A study of consumer behavior using laboratory data*

Philippe Février[†]and Michael Visser[‡]

March 1999
Revision, January 2000

Abstract

This paper reports the results of an experiment on individual consumer behavior. The experiment was designed to address the following questions. Do participants behave as utility-maximizing agents? Are there variables (socio-economic characteristics, experimental conditions) that have an effect on the probability of being non-rational? And finally, to what extent does the presence of non-rational individuals affect the estimation results of commonly used demand equations? Revealed preference tests indicate that 29% of the individuals do not behave as utility-maximizing agents. Gender and the times spent on performing experimental tasks affect the likelihood of being non-rational, but the level of remuneration does not. **remaining results here**

*The experiment on which this paper is based was designed and carried out with P. Combris, S. Issanchou and C. Lange. We are very grateful for their help. We also thank seminar participants at CREST and the University of Bourgogne for their comments and suggestions.

[†]Centre de Recherche en Economie et Statistique (CREST), 28 rue des Saints Pères, 75007 Paris, France. Email: fevrier@ensae.fr.

[‡]CREST. Email: visser@ensae.fr.

1. Introduction

The neoclassical theory postulates that consumers have preferences on a set of consumption bundles. It is assumed that these preferences satisfy certain properties—preferences are assumed to be complete, reflexive, transitive, etc...—which imply the existence of a well defined utility function. The basic hypothesis is then that consumers always choose the most preferred bundle from the set of feasible alternatives, i.e. they are assumed to maximize their utility function under a budget constraint. As is well known, the utility-maximizing hypothesis leads to the main implication of the neoclassical theory of consumer behavior, namely that the matrix of substitution terms is a negative semidefinite matrix. In empirical studies of individual consumer behavior, these Slutsky restrictions are imposed on the demand equations, but they are, however, generally rejected by the data (see for example Blundell, Pashardes and Weber (1993)).

The fact that the Slutsky conditions are generally rejected raises a number of questions. First of all, do consumers actually behave as utility-maximizing agents? If the answer to this question is "partly yes, and partly no", who are then the utility-maximizers and who are those that adopt other behavioral strategies? Is it possible to say precisely what kind of alternative behavioral strategy explains the choices of non-rational individuals? And finally, to what extent does the presence of non-rational individuals affect the estimation results of demand equations and the tests of the Slutsky restrictions?

The purpose of this paper is to address some of these questions. We use data on individual purchase decisions that were generated in a laboratory. A random sample of 120 individuals from the French city of Dijon participated in our experiment. In the first part of the experiment the participants were required to evaluate 6 real food products. In the second part they were given the possibility to buy the products, under 5 different price/budget configurations. We test whether the individuals in our data set are utility-maximizers by applying revealed preference tests (the revealed preference approach to testing for consistency with utility-maximization is mainly due to Afriat (1967, 1973) and Varian (1982)). Revealed preference tests consists in checking if the behavior of an individual verifies certain conditions known as revealed preference conditions. If the choices made by an individual verify the conditions, there exists an utility function that is compatible with her/his behavior; if, on the contrary, the individual's observed behavior is not coherent, there does not exist an utility function that could have generated the individual's choices. In the first case the individual is said to be rational, and in the second case irrational. After establishing who are the rationals and who the irrationals in our sample, we proceed by investigating if there are variables that affect the likelihood of being non-rational. In our analysis of the determinants of non-rationality, we exploit the fact that several socio-economic characteristics of the experimental subjects were recorded; we also study whether experimental conditions, such as the level of remuneration and the individual's times spent on performing the experimental tasks, have an effect on the probability of being non-rational. The final part of the paper investigates whether the presence of irrationals matters on aggregate. Since consumption economics, even at the micro-level, is less concerned with the activities of a particular individual than with the average behavior of groups of economic agents, it is of interest to learn if the aggregate behavior of rationals and irrationals differs, or if, on the contrary, the "errors" made by the irrationals at the individual level cancel out on average, rendering in a sense their aggregate purchase behavior indistinguishable from those of the rationals. For this purpose, we estimate two systems of demand equations that are frequently used in the empirical consumption literature: the almost ideal demand system of Deaton and Muellbauer (1980), and the Translog demand system of Christensen, Jorgenson and Lau (1975). Since we know who are the rationals and the irrationals in the sample, it is possible to identify the respective effects of these two groups of individuals on tests of the Slutsky restrictions and on estimates of the demand equations. We test if the restrictions implied by the economic theory are accepted once the non-rational subjects are excluded from the estimation sample, and we test whether the estimated coefficients of the demand equations are identical for the rationals and irrationals.

To deal with the issues raised in this paper, our experimental data have a number of advantages over field data. In field data, like individual panel data for instance, it is typically the case that successive budget levels for an individual increase over time but prices remain relatively stable, implying that the power of revealed preference tests is low. In contrast, we had full control over the prices and budget levels with which we confronted the participants of our experiment. As will be explained in the next section, the prices and budget levels were chosen so that the power of the tests against certain behavioral alternatives is very high given the relatively small number of different price/budget configurations. Another major advantage of our laboratory data is that the different product choices were made during a short span of time (one hour), so that unlike panel data it can reasonably be assumed that the underlying preference structure was the same for all the observed choices. Still another advantage is that all goods were observed in our laboratory environment, whereas in field data one typically observes only a subset of the commodities chosen by consumers. This means that, unlike studies based on field data, it is not necessary to make additional assumptions about the utility function in order to apply revealed preference tests.[1]

The paper is closely related to Cox (1997) and especially Sippel (1997). Both papers reported revealed preference tests using experimental data.[2] We will describe in what respects our experimental design differs from Sippel's—Cox's study is based on data from the token economy experiment in a psychiatric hospital (see Battalio *et al.* (1973)), which is rather a field experiment, and is therefore not directly comparable with our laboratory experiment—and we will compare in detail the revealed preference test results obtained by Cox and Sippel with our's. Compared to Cox and Sippel, the analysis in this paper is more extended in the sense that we not only test for utility maximization but, as mentioned above, we also try to identify the causes of non-rationality, and look at the consequences of non-rationality on the estimation of demand equations and tests for the Slutsky conditions.

Section 2 describes the design of the experiment and contains a descriptive analysis of the data. Section 3 presents the theory, section 4 the results and section 5 concludes.

2. Data

Two slightly different experiments were carried out. In the next subsection we describe the design of the experiments, and subsection 2.2 contains an informal descriptive analysis of the data.

2.1. Design of the experiments

The experiments will be referred to as Experiment 1 and Experiment 2. Both were held in Dijon, which is a medium sized city of around 300,000 inhabitants situated in eastern France. A different group of individuals was recruited for each experiment. The individuals for Experiment 1 were recruited in May 1997, those for Experiment 2 in September 1997. Apart from this difference in the recruitment period, the selection of the two groups took place in an identical way. Unlike most

[1] Varian (1988) shows that if the set of goods from which a consumer chooses is larger than the set of observed goods, then any sequence of choices is compatible with utility maximization, i.e. the revealed preference theory puts no restrictions whatsoever on the observed data. Only if additional assumptions are made about the utility function, such as (weak) seperability in the observed goods (see Varian (1983)), does the utility hypothesis have implications for subsets of goods.

[2] The book by Kagel, Battalio, Rachlin and Green (1995) summarizes a number of studies based on experimental data from animals. In these studies only the Weak Axiom of Revealed Preference (WARP) test was performed. Coherency with WARP is a necessary but not sufficient condition for utility maximization.

experimental studies, the participants for our experiments were not recruited from the pool of faculty colleagues or university students, but were randomly selected from the population of Dijon. Selected individuals were contacted by telephone and were informed about the goal of the study and the remuneration. The information given on the telephone was deliberately vague: the study was described as "a study on consumer's purchase behavior of orange juice" and the remuneration was formulated as "payment in kind," without mention of the exact amount of juice or the corresponding monetary value. Several questions were then asked about their consumption behavior of orange juice. Individuals who satisfied a number of conditions[3] were invited to participate in the experiment. For both experiments approximately 400 individuals were contacted; roughly 80 satisfied the conditions and agreed to participate. At both experiments exactly 60 individuals actually showed up. Experiment 1 was conducted in June 1997, and Experiment 2 in October 1997.

First we describe Experiment 1. It took place at the tasting laboratory of the *Institut National de la Recherche Agronomique* in Dijon. The 60 individuals participated in an experimental session that lasted about one hour.[4] The subjects performed their tasks in individual booths—resembling the type of booths that can be seen in language classes—each one equipped with a chair and a desk, a television screen and a computer mouse. While seated in their booths, the subjects could not see each other. Subjects could thus not communicate with each other unless they stood up or spoke aloud; we asked them not to do this during the session.

In the first part of the session the participants evaluated six orange juices. The six juices were chosen—out of a larger group of orange juices that were available to consumers in a particular supermarket of Dijon—on the basis of four criteria: they had to be different from a sensory point of view—this was determined in a preliminary study by the sensory experts—they had to be different in nature, their type of packaging had to be different, and they had to be different in price.

The subjects evaluated the six juices[5] one after the other. The order in which

[3] The individuals had to be regular consumers of orange juice, they had to take part in the shopping activities of the household, and it was required that they had never before participated in a study on sensory evaluation of food products.

[4] Because of space limitations at the laboratory, the 60 individuals could not come all at once, and we were thus obliged to organize several sessions. The average number of individuals per session was about 6.

[5] Just before the six products were evaluated, the subjects examined an additional (and identical for all) juice to familiarize them with the procedures in this part of the session. The purpose of this dummy juice was also to counter the so-called first-order carry-over effect in sensory studies (see MacFie et al. (1989)).

the juices were evaluated was randomly determined for each subject, i.e. the sequence of juices was not necessarily the same for all individuals. Each juice was evaluated in two phases: first some of the characteristics of the product were studied, then it was tasted. Subjects were informed about some of the product characteristics via an image appearing on the television screen. Depending on the actual type of packaging of the juice, the image represented either a cardboard pack or a bottle—three of the orange juices chosen for this study were sold in cardboard packs, and three in bottles. The images of the three cardboard packs, respectively the three bottles, were identical and designed by us. They were constructed from photos of real commercial juice packages, which we scanned using a photographic retouching software package. The information written by us on each image was true information, i.e. we did not provide the participants with false information. The images informed the participants about the nature of the juice—there were two pure orange juices, two concentrated orange juices, and two nectars—the country from which the juice or the oranges originated, and in the case of nectar the percentage of pure orange. The image corresponding to product 3 indicated a guaranteed content of vitamin C. Table 1 summarizes the information given to the participants. Note that the subjects were not informed about the market price nor of the brand of the juice, but otherwise the images revealed the most noticeable facts of each juice.

Once the participants had finished observing an image, the corresponding juice was served to them, in a plastic and transparent cup, containing approximately 10 cl of juice at a temperature of about 12°C. Participants were requested to grade the juice in view of its taste and the image-information. They had to do this by putting with their mouse a mark on a horizontal bar shown on the screen. The left hand side of the bar mentioned “I do not like this product at all” and the right hand side “I like this product a lot.” These grades were instantly converted by the computer into numbers on a scale from 1 to 10, by measuring the distance between the mark and the left hand side of the bar. Adopting the terminology used in sensory studies, these numerical grades are referred to as the *actual scores.* It is assumed that they reflect preferences, i.e. if say juice A has a higher actual score than juice B, then juice A is preferred to juice B.

In the second part of the session the economic aspects of the experiment began. Subjects were told that they could buy the orange juices they had just evaluated. It was explained that they would be confronted with 5 different price/budget situations, and that they had to state their demand for the juices in each situation. We told them that one situation would be selected randomly and that they would receive the juices chosen in this situation as their remuneration for taking part in the experiment. Each situation thus had a chance of one in five of being selected, inciting the participants to "play" well in all situations.

Table 2 shows how the prices were chosen in each situation. In situation 1 the juice prices were equal to the market prices, i.e. the prices at which they were bought by us in the supermarket, while in situations 2, 3, 4 and 5 they were determined as functions of the actual scores and the market prices. In situation 1 the prices of the six juices were thus the same for all subjects. In each of the four other situations, the prices differed among individuals because the price calculations performed by the computer were based on individual-specific actual scores. The budget levels were determined as follows. In situations 1, 2, 3, and 4 we let the budget level correspond to the household's monthly expenditures on orange juice as declared by the consumer on the telephone.[6] In situation 5 we added FFr35 to the budget if the original budget level was below FFr 100, and substracted FFr35 otherwise. Below we explain what our motivation was for chosing these particular budget levels and the price formulas given in Table 2.

The 5 budget/price situations appeared successively and in random order on the screen, i.e. the situations did not necessarily show up in the same order for all individuals. For each situation, a table appeared on the screen with 6 columns corresponding to the 6 products. The columns were arranged according to the order in which the subject had evaluated the products. Each column was headed by a reduced size image of the product—which subjects could zoom if they desired to read the product information again. The scores attributed by the individual and the prevailing prices appeared below the images, while the available budget level was indicated in the lower left corner of the table. Subjects had to state their demand by entering the desired number of liters for each juice in the appropriate boxes of the table. A calculator was at their disposal to help them with the calculations. Once they had entered the desired amounts, they had to confirm their choice to the computer. If the cost of the chosen bundle of juices exceeded the budget, a warning message showed up on the screen, *forcing* the subjects to reconsider their choice. If the cost was such that at least one other juice could be purchased with the remaining budget, the computer *asked* but not obliged them to reconsider their choice. The computer recorded automatically the time spent by each subject to complete their demand table. When all participants had successfully completed the 5 demand tables, the remuneration situation was randomly selected.[7] At the end of the session the participants were asked to

[6] For a small number of individuals the declared level of expenditures was lower than FFr30 (resp. higher than FFr150). In those cases we attributed a budget level of FFr30 (resp. FFr150).

[7] A number between 1 and 5 was selected by a random number generator on the computer. If say number 2 was drawn, the participant received the bundle of juices chosen in the second situation he/she had faced. Since the budget/price situations appeared in a random order, the selected remuneration situation could correspond to any of the 5 situations.

answer a number of questions concerning their socio-economic characteristics.

Now we turn to the design of Experiment 2. Experiment 2 was held under the same conditions and at the same location as Experiment 1. Unlike Experiment 1, the subjects of Experiment 2 participated in two sessions, each one lasting about one hour. The duration separating the two sessions varied per individual, but was never less than one week and never more than three weeks. Since the data generated in the second session are not used in this paper, we primarily describe the design of the first session.

The session started with an evaluation of the products. Subjects were requested to evaluate the six orange juices—plus the dummy juice—on the basis of the product images only. The images were identical to the ones presented in Experiment 1, and appeared successively and in random order on the screen. Once the participants had finished looking at the characteristics of the product, they were asked to grade the juice, again by putting a mark on a horizontal bar. Here the left hand side of the bar mentioned "I will certainly not like this product at all" and the right hand side "I will certainly like this product a lot." Similarly as above the grades were converted into numbers varying between one and ten. Henceforth these numerical grades are referred to as the *expectation scores.* Like the actual scores they are assumed to reflect preferences among the orange juices.

The session then proceeded like the second part of Experiment 1: participants were told that they could purchase the products, that they would be confronted with 5 different price/budget situations, that one of the situations would be drawn at random, etc... Note that the juices were chosen and bought without knowledge of their taste, so the individuals were clearly less informed than those that participated in Experiment 1. The budget levels were determined as in Experiment 1. In calculating the prices, the formulas given in Table 2 were again applied, but here the expectation scores were used instead of the actual scores. Once all subjects had successfully filled in the 5 demand tables, the remuneration situation was randomly selected and several questions were asked about their socio-economic characteristics. The session ended with a blind tasting of the six orange juices: subjects were required to taste and grade each product but they could not see the corresponding juice-image. The data resulting from this analysis are not used in the paper, nor do we use the data generated in the second session of Experiment 2. Basically this session was identical to Experiment 1. The total earnings for the participants of Experiment 2 consisted in the remuneration situation of the first session plus the remuneration situation of the second session. It is important to note that when individuals made their juice choices in the first session, they were unaware of the fact that additional juices could be earned in the second session. Therefore, first session choices were not affected by income effects.

Before turning to a descriptive analysis of the data, we want to discuss some of the aspects of the experiments and compare the design of our experiments with the one of Sippel (1997). Our experiments were designed and carried out by economists and sensory scientists.[8] The collaboration with sensory scientists explains why real products were chosen for this study, i.e. products that could be smelled and tasted, and not fictitious ones. The disadvantage of this choice was that it implied a loss of experimental control over certain economic variables. Unlike laboratory auctions for instance, we had no control over the individuals' valuations for the products—in laboratory auctions, the value of a fictitious product is typically drawn from some distribution function, implying that valuations and hence also preferences are observed and controlled by the experimenter. The obvious advantage of real goods over fictitious ones was that they rendered the experiment more realistic and closer to a daily-life context. We felt that this was important given the fact that the participants were sampled from the Dijon population, and thus not necessarily as familiar with abstract forms of reasoning as students or faculty colleagues.

In determining the price formulas and budget levels, we had several objectives in mind. First of all the prices and budgets had to be realistic. This explains why we let the budget levels (in four of the five situations) correspond to the participants' declared household expenditures, and why the formulas in Table 2 were chosen such that the prices always fluctuated closely around the market prices of the juices. Second, the prices and budgets had to vary between individuals. Inter-individual variation of these variables is not necessary for the rationality tests reported in this paper, but is needed to identify the effects of prices and budget levels in the demand equations that we estimate. Finally, the prices had to be such that the power of the rationality test was high. Before the experiments were conducted, a Monte Carlo study was performed—in simulating purchase patterns we used data from Lange, Rousseau and Issanchou (1999); in that paper a random sample of individuals from Dijon reported their monthly juice expenditures, and graded a set of juices which was almost identical to the set of juices considered in our paper—to determine the price formulas in Table 2 such that the power was as high as possible against certain behavioral alternatives (these are defined in section 3).

Unlike the majority of experimental studies in economics, the participants in our study did not receive money for their participation but were paid in kind (Bateman *et al.* is another exception we are aware of; they partly remunerated

[8] The sensory scientists have analyzed the data in Lange, Issanchou and Combris (1998). They study the effects of prices, scores, budget levels and especially the level of product information on purchase behavior.

their experimental subjects with chocolate and Coke). They could fetch their juices at the tasting laboratory, about one week after the end of the experiment. Among the 120 participants of our experiment only 1 did not come to collect the products, indicating that the subjects were definitely not indifferent to the products we offered them. Note that our way of remunerating is actually quite similar to giving cash money. Indeed, because the subjects in our study were regular consumers of orange juice, the juices they received amounted to future monetary savings.

There are many differences between Sippel's experiment and our's, but here we mention only the major ones. Unlike Sippel, we did not recruit the subjects among students in law or economics but from the population of a city. This seems like an important advantage to us, since one is after all more interested in the consumer behavior of the population as a whole than in the consumer patterns of a very specific and non-representative sub-population. Students are likely to behave in a non-representative way, because they are relatively young and well educated, and also because their better knowledge of microeconomic theory can consciously or unconsciously affect their behavior (for instance, a well known phenomenon in experimental economics is that participants who are able to guess the goal of the study try to respond coherently with the theory to please the experimenter). Another important difference with Sippel's experiment is that we put more effort in rendering the experiment as close as possible to reality. The participants were familiar with the products, since they were only invited to participate in the study if they were regular consumers of the products, and in addition they received detailed information about the products during the experiment. Also, as mentioned above, the prices and subjects were chosen "realistically." Finally, his design differs in the number of price/budget situations with which the individuals were confronted (10 in his case and 5 in our case). While a larger number of situations is attractive in the sense that it increases the power of revealed preference tests, we show that there is some evidence that the number of revealed preference violations tends to increase disproportionally towards the end of our experiment, suggesting that one should keep the number of situations limited.

2.2. Descriptive analysis

Table 3 presents summary statistics for the socio-economic variables in our data set. The participants varied considerably in their observed characteristics. In both experiments the youngest person that participated was 19 years old, and the oldest 73 years. In Experiment 1 there were slightly more women than men, while in Experiment 2 the male participants were slightly in the majority. In both experiments there were single persons as well as individuals belonging to large households. In Experiment 2 the monthly household income varied between FFr1600 and FFr39040, and the average income was around FFr12500. In Experiment 1 the income ranged between FFr950 and FFr31600, and the mean income was around FFr14000. Here the descriptive statistics are based on 59 observations, since one person had not declared his/her income. In both experiments the average budget level for situations 1-4 was around FFr80.[9] As mentioned above, the budget level in situation 5 was constructed from the budget level in situations 1-4 by adding or substracting FFr35. As can be seen from Table 3, the resulting mean budget level for situation 5 was around FFr100 in Experiment 1, and around FFr90 in Experiment 2. Although the variable "budget level" is listed in Table 3 among the socio-economic characteristics—this seems appropriate given that, at least in situations 1-4, it corresponds to the household expenditures on orange juice—it will in the sequel of the paper act and be considered as the level of remuneration, i.e. rather as an experimental condition. According to the t-test, each variable appearing in Table 3 has the same mean in the two samples.[10]

Table 4 reports summary statistics for the scores. Recall that the actual scores were attributed by participants of Experiment 1, and the expectation scores by participants of Experiment 2. Note that the average expectation score is higher than the average actual score for all products except product 4. Table 5 reports mean prices and purchased quantities. Recall that in situation 1 the prices were equal to the market prices, so all individuals faced the same juice prices in that situation. In situations 2-5 prices were determined according to the formulas in Table 2 and, to facilitate the calculations for the participants, were rounded off to the nearest half-point (e.g. 6.67 was rounded off to 6.5). If the scores of an individual were such that the resulting price of one of the products was lower than FFr2 (resp. higher than FFr23), we set the price equal to FFr2 (resp. FFr23). Using Tables 4 and 5, the products can be ranked according to mean actual scores, according to mean expectation scores, or according to market prices. The ranking based on market prices is quite similar to the one based on mean

[9] One individual (with declared juice expenditures equal to FFr155) was wrongly attributed a budget level of FFr155 in situations 1-4 (instead of FFr150), and FFr120 in situation 5 (instead of FFr115).

[10] The average values reported in Table 3 are quite close to French national statistics. In 1997 the mean age of individuals living in France was 38.13 years, and women made up 52.70 % of the population; in 1990 the average number of persons per household was 2.57; in 1993 the average monthly household income was FFr14185. Source: INSEE, *Annuaire Statistique de la France, édition 1998.* Unfortunately, there are no national statistics on household expenditures on orange juice, so a comparison with the average budget levels in our data can not be made.

actual scores, but differs somewhat from the one based on mean expectation scores, suggesting that product evaluation reflects the price-ranking of products more accurately when the evaluation is conducted under complete information. Table 6 gives some additional information about the purchase behavior of the participants. To avoid too many details, the statistics of the variables are based on all observations, i.e. they are not calculated for each situation separately. Note that the subjects exhausted most of the budgets given to them: on average the participants of Experiment 1 only left 3% of their budget unspent, and those of Experiment 2 only 4%. In spite of the fact that the six products were fairly close substitutes, the subjects generally bought several juices at the same time: the expected number of different products bought in a situation was two and a half for participants of Experiment 1, and almost three for those of Experiment 2. Although the participants devoted the largest part of their budgets to either their most preferred product (Experiment 1) or their second most preferred product (Experiment 2), the expenditure shares for the less preferred juices were not negligeable.

Table 7 presents summary statistics for the times the individuals spent on filling in their demand tables. The calculation times varied between 21 and 1034 seconds. Given that the demand tables could, in principle, be completed in just a couple of seconds, this seems to suggest that the subjects chose their juices with care and took their task seriously. All mean calculation times are higher in Experiment 1 than in Experiment 2. Apparently, individuals had more difficulties to make up their mind when they were fully informed about the products. In both experiments the mean calculation time is highest for situation 5.[11] This might be due to the the fact that in situation 5 the price of the most preferred product was highest, suggesting that subjects needed the extra time to modify their purchase strategy and shift a larger part of their expenditures to less preferred products. Another explanation could be that the budget level attributed in situation 5 was different from the budget level attributed in the other situations, suggesting that participants needed the extra time to adapt to a new budget constraint.

3. Theory

3.1. Revealed Preference Theory

In this subsection, we present the various revealed preference axioms and their link with the utility maximization problem. Suppose that we have a dataset

[11] In a regression analysis that controls for the socio-economic variables and the order of appearance of the situations, the above effects persist: both an indicator variable for Experiment 2 and one for situation 5 have a significant impact on the calculation time.

of N individuals, and that there are S observations (or situations in the terminology of section 2) for each individual. Suppose there are K goods and let $p_s = (p_{1s}, ..., p_{Ks})'$ and $q_s = (q_{1s}, ..., q_{Ks})'$ denote the $K \times 1$ vectors of prices and associated quantities purchased by a consumer in situation s (we omit the individual-specific index in this and the next subsection). Let q_s and q_t be two bundles of goods with $s, t \in \{1, ..., S\}$ and consider the following definitions:[12]

1. q_s is directly revealed preferred to q_t, written $q_s R^0 q_t$, if $p_s' q_s \geq p_s' q_t$
2. q_s is strictly directly revealed preferred to q_t, written $q_s P^0 q_t$, if $p_s' q_s > p_s' q_t$
3. q_s is revealed preferred to q_t, written $q_s R q_t$, if there exists a sequence of bundles $(q_u, q_v, ..., q_w)$ such that $p_s' q_s \geq p_s' q_u$, $p_u' q_u \geq p_u' q_v, ..., p_w' q_w \geq p_w' q_t$
4. q_s is strictly revealed preferred to q_t, written $q_s P q_t$, if there exist observations q_u and q_v such that $q_s R q_u, q_u P^0 q_v, q_v R q_t$

We can now set out the main axioms of the revealed preference theory. The Weak Axiom of Revealed Preference (WARP) was first stated by Samuelson (1938): a set of observations satisfies WARP if $\forall\ (s,t) \in \{1, ..., S\}$, $q_s R^0 q_t$ and $q_s \neq q_t$ implies *not* $q_t R^0 q_s$. It is clear that a dataset resulting from a utility maximisation problem satisfies WARP but the converse is not true.

To resolve this problem, Houthakker (1950) formulated a stronger version of WARP, using revealed preference cycles: a set of observations satisfies the Strong Axiom of Revealed Preference (SARP) if $\forall\ (s,t) \in \{1, ..., S\}$, $q_s R q_t$ and $q_s \neq q_t$ implies *not* $q_t R q_s$. Houthakker showed that a set of observations (p_s, q_s), $i \in \{1, ..., S\}$, satisfies SARP if and only if there exists a utility function u (continuous, concave and monotonic) that rationalizes the data, i.e which verifies $u(q_s) \geq u(q_t)$ for all q_t such that $p_s' q_s \geq p_s' q_t$.

Varian (1982) generalized this result when he defined the Generalized Axiom of Revealed Preference (GARP): a set of observations satisfies GARP if $\forall\ (s,t) \in \{1, ..., S\}$, $q_s R q_t$ implies *not* $q_t P^0 q_s$. It is straightforward to demonstrate that SARP implies GARP but not vice versa. Indeed GARP allows multi-valued demand functions whereas SARP is only compatible with single-valued ones. To test if the data are compatible with GARP, Varian described Warshall's algorithm. This algorithm is quick and easy to program and is used in this paper.

[12] Our notation is identical to Varian's (1982), except that the situation is indexed by a subscript and not a superscript. It is important to note that all notions, tests, simulation procedures, and estimation methods described in this and the next two subsections are valid for the set of all situations as well as for arbitrary subsets. Therefore, in what follows, the set $\{1, ..., S\}$ can be replaced by any subset of situations.

The equivalence between GARP (or SARP) and the existence of a utility function which rationalizes the data is based on Afriat's theorem (1967). Unfortunately Afriat's utility function is neither differentiable nor strictly concave. Chiappori and Rochet (1987) introduced another axiom to avoid these problems: a set of observations satisfies the Strong version of the Strong Axiom of Revealed Preference (SSARP) if it satisfies SARP and if $\forall\ (s,t) \in \{1,...,S\}$, $p_s \neq p_t$ implies $q_s \neq q_t$. Under this additional restriction—the same bundle can not be obtained for two different price vectors—the authors show that it is possible to obtain a strictly increasing, infinitely differentiable, strongly concave utility function.

Afriat (1973) introduced the efficiency index e ($e \in [0,1]$), redefined the direct and strictly direct revealed preference relations as $q_s R^0(e) q_t$ if $e \cdot p_s' q_s \geq p_s' q_t$, and $q_s P^0(e) q_t$ if $e \cdot p_s' q_s > p_s q_t$, and introduced GARP(e): a set of observations satisfies GARP(e) if $\forall\ (s,t) \in \{1,...,S\}$, $q_s R(e) q_t$ implies *not* $q_t P^0(e) q_s$ ($R(e)$ is defined from $R^0(e)$ in the same way as R is from R^0). Note that at full efficiency, i.e. when $e = 1$, one has the original revealed preference relations and GARP. Afriat's efficiency index is useful to measure the "degree of inconsistency" of GARP violators. Indeed, defining e^* as the largest value of e such that the data of an inconsistent individual satisfies GARP(e^*), a GARP violator with e^* equal to say 0.8 is "more inconsistent" than a GARP violator with e^* equal to say 0.95.

3.2. Power of GARP tests

In this subsection we outline how to estimate the power of the GARP tests (the principle is the same for the other revealed preference axioms and GARP(e)) against behavioral alternatives. We consider four models of alternative behavior, all inspired by Becker (1962).

According to the first model, the random money choice model (Chant (1963)), the consumer spends money randomly among the goods, i.e. with a certain probability the consumer spends each unit of money on a particular good. Denoting π_{ks} the probability of choosing the good k in situation s and m_s the budget level in situation s, this yields the following model:

$$q_{ks} = \pi_{ks} \frac{m_s}{p_{ks}} \quad \forall s \in \{1,...,S\}, \quad \forall k = 1,...,K. \tag{3.1}$$

According to the second model, the random good choice model (Battalio, Dwyer and Kagel (1987)), the consumer randomly selects a collection of goods from those available. Denoting π_{ks} the random proportion associated with good

k in situation s, the model is:

$$q_{ks} = \pi_{ks} \sum_{l=1}^{K} q_{ls} \quad \forall s \in \{1, ..., S\}, \quad \forall k = 1, ..., K.$$

Since $\sum_{l=1}^{K} p_{ls} q_{ls} = m_s$, it follows immediately that in this model:

$$q_{ks} = \pi_{ks} \frac{m_s}{\sum_{l=1}^{K} \pi_{ls} p_{ls}} \quad \forall s \in \{1, ..., S\}, \quad \forall k = 1, ..., K. \tag{3.2}$$

The third and fourth models are variations on respectively model (3.1) and (3.2). Since we know (see section 2) how each individual ranked the products, we can construct for each consumer a partition (A, B) of $\{1, ..., K\}$ where A regroups the three most prefered goods, while B regroups the three least preferred goods. The third model, the random money choice model b), is then defined as model (3.1) but with $\pi_{ks} = 0 \ \forall k \in B$. According to this model, consumers adopt a two-stage strategy: first they decide which are their three most preferred products, and then they spend all their money exclusively on these three goods, but in a random way. The fourth model, the random good choice model b), is defined in a similar way, i.e. it is defined as model (3.2) but with $\pi_{ks} = 0 \ \forall k \in B$. Our partition (A, B) was loosely motivated by the results in Table 6, but is of course somewhat arbitrary and could have been chosen differently (for example by letting A correspond to the set of the four most preferred products, or the five most preferred products, etc.).

Like in Bronars (1987), the power of the GARP test against these four types of random behavior can be approximated using simulation methods. This is illustrated for model (3.1) (the simulation methods are similar for the other models). For each situation s, K i.i.d. uniform random variables u_{ks} are drawn and the probabilities are defined as $\pi_{ks} = u_{ks} / \sum_{l=1}^{K} u_{ls}$. Using the observed expenditure and prices in situation s and using (3.1) gives the simulated quantities q_{ks}^{*}. For each consumer, a set of simulated data $(p_{ks}, q_{ks}^{*}), s \in \{1, ..., S\}, k = 1, ..., K$, can be obtained in this way. Replicating the above procedure 100 times, i.e. generating for each individual 100 simulated data sets, the power of the test for a given individual can then be approximated by taking the average number (over the 100 replications) of times the individual violated GARP at least once. The average number of GARP violations in the sample can be approximated by the sum of the individual power estimates, and the power for the sample as a whole by the mean of the individual power estimates.

3.3. Estimation of demand systems with binding nonnegativity constraints

This subsection explains how the three systems of demand equations that were mentioned in the introduction are estimated. Estimation of the demand systems is not straigthforward in our case because the experimental sample contains many individuals with so-called corner solutions, i.e. individiduals with zero expenditure on one or more products in a given situation. Using the concept of virtual prices (see Neary and Roberts (1980)), Lee and Pitt (1982) proposed an estimation method that consists in transforming the binding nonnegativity constraints into nonbinding constraints. Their procedure requires, for each contribution to the likelihood function, the calculation of multiple integrals (with the number of integrals equal to the number of goods not purchased). In spite of recent advances in simulation methods that allow highly-dimensional integrals to be approximated quite accurately, the empirical implementation of their approach would remain very complex in our case. We therefore propose a different estimation strategy. Our method also uses the concept of virtual prices, but is much easier to implement because it avoids the evaluation of multiple integrals. The method amounts to estimating each demand equation separately by an iterative least squares procedure. Apart from the relative computational simplicity of our method, another advantage is that unlike Lee and Pitt it is not necessary to fully specify the distribution of the error terms in the demand system.

First we consider the translog demand system. Let $v^i_{ks} = p^i_{ks}/m^i_s$ and $w^i_{ks} = p^i_{ks}q^i_{ks}/m^i_s$ represent respectively the normalized price and the budget share of good k in situation s for individual i, $s \in \{1,...,S\}$, $k = 1,...,K$, $i = 1,...,N$ (the right-hand-side variables are those introduced in the previous subsections, except that an individual-specific index is now added). When the maximum value of the indirect translog utility function for individual i in situation s is such that the corresponding optimal quantities are strictly positive for all k (an interior solution), an application of Roy's identity implies the following form for the budget share w^i_{ks} (see Christensen, Jorgenson and Lau, 1975)

$$w^i_{ks} = \frac{\alpha^i_k + \sum_{j=1}^K \beta_{kj} \log(v^i_{js})}{\sum_{k=1}^K \alpha^i_k + \sum_{k=1}^K \sum_{j=1}^K \beta_{kj} \log(v^i_{js})}. \tag{3.3}$$

The scalar parameters α^i_k and β_{kj} are the preference parameters of the underlying indirect utility function. Note that none of the preference parameters varies with the situation-specific index s. This reflects our basic hypothesis that the preference structure for a given individual remained constant during the course of the experiment. Note also that the parameter α^i_k is allowed to differ over individuals, while β_{kj} is assumed constant. Since the parameters in the above share equation are only identified up to a multiplicative constant, a normalization of the parameters is necessary. A convenient normalization is $\sum_{k=1}^{K} \alpha_k^i = -1$ for all i. For a reason given below, we impose the homogeneity restrictions, i.e. we impose the restrictions $\sum_{k=1}^{K} \beta_{kj} = 0$ for all j. The parameter α_k^i is assumed to be of the form $\alpha_k^i = \alpha_{0k} + \alpha_{1k}' x^i + \varepsilon_k^i$, where α_{0k} is a scalar parameter, α_{1k} a vector of parameters, x^i a vector containing the observed characteristics of individual i, and ε_k^i a scalar random variable. The variable ε_k^i represents the effect of the omitted characteristics of individual i on the budget share of good k. The omitted characteristics are unobserved to us—which is why ε_k^i is considered as a random term—but known to individual i. The error terms are assumed to satisfy the two following standard assumptions: the vectors $\varepsilon^i = (\varepsilon_1^i, ..., \varepsilon_K^i)'$, $i = 1, ..., N$, are independent and follow the same distribution;[13] for each individual i, the variables $\varepsilon_1^i, ..., \varepsilon_K^i$ are independent from x^i and all normalized prices $v_{11}^i, ..., v_{KS}^i$. Given the normalization, the homogeneity assumption and the chosen form for α_k^i, the share equation (3.3) becomes

$$-w_{ks}^i = \alpha_{0k} + \alpha_{1k}' x^i + \sum_{j=1}^{K} \beta_{kj} \log(v_{js}^i) + \varepsilon_k^i. \qquad (3.3')$$

As mentioned above, many points in our sample are not interior solutions. When for an individual i in situation s at least one of the observed quantities q_{ks}^i equals zero, the nonnegativity constraints are binding, which means that an application of Roy's identity is inappropriate, which in turn means that the budget share w_{ks}^i is not defined as in (3.3'). Neary and Roberts have shown that there nonetheless exists a vector of normalized virtual prices that exactly support the observed quantities q_{ks}^i, i.e. there exists virtual prices—equal to v_{ks}^i if the purchased quantity of good k is nonzero and smaller than v_{ks}^i otherwise—such that the nonnegativity constraints are no longer binding. This means that Roy's identity *evaluated at the virtual prices* yields the correct expression for the budget shares. In the appendix it is shown how the virtual prices can be calculated for the translog model. It also shown that Roy's identity evaluated at the virtual prices yields a share equation of the form

$$-w_{ks}^i = \alpha_{0k} + \alpha_{1k}' x^i + \sum_{j=1}^{K} \beta_{kj} \log(\tilde{v}_j(x^i, v_s^i, r_s^i, \theta)) + \tilde{\varepsilon}_k^i(\varepsilon^i, r_s^i, \theta) \qquad (3.3")$$

where $\tilde{v}_j(\cdot)$ is a function that has the vectors $x^i, v_s^i = (v_{1s}^i, ..., v_{Ks}^i)'$, and θ as its arguments. Here θ is defined as the vector containing all parameters of the demand

[13] This assumption does not exclude dependence between ε_k^i and ε_l^i.

system, that is $\alpha_{0k}, \alpha'_{1k}$, and β_{kj} for all $j, k = 1, ..., K$. The form of the function $\tilde{v}_j(\cdot)$ depends[14] on the demand regime r^i_s. This scalar variable indicates which products are purchased and which are the ones not purchased by individual i in situation s, so that r^i_s can take the values $1, 2, ..., 2^K - 1$ (the regime where all products are purchased, the regime where only good 1 is not purchased, etc.). In equation (3.3"), $\tilde{\varepsilon}^i_k(\cdot)$ is a function with $\varepsilon^i_1, ..., \varepsilon^i_K$, and θ as its arguments, and its form depends on the demand regime r^i_s.

TO BE FINISHED

The Translog demand system to be estimated is made up of the share equations (3.3"), $k = 1, ..., K$. Each share equation [15] Estimating the unknown parameters in the demand system of K equationsmodel (3.3') and testing the various hypotheses that are of interest in this paper (equality of the parameters for the rationals and non-rationals, symmetry of the Slutsky matrix, etc.) is straightforward when the observed quantities q^i_{ks} are strictly positive for all i, k and s, i.e. when all points in the sample are interior solutions.

In the almost ideal demand system (see Deaton and Muellbauer, 1980) the budget share w^i_{ks} has the form

$$w^i_{ks} = \alpha^i_k + \sum_{j=1}^{K} \beta_{kj} \log(p^i_{js}) + \gamma_k \log(m^i_s / P^i_s). \tag{3.4}$$

In the share equation (3.4) P^i_s is a price index defined by

$$\log(P^i_s) = \sum_{k=1}^{K} \alpha^i_k \log(p^i_{js}) + \frac{1}{2} \sum_{j=1}^{K} \sum_{k=1}^{K} \beta_{kj} \log(p^i_{ks}) \log(p^i_{js}). \tag{3.5}$$

4. Results

Table 8 summarizes the results for the GARP tests. We report test results that are based on the real data, i.e. the data generated in our two experiments, and

[14] For notational simplicity, we write $\tilde{v}_j(x^i, v^i_s, r^i_s, \theta)$ instead of $\tilde{v}_{jr^i_s}(x^i, v^i_s, \theta)$, and $\tilde{\varepsilon}^i_k(\varepsilon^i, r^i_s, \theta)$ instead of $\tilde{\varepsilon}^i_{kr^i_s}(\varepsilon^i, \theta)$.

[15] This explains why the homogeneity hypothesis was made. When the homogeneity restrictions are not imposed, Roy's identity evaluated at the virtual prices leads to a share equation that is untractable compared to (3.3"). The share equation is a ratio in which both the nominator and denominator depend on the error terms, and statistical inference can then no longer be based on regression techniques. Each share equation still has a regression-equation form, in the sense that the dependent variable can be written as the sum of a function that depends on the exogenous variables, and an error term.

results based on simulated data. The column headed "Number of GARP violations" lists, for all possible combinations of price/budget situations, the number of individuals that violated GARP using the real data. For example, when the GARP test is based on the first, the second and the third situation with which we confronted the experimental subjects, i.e. when the last two situations that appeared are omitted, 13 out of 120 individuals violated GARP (the inconsistency rate in this case is thus approximately 11%). The column headed "Number of GARP violations: random good choice" lists, again for all possible combinations of situations, the average number of GARP violations in the sample—calculated according to the simulation strategy outlined in section 3—under the assumption that the individuals behaved as random money deciders. For example, when the fourth and fifth situation are omitted, the average number of individuals violating GARP is 9.2, so that the estimated power for this particular combination of situations equals almost 8% (9.2/120). The last three columns show the results for the other behavioral models that were presented in section 3.

Let us first look at the results of the GARP tests that are based on all five situations. As Table 8 shows, 35 out of 120 individuals violated GARP at least one time—15 individuals from Experiment 1, and 20 from Experiment 2—so that for 29% of the individuals the observed choices were not compatible with utility-maximizing behavior. In comparison, the inconsistency rates found by Cox (1997) and Sippel (1997) were respectively 37% and 42% (and 63% for his second experiment). The total number of GARP violations among the 35 non-rational individuals varied from 1 to 14, and the mean number of violations was 3.5. The estimated power against the four types of random behavior was 22%, 26%, 58% and 46% respectively. Note that in restricting the choice set from 6 to 3 goods, the power against random choice behavior increased, but the power against random goods choice behavior decreased. Cox and Sippel also reported the estimated power, but only against the hypothesis that consumers behave as random money deciders. The estimated power obtained by Cox (he had 5 products and 7 situations) was 47% for the GARP violators in his sample, and 51% for the non-violators. The estimated power obtained by Sippel (8 products and 10 situations) was 61% for his first experiment and 97% for his second experiment.

Next consider the results when the GARP tests are carried out on subsets of the 5 situations, i.e. when one, two or three situations are omitted. Table 8 shows that the percentage of inconsistent individuals and the power of the tests decrease with the number of omitted situations. The decrease in the percentage of inconsistent individuals depends on which situation(s) is(are) actually omitted. The first and the fifth situation account for a disproportionate part of the GARP violations in our sample. For instance, when the first or the fifth situation is omitted, the number of individuals violating GARP is only 21, whereas it equals 28 when the fourth situation is omitted. Similarly, when both the first and the fifth situation are omitted, only 8 individuals are non-rational, whereas there are 19 non-rationals when the third and fourth situation are omitted. The fact that the incidence of GARP violations is relatively high in the first and fifth situation is somewhat surprising given that the decrease in power is practically independent of which situation(s) is(are) omitted. Admittedly, we have only computed the power of the GARP test against a limited number of alternative types of behavior, so that our findings could simply reflect higher power—in the first and fifth situation—against some other behavioral strategy. But our findings can also be interpreted in another way. The relatively high incidence of GARP violations in the first situation might reflect learning effects, while the higher incidence in the fifth situation might result from weariness or loss of concentration among our experimental subjects. Note that these interpretations lead to opposite and conflicting recommendations. On the one hand, they seem to suggest that in experiments like our's one should add "warming-up" situations to avoid learning effects, while on the other hand one should keep the number of situations small—perhaps not more than 4 or 5—to avoid concentration problems.

Table 9 gives the results for the GARP(e) tests. A small diminution of the Afriat efficiency index e induces only a small reduction in the number of irrational individuals. For instance, when $e = 0.99$, still 30 individuals are irrational (inconsistency rate of 25%), and when $e = 0.95$, still 18 individuals are irrational (15%). In the data used by Cox, the inconsistency rate was 22% when $e = 0.99$ and 15% when $e = 0.95$ (these results are obtained from Sippel (1996)). In Sippel's first experiment the inconsistency rate was 25% when $e = 0.99$ and 8.3% when $e = 0.95$; in his second experiment he found 26.7% when $e = 0.99$ and 10% when $e = 0.95$. The inconsistency rates found by Cox and Sippel are clearly less robust to small changes in the efficiency index. Although their rates of GARP violations are relatively higher at full efficiency, they are comparable to our's (or lower) even for small inefficiency levels. This means that in the samples of Cox and Sippel there were comparatively many non-rational individuals that "almost" satisfied GARP. Consider now the power estimates for the different values of Afriat's efficiency index. For simplicity, Table 9 only reports estimates of the power against the random money choice model and the random good choice model. The estimated average number of individuals violating GARP—under the assumption that they behaved as random money deciders—equals 20.22 when $e = 0.99$ (estimated power is 20.22/120=17%) and 5.68 when $e = 0.95$; under the assumption that individuals behaved as random good deciders, the estimates are 59.47 when $e = 0.99$ (50%) and 27.62 when $e = 0.95$ (23%). In the data used by Cox, the estimated power was 25% when $e = 0.99$ and 12% when $e = 0.95$ (these results are again taken from Sippel (1996)[16]). The estimated power obtained by Sippel in his first experiment was 47% when $e = 0.99$ and 17% when $e = 0.95$; in his second experiment he found 65% when $e = 0.99$ and 13% when $e = 0.95$. Our estimated power using the random money choice model decreases rapidly with the level of inefficiency—like in Sippel's data—and is rather low when $e = 0.95$. On the contrary, our estimated power using the random good choice model decreases slowly—like in Cox's data—and at the 95% efficiency level it is considerably higher than the power estimates obtained by Cox and Sippel.

Let us now turn to the analysis of the determinants of non-rationality. To study whether there are variables that have an impact on the probability of being non-rational, we created for each subject the dummy variable garp, where garp=0 for an individual whose choices did not satisfy GARP and 1 otherwise, and estimated a logit model with garp as the independent variable, and the socio-economic characteristics and experimental conditions (the calculation times and the remuneration level) as the independent variables. The results of the logit regressions—one for each experiment—are presented in Table 10. The variables "age", "size household" and "income" have no significant impact on the probability of being non-rational. The variable "gender" is significantly positive in Experiment 2, suggesting that when only partial information was given to the participants, women were less likely to be irrational than men. The variable "budget level" (defined as the budget level attributed in situations 1, 2, 3, 4; results were not different when this variable was defined as the budget in situation 5 or as the (weighted) average of the budget levels) is insignificant in both regressions, so that the likelihood of violating the utility-maximization hypothesis was (fortunately) not influenced by the height of remuneration. We initially included the original calculation times—those presented in Table 7—in the logit models. None of these variables had a significant effect though. The mean of the calculation times did not have a significant effect either. We subsequently defined the calculation times ordered according to the sequence in which the situations arrived, and when these variables were included in the logit models, two appeared to be significant in the regression for Experiment 1. Since the results with the ordered calculation times are the most successful, these are the ones that are retained and presented in Table 10.[17]

[16] Sippel (1996) reported power estimates based on all individuals in Cox's data set, and not separately for the violators and non-violators. He found an estimate equal to 28% when $e = 1$, which is somewhat low given the results reported in Cox (1997).

[17] The results regarding the other variables in Table 10 were not fundamentally different when the original calculation times or their mean value were included in the model.

To evaluate the impact of the calculation times, the idea to regress garp on the (ordered) calculation times seems natural at first sight. However, the coefficients corresponding to the time variables can be difficult to interpret when there are different types of GARP violators (as is the case in our sample). To see this, consider the following simple example. Suppose there are 3 groups of individuals. The first group is made up of rational individuals, while the second and third group are made up of irrational individuals, all with just one GARP violation. The individuals in the second group have GARP violation $q_{1st}Rq_{2nd}$ and $q_{2nd}P^0q_{1st}$, while those in the third group have GARP violation $q_{4th}Rq_{5th}$ and $q_{5th}P^0q_{4th}$. The individuals in the second group violate GARP in the first 2 situations, but behave in a manner that is not incompatible with utility-maximization in the last 3 situations. Inversely, the individuals in the third group behave like utility-maximizing agents in the first 3 situations, but violate GARP in the last 2 situations. Because of this asymmetry, the effects of the ordered calculation times (on the probability of being non-rational) are likely to differ for these 2 types of GARP violators. The logit model cannot capture this since the coefficients of the time variables are identical, regardless of the type of GARP violation. As a consequence of aggregation over the different types of GARP violators, the logit estimates (based on the individuals from the 3 groups) then represent a mixture of group-specific effects, and not the pure causal effects of ordered calculation times on the probability of violating GARP.

To better understand the impact of the calculation times, we proceeded in the following way. For each individual two categories of dummy variables were created: the gexp_i variables and the gdexp_s variables, $s = 1, ..., 5$. The variable $\text{gexp}_s = 0$ if the individual violated GARP and if in addition the sth situation played a role in (one of) the violation(s), and 1 otherwise. For example, when $q_{1st}R^0q_{2nd}$ and $q_{2nd}P^0q_{1st}$, then $\text{gexp}_1 = 0$ and $\text{gexp}_2 = 0$; when $q_{1st}Rq_{2nd}$ (with $q_{1st}R^0q_{3rd}$ and $q_{3rd}R^0q_{2nd}$) and $q_{2nd}P^0q_{1st}$, then $\text{gexp}_1 = \text{gexp}_2 = \text{gexp}_3 = 0$. The second category of variables was defined in a similar way. The variable $\text{gdexp}_s = 0$ if the individual violated GARP and if in addition the sth situation played a "direct" part in (one of) the violation(s), and 1 otherwise (the sth situation is said to play a "direct" part in a GARP violation if and only if there exists u such that $q_{sth}R^0q_{uth}$ and $q_{uth}P^0q_{sth}$, or $q_{uth}R^0q_{sth}$ and $q_{sth}P^0q_{uth}$). We then estimated logit models, with gdexp_s (resp. gexp_s) as the independent variable and the calculation time in the sth situation as the dependent variable. Note that unlike the logit regressions on garp, the estimated coefficients have a direct and unambiguous interpretation. The results are given in Table 11 (only the estimates based on the entire sample are given; the regressions on the 2 experimental subsamples gave similar results). The ordered calculation time is significantly positive in the logit regressions on $gdexp_1$ and $gexp_1$, and significantly negative in the regressions on $gdexp_2$ and $gexp_2$. The ordered calculation time is weakly significant (p-value is 0.12) and negative in the logit regression on $gdexp_3$. These results suggest that the calculation times had an effect on the probability of violating GARP: individuals who spent relatively much time on the first situation which they encountered, and who acted quickly in the subsequent situations, were less likely to violate the utility-maximization hypothesis.

REMAINING RESULTS HERE

5. Conclusion

References

[1] Afriat, S. (1967), "The construction of a utility function from expenditure data," *International Economic Review,* 8, 67-77.

[2] Afriat, S. (1973), "On a system of inequalities on demand analysis: an extension of the classical method," *International Economic Review*, 14, 460-472.

[3] Bateman, I., A. Munro, B. Rhodes, C. Starmer and R. Sugden (1997), "A test of the theory of reference-dependent preferences," *Quarterly Journal of Economics*, 112, 479-505.

[4] Battalio, R.C., G.P. Dwyer and J.H. Kagel (1987), "Tests of competing theories of consumer choice and the representative consumer hypothesis," *Economic Journal*, 97, 842-856.

[5] Battalio, R.C., J.H. Kagel, R.C. Winkler, E.B. Fisher, R.L. Basmann and L. Krasner (1973), "A test of consumer demand theory using observations of individual consumer preferences," *Western Economic Journal*, 11, 411-428.

[6] Becker, G.S. (1962), "Irrational behavior and economic theory," *Journal of Political Economy,* 70, 1-13.

[7] Blundell, R., P. Pashardes and G. Weber (1993), "What do we learn about consumer demand patterns from micro data," *American Economic Review*, 83, 570-597.

[8] Blundell, R. and J.-M. Robin (1999), "Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems," *Journal of Applied Econometrics,* 14, 209-232.

[9] Bronars, S.G. (1987), "The power of nonparametric tests of preference maximization," *Econometrica,* 55, 693-698.

[10] Camerer, C. (1995), "Individual decision making" in J.H. Kagel and A.E. Roth, eds., *The handbook of experimental economics.* Princeton: Princeton University Press, 587-703.

[11] Chant, J.F. (1963), "Irrational behavior and economic theory: a comment," *Journal of Political Economy,* 71, 505-510.

[12] Chiaporri, P.A. and J.C. Rochet (1987), "Revealed preferences and differentiable demand," *Econometrica*, 55, 687-691.

[13] Christensen, L.R., D.W. Jorgenson and L.J. Lau (1975), "Transcendental logarithmic utility functions," *American Economic Review,* 65, 367-383.

[14] Cox, J.C. (1997), "On testing the utility hypothesis," *Economic Journal,* 107, 1054-1078.

[15] Deaton, A.S. and J. Muellbauer (1980), "An almost ideal demand system," *American Economic Review,* 70, 312-326.

[16] Houthakker, H.S. (1950), "Revealed preference and the utility function," *Economica*, 17, 159-174.

[17] Kagel, J.H., R.C. Battalio and L. Green (1995), *Economic choice theory: an experimental analysis of animal behavior.* Cambridge: Cambridge University Press.

[18] Lange, C., S. Issanchou and P. Combris (1998), "Expected or experienced quality: trade-off with price," *Food Quality and Preference* (forthcoming).

[19] Lange, C., F. Rousseau and S. Issanchou (1998), "Expectation, liking and purchase behaviour under economical constraint," *Food Quality and Preference,* 10, 31-39.

[20] Lee, L.-F. and M.M. Pitt (1986), "Microeconometric demand systems with binding nonnegativity constraints: the dual approach," *Econometrica*, 54, 1237-1242.

[21] MacFie, H.J., N. Bratchell, K. Greenhoff and L.V. Vallis (1989), "Designs to balance the effect of order of presentation and first-order carry-over effects in Hall tests," *Journal of Sensory Studies,* 4, 129-148.

[22] Neary, J.P. and K.W.S. Roberts (1980), "The theory of household behaviour under rationing," *European Economic Review*, 13, 25-42.

[23] Samuelson, P.A. (1938), "A note on the pure theory of consumer's behaviour," *Economica,* 5, 61-71.

[24] Sippel, R. (1996), "A note on the power of revealed preference tests with Afriat inefficiency," *Bonn University Discussion Paper,* No. A-528.

[25] Sippel, R. (1997), "An experiment on the pure theory of consumer's behaviour," *Economic Journal*, 107, 1431-1444.

[26] Varian, H. (1982), "The nonparametric approach to demand analysis," *Econometrica*, 50, 945-972.

[27] Varian, H. (1983), "Non-parametric tests of consumer behaviour," *Review of Economic Studies,* 50, 99-110.

[28] Varian, H. (1988), "Revealed preference with a subset of goods," *Journal of Economic Theory*, 46, 179-185.

APPENDIX (TO BE FINISHED)

1. Calculation of virtual prices

To see how this works, consider the regime where only good 1 is not purchased, i.e. $q^i_{1s} = 0$, $q^i_{ks} > 0$, $k = 2, ..., K$. The virtual price for good 1, $\tilde{v}^i_{1s}$, is solved from the equation

$$0 = \alpha_{01} + \alpha'_{11}x^i + \beta_{11}\log(\tilde{v}^i_{1s}) + \sum_{j=2}^{K}\beta_{1j}\log(v^i_{js}) + \varepsilon^i_1,$$

so that

$$\log(\tilde{v}^i_{1s}) = -\frac{(\alpha_{01} + \alpha'_{11}x^i + \sum_{j=2}^{K}\beta_{1j}\log(v^i_{js}))}{\beta_{11}} - \frac{\varepsilon^i_1}{\beta_{11}}.$$

In this regime the virtual prices for the other goods equal the observed prices, i.e. $\tilde{v}^i_{ks} = v^i_{ks}$, $k = 2, ..., K$. Consider next the regime where the first two goods are zero and all remaining goods are purchased. The virtual prices $\tilde{v}^i_{1s}$ and $\tilde{v}^i_{2s}$ are then solved from

$$\begin{pmatrix} \log(\tilde{v}^i_{1s}) \\ \log(\tilde{v}^i_{2s}) \end{pmatrix} = -\begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}^{-1} \left\{ \begin{pmatrix} \alpha_{01} + \alpha'_{11}x^i + \sum_{j=3}^{K}\beta_{1j}\log(v^i_{js}) \\ \alpha_{02} + \alpha'_{12}x^i + \sum_{j=3}^{K}\beta_{2j}\log(v^i_{js}) \end{pmatrix} + \begin{pmatrix} \varepsilon^i_1 \\ \varepsilon^i_2 \end{pmatrix} \right\}$$

and $\tilde{v}^i_{ks} = v^i_{ks}$, $k = 3, ..., K$.

As can be seen, for both regimes, the logarithm of the virtual prices can be expressed as the sum of two terms. The first term depends only on the normalized prices and the observed characteristics, and the second term only on the unobserved variables. It is not difficult to show that this is true for all possible demand regimes: the logarithm of the virtual prices can always be written as the sum of a function of the exogenous variables and a function of the unobserved variables.

2. Asymptotic properties of the estimator

The estimations of our three models are quite similar (we present here only the estimation of the Translog model) and are based on a method develloped by R.Blundell and J.M.Robin (1999). In their article, they estimate a model

$$y_{kt} = g(x_t, \theta^0)'\theta^0_k + u_{kt}, \quad k = 1..K, \quad t = 1..T$$

using iterated least squares ($\theta_i^0 \in \mathcal{P} \subset IR^L$ is a parameter with $\theta^0 = (\theta_1^{0'}, ..., \theta_K^{0'})$ and g is a L-vector of functions of x_t and θ^0 (see page 211, equation 1)).

The $(p+1)$th iteration of their algorithm yields the following value for the parameter associated with the kth equation :

$$\theta_k^{(P+1)} = [G(\theta^{(p)})' G(\theta^{(p)})]^{-1} G(\theta^{(p)})' \mathbf{y}_k$$

where $\mathbf{y}_k = (y_{k1}, ..., y_{kT})'$ and $G(\theta) = \begin{pmatrix} g(x_1, \theta)' \\ \vdots \\ g(x_T, \theta)' \end{pmatrix}$.

Noting $\widehat{\theta}$ the limit value of such recursive sequence, they obtain (theorems 1 and 2 page 213) that

Proposition 2.1. *$\widehat{\theta}$ converges almost surely to θ^0 and that $\widehat{\theta}$ is asymptotically normal :*

$$\sqrt{T}(\widehat{\theta} - \theta^0) \rightsquigarrow N(0, J_0^{-1}(\Sigma_0 \otimes L_0)(J_0')^{-1})$$

with

$$L_0 = E[g(x_t, \theta^0) g(x_t, \theta^0)']$$

$$\Sigma_0 = E(u_t u_t' | x_t)$$

and

$$J_0 = I_K \otimes L_0 + E\left[\left(\Theta_0' \frac{\partial g(x_t, \theta^0)}{\partial \theta'}\right) \otimes g(x_t, \theta^0)\right]$$

where $u_t = (u_{1t}, .., u_{Kt})'$ and $\Theta_0 = (\theta_1^0, ..., \theta_K^0)$.

Consider now the Translog model :

$$-w_{ks}^i = \alpha_{0k} + \alpha_{1k}' x^i + \sum_{j=1}^K \beta_{kj} \log(\tilde{v}_j(x^i, v_s^i, r_s^i; \theta)) + \tilde{\varepsilon}_k^i(r_s^i) \qquad (3.3")$$

One can rewrite it in the desired form because :

$$-w_{ks}^i = \begin{pmatrix} 1 & x^{i\prime} & \log(\tilde{v}_1(x^i, v_s^i, r_s^i; \theta)) \cdots \log(\tilde{v}_K(x^i, v_s^i, r_s^i; \theta)) \end{pmatrix} \begin{pmatrix} \alpha_{0k} \\ \alpha_{1k} \\ \beta_{k1} \\ \vdots \\ \beta_{kK} \end{pmatrix} + \tilde{\varepsilon}_k^i(r_s^i) \qquad (2.1)$$

If we note $t = (i, s)$ (an observation is an individual i in a given situation s), e a vector that has every components equal to one and g a $K+1+?$ vector of functions of e_t, x_t', v_t, r_t and θ and if we remark that in our model $\theta = (\theta_1', ..., \theta_K')$ where $\theta_k = (\alpha_{0k}, \alpha_{1k}, \beta_{k1}, ..., \beta_{kK})'$, we obtain the following form for the Translog model :

$$-w_{ks}^i = h(e_t, x_t', v_t, r_t; \theta)\theta_k + \tilde{\varepsilon}_{kt}(r_t) \quad (3.3")$$

Unfortunately, $E(\tilde{\varepsilon}_{kt}(r_t)|e_t, x_t', v_t, r_t) = C(r_t)$ and one can not apply the results of Blundell and Robin because they need $E(u_{kt}|x_t) = 0$ (assumption 2 page 211). To solve this problem, we substract from each observation the average of the observations in the subpopulation $R(r_s^i)$ ($R(r)$ is the subset of all the couples (i, s) in the regime r i.e. $R(r) = \{(i, s) \text{ such that } r_s^i = r\}$).

Introducing

- $\overline{w}_{ks}^i = \frac{1}{card(R(r_s^i))} \sum_{(i',s') \in R(r_s^i)} w_{ks'}^{i'}$ the average of the budget share on $R(r_s^i)$,
- $\overline{x}^i = \frac{1}{card(R(r_s^i))} \sum_{(i',s') \in R(r_s^i)} x^{i'}$ the average of the exogeneous variables,
- $\log(\overline{\tilde{v}_j}) = \frac{1}{card(R(r_s^i))} \sum_{(i',s') \in R(r_s^i)} \log(\tilde{v}_j(x^{i'}, v_{s'}^{i'}, r_{s'}^{i'}; \theta))$ the average of the virtual normalized prices,
- $\overline{\log(v_s^i)} = \frac{1}{card(R(r_s^i))} \sum_{(i',s') \in R(r_s^i)} \log(v_{s'}^{i'})$ the average of the normalized prices,
- $\overline{\tilde{\varepsilon}}_k^i(r_s^i) = \frac{1}{card(R(r_s^i))} \sum_{(i',s') \in R(r_s^i)} \tilde{\varepsilon}_k^{i'}(r_{s'}^{i'})$ the average of the error terms,

the Translog model becomes :

$$-(w_{ks}^i - \overline{w}_{ks}^i) = \alpha_{1k}'(x^i - \overline{x}^i) + \sum_{j=1}^{K} \beta_{kj} \left(\log(\tilde{v}_j(x^i, v_s^i, r_s^i; \theta)) - \log(\overline{\tilde{v}_j}) \right) + \tilde{\varepsilon}_k^i(r_s^i) - \overline{\tilde{\varepsilon}}_k^i(r_s^i) \quad (3.3")$$

Looking back at the definition of the virtual normalized prices, it is possible to remark that

$$\log(\tilde{v}_j(x^i, v_s^i, r_s^i; \theta)) - \log(\overline{\tilde{v}_j}) = l(x^i - \overline{x}^i, \log(v_s^i) - \overline{\log(v_s^i)}, r_s^i; \theta)$$

and thus we obtain

$$-(w_{kt} - \overline{w}_{kt}) = g(x_t - \overline{x}_t, \log(v_t) - \overline{\log(v_t)}, r_t; \theta)\theta_k + \tilde{\varepsilon}_{kt}(r_t) - \overline{\tilde{\varepsilon}}_{kt}(r_t) \quad (3.3")$$

This time the assumption $E(u_{kt}|x_t) = 0$ is verified because

$$E(\tilde{\varepsilon}_{kt}(r_t) - \overline{\tilde{\varepsilon}}_{kt}(r_t)|x_t - \overline{x}_t, \log(v_t) - \overline{\log(v_t)}, r_t) = C(r_t) - C(r_t) = 0$$

Nevertheless, there still exists a difference between our model and the model of Blundell and Robin as exposed in their paper because the residuals $\tilde{\varepsilon}_{kt}(r_t) - \overline{\tilde{\varepsilon}}_{kt}(r_t)$ are not idependent (assumption 1 page 211). Indeed $\tilde{\varepsilon}_k^i(r_s^i)$ and $\tilde{\varepsilon}_k^i(r_{s'}^i)$ are correlated because they result from the omission of non observed variables. As the authors mentionned it, it is straitforward to extend their results in this case.

Proposition 2.2. *In our model, the estimator $\widehat{\theta}$ converges almost surely to θ^0 and $\widehat{\theta}$ is asymptotically normal :*

$$\sqrt{T}(\widehat{\theta} - \theta^0) \rightsquigarrow N(0, J_0^{-1}\Omega(J_0')^{-1})$$

where $\Omega = (\Omega_{kl})_{k=1,..,K;l=1,..,K}$ and

$$\Omega_{kl} = E\left(\frac{1}{T}\sum_i\sum_s\sum_{s'} g(Z_{(i,s)};\theta)g(Z_{(i,s')};\theta)'\eta_{k(i,s)}\eta_{l(i,s')}\right)$$

with the notations

$$g(Z_{(i,s)};\theta) = g(x_{(i,s)} - \overline{x}_{(i,s)}, \log(v_{(i,s)}) - \overline{\log(v_{(i,s)})}, r_{(i,s)};\theta)$$

and

$$\eta_{k(i,s)} = \tilde{\varepsilon}_{k(i,s)}(r_{(i,s)}) - \overline{\tilde{\varepsilon}}_{k(i,s)}(r_{(i,s)})$$

Proof. The proof is identical to Blundell and Robin's one. The only difference is the expression of the term

$$\Omega_{kl} = E\left(\frac{1}{T}\sum_t\sum_{t'} g(x_t;\theta)g(x_{t'};\theta)'u_{kt}u_{lt}\right)$$

in the variance. In their case, the hypothesis $E(u_{kt}u_{lt'}|x_t, x_{t'}) = 0$ allows them to simplify this expression to obtain

$$\Omega_{kl} = E\left(\frac{1}{T}\sum_t\sum_{t'} g(x_t;\theta)g(x_{t'};\theta)'u_{kt}u_{lt}\right) = E(g(x_t;\theta)g(x_t;\theta)')\sigma_{kl}$$

and consequently

$$\Omega = \Sigma_0 \otimes L_0$$

In our case, because of the correlation between u_{kt} and u_{lt}, we only obtain

$$\Omega_{kl} = E\left(\frac{1}{T}\sum_i\sum_s\sum_{s'} g(Z_{(i,s)};\theta)g(Z_{(i,s')};\theta)'\eta_{k(i,s)}\eta_{l(i,s')}\right)$$

■

Proposition 2.3. *An estimator of this asymptotique variance-covariance matrix is*

$$\widehat{V}_{as}\widehat{\theta} = \widehat{J}^{-1}\widehat{\Omega}\widehat{J'}^{-1}$$

where

$$\widehat{J} = [I_K \otimes G(\widehat{\theta})'] \frac{\partial (I_K \otimes G(\widehat{\theta}))\widehat{\theta}}{\partial \theta'}$$

and

$$\widehat{\Omega} = \frac{1}{T} \sum_i \sum_s \sum_{s'} g(Z_{(i,s)}; \widehat{\theta}) g(Z_{(i,s')}; \widehat{\theta})' \widehat{\eta}_{k(i,s)} \widehat{\eta}_{k(i,s')}$$

with

$$\widehat{\eta}_{kt} = -(w_{kt} - \overline{w}_{kt}) - g(x_t - \overline{x}_t, \log(v_t) - \overline{\log(v_t)}, r_t; \widehat{\theta})$$

Proof. The estimator $\widehat{J}$ of J_0 is the estimator introduced by Blundell and Robin (page214). The estimator

$$\widehat{\Omega} = \frac{1}{T} \sum_i \sum_s \sum_{s'} g(Z_{(i,s)}; \widehat{\theta}) g(Z_{(i,s')}; \widehat{\theta})' \widehat{\eta}_{k(i,s)} \widehat{\eta}_{k(i,s')}$$

of Ω is the White's estimator.

■

Table 1
The six products

Product	Information
1	Pure orange juice; bottle
2	Pure orange juice; cardboard pack; product from Morocco
3	Concentrated orange juice; bottle; guaranteed content of vitamin C
4	Concentrated orange juice; cardboard pack; product from Florida
5	Nectar; bottle; 55% of pure orange
6	Nectar; cardboard pack; 50% of pure orange

Table 2
Calculation of prices

Situation	Price most preferred product	Price second most preferred produc
1	MP(1)	MP(2)
2	$\text{MP}(1)\cdot\left\{1+\frac{score(1)}{\sum score(j)}\right\}$	$\text{MP}(2)\cdot\left\{1+\frac{score(2)+score(3)+score(4)}{\sum score(j)}\right.$
3	$\text{MP}(1)\cdot\left\{1+\frac{score(1)+score(2)}{\sum score(j)}\right\}$	$\text{MP}(2)\cdot\left\{1+\frac{score(2)+score(3)+score(4)+sco}{\sum score(j)}\right.$
4	$\text{MP}(1)\cdot\left\{1+\frac{score(1)+score(2)+score(3)}{\sum score(j)}\right\}$	$\text{MP}(2)\cdot\left\{1+\frac{score(2)}{\sum score(j)}\right\}$
5	$\text{MP}(1)\cdot\left\{1+\frac{score(1)+score(2)+score(3)+score(4)}{\sum score(j)}\right\}$	$\text{MP}(2)\cdot\left\{1+\frac{score(2)+score(3)}{\sum score(j)}\right\}$

Situation	Price fourth most preferred product	Price fifth most preferred product
1	MP(4)	MP(5)
2	$\text{MP}(4)\cdot\left\{1-\frac{score(6)}{\sum score(j)}\right\}$	$\text{MP}(5)\cdot\left\{1-\frac{score(3)+score(4)+score(5)}{\sum score(j)}\right.$
3	$\text{MP}(4)\cdot\left\{1-\frac{score(5)+score(6)}{\sum score(j)}\right\}$	$\text{MP}(5)\cdot\left\{1-\frac{score(2)+score(3)+score(4)+sco}{\sum score(j)}\right.$
4	$\text{MP}(4)\cdot\left\{1-\frac{score(4)+score(5)+score(6)}{\sum score(j)}\right\}$	$\text{MP}(5)\cdot\left\{1-\frac{score(5)}{\sum score(j)}\right\}$
5	$\text{MP}(4)\cdot\left\{1-\frac{score(3)+score(4)+score(5)+score(6)}{\sum score(j)}\right\}$	$\text{MP}(5)\cdot\left\{1-\frac{score(4)+score(5)}{\sum score(j)}\right\}$

Notes: MP(j) is the market price of the j-th most preferred product; $score(j)$ is the score (actual score or expectation score) of the j-th most preferred product, i.e. $score(1) \geq score(2) \geq ... \geq score(6)$.

Table 3
Summary statistics socio-economic variables

Variable	Experiment 1				Experiment 2	
	Mean	Std. dev.	Min	Max	Mean	Std. dev.
Age in years	37.85	14.82	19	73	36.20	13.07
Gender (=1 if woman)	0.57	0.50	0	1	0.48	0.50
Size household	2.75	1.29	1	7	2.88	1.49
Income (in FFr)	13946.10	8454.53	950	31600	12305.25	7289.91
Budget level (in FFr) for situations 1, 2, 3, 4	80.75	39.32	30	155	75.17	41.78
Budget level (in FFr) for situation 5	98.42	21.79	65	130	92.67	19.90

Table 4
Summary statistics actual scores and expectation scores

	actual score				expectation score			
Product	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max
1	4.61	2.43	1.00	9.82	7.25	1.99	1.99	9.91
2	6.32	2.29	1.27	9.91	6.74	1.85	1.45	9.91
3	4.20	2.04	1.00	8.20	5.72	2.15	1.72	9.82
4	5.08	2.10	1.09	9.82	4.91	2.06	1.00	9.28
5	4.19	2.09	1.00	9.64	4.81	2.30	1.09	9.82
6	3.59	2.10	1.00	9.64	4.89	2.16	1.00	10.00

Table 5
Mean prices (in FFr) and quantities: Experiment 1; Experiment 2

	Situation 1	Situation 2	Situation 3	Situation 4	Situation 5
			Prices		
Product 1	7;7	6.76;7.98	7.01;8.61	7.48;8.98	7.72;9.42
Product 2	11;11	14.41;13.02	15.29;13.98	14.82;13.18	15.49;13.92
Product 3	10;10	9.03;10.86	8.74;10.82	9.82;10.71	9.30;10.43
Product 4	5;5	6.13;4.43	6.08;4.23	5.73;4.72	5.51;4.53
Product 5	6;6	5.93;5.07	5.89;4.95	5.98;5.43	5.98;5.52
Product 6	3;3	2.63;3.19	2.68;3.02	2.72;2.96	2.87;2.87
			Quantities		
Product 1	2.15;2.97	2.08;2.78	1.35;2.83	2.23;2.45	2.76;3.30
Product 2	3.35;1.92	2.38;1.62	2.25;1.35	2.07;1.73	1.95;1.90
Product 3	0.67;1.45	0.53;1.17	0.47;1.00	0.33;0.83	0.77;0.95
Product 4	1.73;1.08	1.57;1.43	1.48;1.33	1.70;1.13	2.52;1.20
Product 5	1.15;0.82	1.03;1.18	1.47;0.88	1.45;0.75	1.62;0.95
Product 6	1.68;1.97	1.62;1.62	1.32;2.25	1.05;2.00	1.62;2.12

Table 6

Purchase behavior of participants

Variable	Experiment 1 Mean	Std. dev.	Min	Max	M
(Budget level minus expenditure)/budget level	0.03	0.06	0	0.4	0.
Number of different products purchased	2.56	1.03	1	5	2.
Expenditure share of most preferred product	0.36	0.33	0	1	0.
Expenditure share of second most preferred product	0.20	0.27	0	1	0.
Expenditure share of third most preferred product	0.15	0.25	0	1	0.
Expenditure shareof fourth most preferred product	0.07	0.18	0	1	0.
Expenditure share of fifth most preferred product	0.07	0.17	0	1	0.
Expenditure share of least preferred product	0.12	0.20	0	1	0.

Table 7

Summary statistics calculation times (in seconds)

	Experiment 1				Experiment 2	
Variable	Mean	Std. dev.	Min	Max	Mean	Std. dev.
Calculation time situation 1	205.42	189.75	21	856	166.88	147.90
Calculation time situation 2	228.03	200.70	40	1034	163.87	115.28
Calculation time situation 3	206.68	153.23	32	592	172.58	122.87
Calculation time situation 4	203.28	173.26	28	785	162.12	141.16
Calculation time situation 5	237.72	174.31	62	843	208.83	146.31

Table 8

GARP violations

Number of omitted situations	Omitted situations	Number of GARP violations (%)	Number of GARP violations: random money choice (%)	Number of GARP violations: random money choice b) (%)	Nu vi go
0	Ø	35 (29)	26.43 (22)	31.59 (26)	
1	1st	21 (18)	17.29 (14)	20.70 (17)	
	2nd	22 (18)	17.07 (14)	21.19 (18)	
	3rd	27 (23)	17.00 (14)	20.80 (17)	
	4th	28 (23)	17.40 (15)	20.83 (17)	
	5th	21 (18)	16.63 (14)	20.24 (17)	
2	1st,2nd	10 (8)	9.17 (8)	11.69 (10)	
	1st,3rd	14 (12)	9.21 (8)	11.23 (9)	
	1st,4th	17 (14)	9.67 (8)	11.53 (10)	
	1st,5th	8 (7)	8.92 (7)	10.72 (9)	
	2nd,3rd	14 (12)	9.26 (7)	11.29 (9)	
	2nd,4th	14 (12)	9.29 (8)	11.83 (10)	
	2nd,5th	10 (8)	8.76 (7)	11.23 (9)	
	3rd,4th	19 (16)	9.25 (8)	11.39 (10)	
	3rd,5th	16 (13)	8.71 (7)	10.87 (9)	
	4th,5th	13 (11)	9.20 (8)	10.83 (9)	
3	1st,2nd,3rd	4 (3)	3.39 (3)	3.86 (3)	
	1st,2nd,4th	7 (6)	3.54 (3)	4.71 (4)	
	1st,2nd,5th	2 (2)	2.92 (2)	3.97 (3)	
	1st,3rd,4th	9 (8)	3.39 (3)	4.35 (4)	
	1st,3rd,5th	4 (3)	3.04 (3)	3.86 (3)	
	1st,4th,5th	4 (3)	3.42 (3)	3.62 (3)	
	2nd,3rd,4th	6 (5)	3.23 (3)	4.19 (4)	
	2nd,3rd,5th	6 (5)	3.17 (3)	3.93 (3)	
	2nd,4th,5th	4 (3)	3.23 (3)	4.13 (3)	
	3rd,4th,5th	7 (6)	3.16 (3)	3.77 (3)	

Table 9
GARP(e) violations

e	GARP(e) (%)	Random money choice (%)	Random good choice (%)
1	35 (29)	26.43 (22)	69.91 (58)
0.99	30 (25)	20.22 (17)	59.47 (50)
0.98	27 (23)	15.43 (13)	49.90 (42)
0.97	24 (20)	11.20 (9)	41.18 (34)
0.96	21 (18)	7.88 (7)	33.82 (28)
0.95	18 (15)	5.68 (5)	27.62 (23)
0.94	14 (12)	3.95 (3)	22.34 (19)
0.93	13 (11)	2.91 (2)	17.82 (15)
0.92	10 (8)	2.08 (2)	14.31 (12)
0.91	7 (6)	1.38 (1)	11.34 (9)
0.90	6 (5)	1.04 (1)	8.96 (7)

Table 10
Logit regressions on garp

Variable	Experiment 1	Experiment 2
Age	-0.005 (-0.174)	-0.037 (-1.459)
Gender	-0.635 (-0.764)	1.189 (1.785)
Size household	0.705 (1.571)	0.149 (0.577)
Income	0.009 (0.174)	0.037 (0.700)
Budget level	-0.030 (-0.283)	0.002 (0.028)
Calculation time 1st situation	0.003 (1.182)	0.002 (0.792)
Calculation time 2nd situation	-0.015 (-2.567)	-0.002 (-0.404)
Calculation time 3rd situation	-0.007 (-1.600)	-0.006 (-0.886)
Calculation time 4th situation	0.021 (2.228)	0.007 (0.960)
Calculation time 5th situation	-0.003 (-0.715)	-0.007 (-0.873)
Constant	0.383 (0.232)	0.769 (0.552)

Notes: t-value in parentheses; estimates of “income” and “budget level” multiplied by 1000 and 10 respectively.

Table 11

Logit regressions on $gdexp_1$,..., $gdexp_5$, and $gexp_1$,..., $gexp_5$

Variable	$gdexp_1$	$gdexp_2$	$gdexp_3$	$gdexp_4$
Calculation time 1st situation	0.004 (1.708)			
Calculation time 2nd situation		-0.005 (-1.744)		
Calculation time 3rd situation			0.001 (0.413)	
Calculation time 4th situation				0.002 (0.636)
Calculation time 5th situation				
Constant	0.377 (0.369)	2.652 (4.296)	1.226 (2.766)	1.099 (2.219)
	$gexp_1$	$gexp_2$	$gexp_3$	$gexp_4$
Calculation time 1st situation	0.004 (1.708)			
Calculation time 2nd situation		-0.005 (-1.845)		
Calculation time 3rd situation			-0.003 (-1.566)	
Calculation time 4th situation				0.003 (0.995)
Calculation time 5th situation				
Constant	0.377 (0.369)	2.625 (4.329)	1.762 (4.282)	0.796 (1.657)

Note: t-value in parentheses.