situation" in the instructions), subjects were compensated for the price changes in the remaining 9 situations by corresponding changes in income so that they always could afford x^0 again. Since the budget was hence endogenous in all situations but the first, this raised the possibility of strategic considerations when deciding on x^0. This problem was dealt with in two ways: first, only half of the subjects were actually informed of the compensation procedure (treatment 2b), to the others (treatment 2a) the budgets appeared to be exogenous. Secondly, prices for the two groups of goods were the same in S0 so that there was no "cheapest" good a subject might want to choose in order to expand the other 9 budget sets.[11]

In both experiments, relative prices changed considerably between the situations though to a somewhat lesser degree in Exp2 where the compensation mechanism would have translated extreme price changes into equally extreme changes in the available budget when a subject had chosen a border solution in S0. Even though income was not under complete control *ex ante* in Exp2, the provisions taken were sufficient to make the available budgets quite tight, forcing the subjects in both experiments to carefully consider their choices.

The experiments were divided into two parts. In the first part, after receiving verbal instructions,[12] the subject had to fill out 10 order sheets, one for each budget situation, with prices and income varying considerably between the situations. The subject had to state his demand for each of the 8 goods offered, the only restriction being the budget restriction. Since this task required a considerable amount of calculation (multiplying prices and demand for each good and adding up to check whether the budget was exhausted), a personal computer was used for this purpose.[13] The subject simply had to enter the amounts desired, and the software informed on the cost of the bundle and checked whether it was inside the budget set. If not, a warning message showed up on the screen. In order to help the subject not to waste anything of his budget, additional information appeared on the screen showing, for each of the 8 goods, the amount that could be ordered to exactly exhaust the budget,

[11] Specifically, the budget in S0 allowed a maximum consumption of 50 minutes of any of the time-consuming goods or 1000 grams of any of the food and drink items.

[12] A translation of the instructions which were given in German can be found in the appendix.

[13] See Fig. 2 in the appendix for a sample screen shot. While in Exp1 subjects had to transfer their choices from the screen to sheets of paper, in Exp2 the order sheet for each situation was printed out immediately after the decision.

given the demand stated for the other 7 goods.[14] Since this information was meant only for fine-tuning the demand it only appeared when the demand vector stated so far was sufficiently close to the budget hyperplane, namely when the corresponding expenditure was within a range of 95–105% of the budget. This information clearly helped the subjects to avoid an annoying trial-and-error mechanism.

No time limit was imposed on the first part of the experiment. Subjects could spend as much time as they liked on their decisions and were free to compare, reconsider, and correct choices already made. This applies to all situations except S0 in Exp2 which could not be corrected once it had been printed out.[15] When they felt that the 10 order sheets filled out represented their actual preferences, they asked the experimenter to start the second part. One of the 10 situations was then drawn at random using a bingo cage. The second part always lasted exactly 60 minutes independent on how much time the subjects had spent on the first part. They received the goods chosen in the situation drawn and could consume them in any order they liked (simultaneous consumption of, say, videos and Coke was possible, too). After the 60 minutes the experiment was over, and the subjects received a fixed show-up fee of DM 25.

3.2 Experimental results

Both experiments took place at Bonn University's Laboratory for Experimental Economics. Exp1 was conducted in June 1993 and involved 12 subjects, while 30 subjects participated in Exp2 during February 1995. Subjects were students mostly of law or economics asking for participation after a public announcement of the experiment. None of them participated in both experiments.

In Exp1, 11 out of the 12 subjects violated the weak (and thus also the strong) axiom of revealed preference. I.e. only one participant can be viewed as a utility maximizer. All of the inconsistent subjects chose different bundles of goods in the identical situations 1 and 7, i.e. they showed an inhomogeneous demand and, hence, violated WARP. However, it might be argued that these subjects were in fact indifferent between the two bundles so that choosing x^1 in situation 1 and $x^7 \neq x^1$ in situation 7 should not be called inconsistent. This is the argument underlying Varian's GARP (see above) which does not ask for a unique demand vector in every situation. Since non-homogeneity of demand was the only "inconsistency" for 6 subjects, it leaves us with 5 subjects (42%) who violated revealed preference even in its weak form of GARP.[16] The percentage of inconsistent subjects was even higher

[14] Subjects were not forced to spend all their budget but were free to spend less. However, since any amount left over was not refunded, subjects clearly had an incentive to exhaust their budget (assuming nonsatiation, of course).

[15] The special nature of the "reference situation" S0 was emphasized in the instructions to subjects in both treatments 2a and 2b.

[16] Whether the subjects really were indifferent between the different bundles x^1 and x^7 is an

in Exp2: 22 (73%) violate WARP and SARP,[17] and 19 (63%) violated GARP.[18]

While these results are not too favourable to the neoclassical theory of consumer behaviour, they deserve some closer inspection. If we count for each subject the number of violations of the axioms of revealed preference we find these numbers to be remarkably low given the high power of the test. Since we performed $\binom{10}{2} = 45$ pairwise comparisons for each subject, and in each comparison budget sets intersected, 45 was the maximum number of SARP violations possible for a subject behaving extremely inconsistent. Yet, the median number of SARP violations in Exp2 was only 2, and for GARP the median was only 1.[19] The fact that nearly all subjects remained well below the maximum number clearly indicates that they did not choose randomly. A closer look at the actual demand data corroborates this view.[20] Every subject showed a marked preference for some of the goods while other goods were not chosen at all, even at low prices.[21] Several subjects substituted cheaper goods for their more expensive counterparts, e.g. Coke for orange juice, sometimes to the extent that they always switched from one to the other, depending upon which was cheaper in the particular situation. There can be no doubt that the subjects tried to select a combination of goods that came as close as possible to what they really liked to consume given the respective budget constraint. They spent a considerable amount of time on their decisions (typically 30–40 minutes) and repeatedly corrected entries on some of their order sheets when they reconsidered previous choices. Of the 11 subjects who changed one or more of their entries in Exp2,[22] this never led to an increase but in most cases to a decrease in the number of violations. Subject 27 actually became consistent with GARP while his original choices would have violated GARP 3 times. In the majority of cases, the budget was spent entirely or at least up to a negligible rest the spending of which would have delivered only a fraction of a gram (or minute) of additional consumption. Repeatedly subjects expressed their discomfort with situations where preferred goods were quite expensive and hoped that another, preferred situation would be drawn. From all this, we can

open question. They were not asked explicitly about their preferences regarding the two bundles since this might have revealed to them the identity of the two situations, a fact which was to be kept unknown until all sessions of Exp1 were completed.

[17] 14 of these subjects actually exhibited an upward-sloping compensated demand between S0 and one or more of the other situations.

[18] There was no significant difference in behaviour between subjects in treatments 2a and 2b. 8 of the 15 participants informed of the compensation violated GARP, as did 11 of the 15 others who were not informed.

[19] In Exp1, the medians were even lower and both smaller than 1. Note, however, that in Exp2 there were actually 2 subjects who showed 45 violations of SARP, i.e. for these subjects we have that for any $x^i \neq x^j, i, j = 0 \ldots 9$, both x^i is (indirectly) revealed preferred to x^j and x^j is (indirectly) revealed preferred to x^i, so that no preference pattern at all emerges from their behaviour.

[20] These data are not reproduced here but are available from the author upon request, as is the software used in the experiment.

[21] Some subjects explicitly stated to the experimenter that they disliked certain goods, such as coffee or playing a computer game; consequently, they never chose these goods.

[22] Unfortunately, in Exp1 the original choices were not recorded.

safely conclude that subjects were highly motivated when making their decisions. Still, a majority of them violated the axioms of revealed preference.

3.3 An evaluation of the results

It is useful to compare our results with previously published findings on the consistency of consumer demand. This, however, is not an easy task since the earlier studies differ in various respects from our approach. As pointed out above, the analysis of panel data is highly inadequate to check whether consumers actually are utility maximizers. In the light of our findings, however, the negative results found in these studies might not be solely ascribed to methodological shortcomings or data inaccuracies. The same is true for Battalio *et al.*'s (1973) token economy study. Since the accuracy of our data cannot be questioned, and yet we also find a considerable number of violations of the revealed preference axioms, we might conjecture that at least some of the inconsistencies detected by Battalio *et al.* are true violations as well.

Given that subjects do violate the axioms of revealed preference, the question arises how serious these deviations from optimizing behaviour really are. It might be the case that the difference in "utility" or satisfaction between a chosen bundle and another one revealed preferred to it is in fact hardly noticeable for the subject. We might then regard the inconsistency of not choosing the seemingly preferred bundle as being of minor importance. There is, however, no way to tell whether a deviation is serious or not since we do not know the "true" preference structure. In fact, without this knowledge we cannot even tell what is a deviation and what is behaviour in accordance with these preferences.

These problems notwithstanding, several authors have suggested measures for rating revealed preference violations. A particularly simple measure is the Afriat efficiency index (see Afriat, 1973; Varian, 1993). Acting inconsistently, i.e. choosing some consumption bundle x^1 in place of another bundle x^0 which has been revealed preferred to it and is affordable as well, obviously amounts to wasting income. This inefficiency is the greater the smaller is the ratio p^1x^0/p^1x^1, where $x^0 \ogreaterthan x^1$. If this ratio, called the Afriat efficiency index and denoted by e, is close to 1, the waste of income might be insignificant for the consumer. This argument suggests reformulating the revealed preference relation $\ogreaterthan$ in a weaker form, so that $x^0 \ogreaterthan x^1$ if and only if x^1 is "distinctly" cheaper than the chosen bundle x^0:

$$x^0 \ogreaterthan x^1 \iff e \cdot p^0x^0 \geq p^0x^1, \quad x^0 \neq x^1$$

If we set $e = .9$, for example, then x^0 has to be more than 11% more expensive than x^1 before we conclude that $x^0 \ogreaterthan x^1$. Table 2 shows that the number of inconsistent subjects and correspondingly the total number of inconsistencies is substantially reduced if the relation $\ogreaterthan$ is weakened in the above way. With an efficiency of 95%, less than 10% of the subjects remain inconsistent.

Table 2:
Number of inconsistent subjects
and total number of inconsistencies
depending upon the Afriat efficiency index e

	number of inconsistent subjects		total number of inconsistencies	
e	Exp1	Exp2	Exp1	Exp2
1.0	11	22	40	165
.999	5	13	26	124
.99	3	8	22	109
.95	1	3	12	34
.90	1	1	12	3

Note that proceeding along these lines severely reduces the power of the test, since with a lower e, fewer preferences are revealed, and there are more pairs of bundles which cannot be compared using the relation $\otimes$. The procedure might be appropriate if there is "noise" in the data, as in panel data studies, but is less convincing in the present context.[23]

Perhaps more satisfying is an approach that takes into account the fact that the revealed preference axioms do not contain any stochastic component. The consumer is assumed not to make the slightest error when choosing his goods. But this might be too heroic an assumption since experimental subjects are known to make errors from time to time. In our case, they might unintentionally enter incorrect amounts on the order sheets, and these errors might translate into an overall inconsistent behaviour. Therefore we should ask how sensitive our results are to slight perturbations of the actual demand vectors. We did this in the following way: In each situation, the subject's demand for one good was increased by $\varepsilon\,\%$, where ε ranged from 2 to 100, and the demand for another good was decreased correspondingly such that the overall expenditure remained unchanged. We did this for each situation separately, and for all goods with a positive demand.[24]

We find that the results are hardly affected by the perturbations. Subjects found to frequently violate the axioms continue to do so with the perturbed data, while

[23]Jerison and Jerison (1993) point out that the inefficiency as measured by the Afriat index will be the greater the higher is the degree of variation in prices. Since in our case relative prices varied considerably the inefficiencies detected, already of a rather small scale, will be further reduced if divided by the length of the vector of price changes. But the exact adjustment, proposed by Jerison and Jerison (1993), is not applicable in our case since it requires knowledge of the Slutsky matrix and thus of the whole demand function.

[24]We did not use Manser and McDonald's (1988) and Patterson's (1991) sensitivity analysis with random shocks applied to the whole data set at once since this is not the kind of possible error we have in mind. After all, our data are obtained without "noise".

subjects found to be consistent rarely become inconsistent even when the demand for one good is doubled ($\varepsilon = 100$) and that for another decreased accordingly. Only for subjects who violated GARP once or twice the perturbation has some significance since for them we can find perturbed demand vectors making them consistent.

Lacking a convincing measure of how far from the optimum inconsistent subjects actually are, we cannot reach a definite conclusion about the seriousness of the detected violations of the revealed preference axioms. But given that even in this relatively simple setting the majority of subjects do not appear to be optimizing we might regard this as a confirmation of the view that human decision making is only boundedly rational (cf. e.g. Selten, 1990).

4 Conclusion

The revealed preference approach to consumer theory, originating with Samuelson's (1938) seminal paper, is based on the assumption that "the individual guinea-pig, by his market behaviour, reveals his preference pattern—if there is such a consistent pattern" (Samuelson, 1948, p. 243). Thus, it is formulated in a positivist vein and is, in principle, subject to empirical refutation. However, only in an experimental environment where tight control of the relevant parameters, in particular the consumer's preferences, can be maintained, does such an empirical test make sense. The experiment reported was designed to reflect the fact that revealed preference theory is concerned with hypothetical choices rather than actual choices over time. In contrast to earlier experimental studies, the possibility that the different choices are made under different preference patterns can almost be ruled out. We find a considerable number of violations of the revealed preference axioms, which contradicts the neoclassical theory of the consumer maximizing utility subject to a given budget constraint. We should therefore pay closer attention to the limits of this theory as a description of how people actually behave, i.e. as a positive theory of consumer behaviour.[25]

Another conclusion regards the usefulness of the experimental setup for further research on properties of consumer demand. It turned out that subjects were highly motivated when making their decisions. Although their choices were not related to monetary payoffs but rather to the possibility of consuming a specific bundle of goods in the laboratory, the latter proved to be of a nontrivial value, inducing the subjects to weigh their alternatives carefully. Thus, an experimental analysis of consumer demand along the lines laid out here can be a fruitful approach especially in cases where real-economy consumption data are either not reliable enough or simply not available.

[25]See Thaler (1980) for further examples of divergence between normative theory and actual behaviour of consumers.

Appendix

Instructions for participants[26]

Welcome to our consumption experiment. If you complete the experiment until the end, you will receive a fixed amount of DM 25,– as a compensation for the time and effort spent here. In addition, you can consume goods offered here in the laboratory.

It is your task first to purchase goods with an amount determined by us which you can then consume here in the laboratory during 60 minutes. In these 60 minutes you can only consume goods previously purchased here. Consumption of food you might carry with you, smoking, or in general preoccupation with anything you did not pay for, is not allowed. The goods purchased cannot be taken out of this room but can only be consumed here during the 60 minutes. When buying the goods, it is therefore wise to carefully think about what you really want to consume—and what not.

Thus, the experiment is divided into two parts. In the first part you order the goods that you would like to consume in the second part. The first part ends when you have decided which goods you would like to consume and you have filled out order sheets accordingly. There is no time limit on the first part. The second part, however, will last exactly 60 minutes, regardless of the duration of the first part.

The following 8 goods can be purchased:

[The 8 goods were shown and explained in detail.]

How do you buy these goods? To buy goods, you have to fill out order sheets like this one:

[The order sheet was explained. The following part of the instructions was slightly different between Exp1 and Exp2.
Instructions in Exp1:]

You have to fill out 10 order sheets altogether. The goods' prices are different on each order sheet, sometimes they are cheaper, sometimes they are more expensive. The amount you have available for spending on the goods is different in each case, too. Thus, it might happen that in one situation buying 30 minutes of the computer game exhausts all your budget, leaving nothing for the other goods, while in another situation the computer game is so cheap that you can afford 60 minutes of it or have something to drink or eat with it. The way you spend the available budget on the different goods, i.e., how much you would like to consume of each of the goods, is up to you. You can spend the whole amount on a single good, you can buy something of every good, or you can make a selection of several out of the 8 goods. It's only your choice. The only restriction you face is that you cannot spend more than the available amount. Of course, you do not have to spend the whole amount. If you like, you can spend less. But note that any amount left will not be refunded and that the only thing you can do during the 60 minutes is consume goods you have purchased before.

Now, which one of the 10 order sheets will be valid in the second part of the experiment? This is not known in advance. One of the 10 order sheets filled out will be drawn at

[26] In the experiment, the instructions were given orally in German.

random using this bingo cage. One of the numbers from 1 to 10 will be drawn, and the order sheet carrying this number will be used. The goods you ordered on this sheet you will actually receive for consumption. I.e., only one of the 10 sheets will eventually be used. In filling out each order sheet, you should bear in mind, however, that this particular sheet might be drawn. Thus you should fill out the forms carefully. Always look at the goods' prices and the amount you have available and consider the resulting alternatives before you fill out the form. You are free to erase and correct entries as often as you like if you change your mind and want to order something different. Only when you are satisfied with your choices on each of the 10 order sheets the second part of the experiment begins by drawing a random number.

To minimize your calculation effort, the order sheets are reproduced here on the computer screen. Each order sheet on paper can be found in the computer as well.

[The usage of the software was explained.]

Do you have any questions? If not, we can start with the experiment. I will be around for the whole time happy to answer any questions which might arise.

[End of instructions in Exp1.
Instructions in Exp2, continued:]

You have to fill out 10 order sheets altogether, corresponding to 10 different price situations. The sheets are denoted S0 through S9, and you will fill them out one after the other on the screen. Look at S0, which is called the reference situation, and note that in S0 the three goods suited for spending your time all cost the same, while the other five goods also have a unique price, so that, within both categories, you can buy the same amount of the different goods. This, however, will not be the case in the other 9 situations S1 through S9. The goods' prices will vary between the situations and relative to each other, sometimes the goods are cheaper, sometimes they are more expensive than in the reference situation S0. The amount you have available for spending on the goods is different in each situation, too. Thus, it might happen, for example, that in one situation buying 30 minutes of the computer game exhausts all your budget, leaving nothing for the other goods, while in another situation the computer game is so cheap that you can afford 60 minutes of it and even have something to drink or eat with it. [Only in treatment 2b: In fact, the available amount in situations S1 through S9 is calculated such that you can again afford exactly the combination of goods you bought in the reference situation S0, no matter how different the prices of the goods in the particular situation are. I.e., if you like to, you can buy the the same combination of goods in each of the 10 situations. Note that this might not be a good idea, since in S1 through S9 you might be able to afford combinations preferred to the one you chose in S0.] The way you spend the available budget on the different goods, i.e., how much you would like to consume of each of the goods, is up to you. You can spend the whole amount on a single good, you can buy something of every good, or you can make a selection of several out of the 8 goods. It's only your choice. The only restriction you face is that you cannot spend more than the available amount. Of course, you do not have to spend the whole amount. If you like, you can spend less. But note that any amount left will not be refunded and that the only thing you can do during the 60 minutes is consume goods you have purchased before.

Now, which one of the 10 order sheets will be valid in the second part of the experiment?

This is not known in advance. One of the 10 situations will be drawn at random using this bingo cage. One of the numbers from 1 to 10 will be drawn, and the order sheet corresponding to this number will be used (1=S1, 2=S2, ..., 10=S0). The goods you ordered on this sheet you will actually receive for consumption. I.e., only one of the 10 sheets will eventually be used. In filling out each order sheet, you should bear in mind, however, that this particular sheet might be drawn. Thus you should fill out the forms carefully. Always look at the goods' prices and the amount you have available and consider the resulting alternatives before you fill out the form. You are free to erase and correct entries as often as you like if you change your mind and want to order something different. This is true for all situations except the reference situation S0. Your choice of goods in the reference situation S0 is a once and for all decision, so you should take special care in this situation. When you are satisfied with your choices on each of the 10 order sheets the second part of the experiment begins by drawing a random number.

Do you have any questions? If not, we can start with the experiment. I will be around for the whole time happy to answer any questions which might arise.

Fig. 2:
Sample screen shot

S8 Güter	Einheit	Preis pro Einheit	maximale Bestellmenge	bestellte Menge	Preis der bestellten Menge	ausschöpfende Menge
Videos	min	40.000	57.500	13.000	520.000	12.000
Computerspiel	min	30.000	76.667	34.000	1020.000	32.667
Magazine	min	50.000	46.000	0.000	0.000	
Coca-Cola	g	1.750	1314.286	200.000	350.000	177.143
Orangensaft	g	2.500	920.000	100.000	250.000	84.000
Kaffee	g	2.000	1150.000	100.000	200.000	80.000
Haribo	g	2.500	920.000	0.000	0.000	
Knabbern	g	2.250	1022.222	0.000	0.000	
				Summe:	2340.000	zu viel!
				max.	2300.000	

References

Afriat, S. N. (1973), On a System of Inequalities in Demand Analysis: An Extension of the Classical Method, *International Economic Review* **14**, 460–472.

Battalio, R. C., Kagel, J. H., Rachlin, H. and L. Green (1981), Commodity-Choice Behavior with Pigeons as Subjects, *Journal of Political Economy* **89**, 67–91.

______, ______, Winkler, R. C., Fisher, E. B., Basmann, R. L. and Krasner, L. (1973), A Test of Consumer Demand Theory Using Observations of Individual Consumer Purchases, *Western Economic Journal* **11**, 411–428.

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

Chalfant, J. A. and Alston, J. M. (1988), Accounting for Changes in Tastes, *Journal of Political Economy* **96**, 391–410.

Härdle, W., Hildenbrand, W. and Jerison, M. (1991), Empirical Evidence on the Law of Demand, *Econometrica* **59**, 1525–1549.

Houthakker, H. S. (1950), Revealed Preference and the Utility Function, *Economica* **17**, 159–174.

Jerison, D. and Jerison, M. (1993), Approximately Rational Consumer Demand, *Economic Theory* **3**, 217–241.

Kagel, J. H., Battalio, R. C., Rachlin, H., Green, L., Basmann, R. L. and Klemm, W. R. (1975), Experimental Studies of Consumer Demand Using Laboratory Animals, *Economic Inquiry* **13**, 22–38.

Knetsch, J. L. (1992), Preferences and Nonreversibility of Indifference Curves, *Journal of Economic Behavior and Organization* **17**, 131–139.

Koo, A. Y. C. (1963), An Empirical Test of Revealed Preference Theory, *Econometrica* **31**, 646–664.

Koo, A. Y. C. and Hasenkamp, G. (1972), Structure of Revealed Preference: Some Preliminary Evidence, *Journal of Political Economy* **80**, 724–744.

MacCrimmon, K. R. and Toda, M. (1969), The Experimental Determination of Indifference Curves, *Review of Economic Studies* **36**, 433–451.

Manser, M. E. and McDonald, R. J. (1988), An Analysis of Substitution Bias in Measuring Inflation, 1959–85, *Econometrica* **56**, 909–930.

Mattei, A. (1994), La cohérence des choix des consommateurs suisses, *Revue suisse d'Economie politique et de Statistique***130**, 3–20. English version: *A Non-Parametric Test of Consumer Behavior Using Panel Data.* Discussion Paper no. 9112, Département d' Econométrie et d' Economie Politique, Université de Lausanne.

Matzkin, R. L. and Richter, M. K. (1991), Testing Strictly Concave Rationality, *Journal of Economic Theory* **53**, 287–303.

Mossin, A. (1972), A Mean Demand Function and Individual Demand Functions Confronted with the Weak and the Strong Axioms of Revealed Preference: An Empirical Test, *Econometrica* **40**, 177–192.

Patterson, K. D. (1991), A Non-Parametric Analysis of Personal Sector Decisions on Consumption, Liquid Assets and Leisure, *The Economic Journal* **101**, 1103–1116.

Polemarchakis, H. M. (1983), Expectations, Demand and Observability, *Econometrica* **51**, 565–574.

Pollak, R. A. (1990), Distinguished Fellow: Houthakker's Contributions to Economics, *Journal of Economic Perspectives* **4**, 2, 141–156.

Russell, R. R. (1992), Remarks on the Power of Non-Parametric Tests of Consumer-Theory Hypotheses, in *Aggregation, Consumption and Trade* (ed. L. Phlips and L. D. Taylor), Dordrecht: Kluwer Academic Publishers, 121–135.

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

_____ **(1948),** Consumption Theory in Terms of Revealed Preference, *Economica* **15**, 243–253.

Selten, R. (1990), Bounded Rationality, *Journal of Institutional and Theoretical Economics* **146**, 649–658.

Sen, A. K. (1973), Behaviour and the Concept of Preference, *Economica* **40**, 241–259.

_____ **(1993),** Internal Consistency of Choice, *Econometrica* **61**, 495–521.

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

Thurstone, L. L. (1931), The Indifference Function, *Journal of Social Psychology* **2**, 139–167.

Varian, H. R. (1982), The Nonparametric Approach to Demand Analysis, *Econometrica* **50**, 945–973.

_____ **(1993),** Goodness-of-Fit for Revealed Preference Tests, mimeo, University of Michigan.