Università Bocconi – Istituto di Economia Politica – Studi e quaderni TPE 2004-2

Early Experiments in Consumer Demand Theory: 1930-1970

Ivan Moscati*

IEP, Università Bocconi, Milano

ivan.moscati@uni-bocconi.it

Abstract

This paper reconstructs the history of experimental research on consumer demand behavior between 1930 and 1970. The backgrounds of the experiments and their impact on the development of consumption theory are also investigated. Among other things, the paper shows that in fact many prominent economists of the period were involved in this stream of research.

Keywords: Consumer Theory, Demand Theory, Experimental economics, Indifference curves, Transitivity of preferences

JEL Codes: B21, B31, C91, D12

1. Introduction

During the second half of the last century, experimental economics gained momentum as a significant field of economic inquiry, and the Nobel Prize in Economics awarded to Daniel Kahneman and Vernon L. Smith in 2002 has sanctioned the full recognition of the experimentalists' work by the economics profession. Some parts of the history of experimental research have been told by the protagonists themselves and by a number of historians of economics[1]. These accounts focus on the experiments performed in game theory, on the tests about individual behavior under uncertainty, or on the laboratory replications of different market structures.

Another important stream of experimental research has received less attention. This concerns individual choices from certain alternatives and encompasses the experiments performed to derive the indifference curves of a subject, to test the assumption that the individual preferences are transitive, or to check whether the axioms of revealed preference are satisfied by actual behavior. These are all matters that in neoclassical economics traditionally pertain to the analysis of consumer demand, such that these experiments are usually referred to as experiments in consumer demand theory. Historical accounts of this branch of experimental economics are short on details or incomplete[2]. The main goal of the present work is therefore to provide a comprehensive reconstruction of the experimental research made in consumption theory in the early period 1930-70.

In 1930 the American psychologist Louis Leon Thurstone performed an experiment to derive empirically the indifference curves of an individual. According to all the available evidence, Thurstone's experiment was the first one conducted on issues pertaining to the neoclassical theory of consumer behavior, making 1930 the natural starting point of the present story. As regards the closing date, before 1970 a limited number of experiments in consumer choice were performed, whose main topics of investigation were the derivation of indifference curves and the testing of preference transitivity. These experimental studies were interwoven, with respect to both their topics and the individuals involved in them. After 1970, experiments on different issues were conducted and new actors entered the scene, making that date a plausible limit for the current narrative.

A distinctive feature of the present reconstruction is that it explores the experimental research on choice behavior in connection with the 1930-70 developments in consumer demand theory. The experiments under consideration were in fact motivated by theoretical questions and can be assessed in a proper historical way only against the background of the contemporary theoretical debate. Moreover, many of the main demand theorists of

the period – H. Schultz, R. Frisch, H. Hotelling, N. Georgescu-Roegen, M. Friedman, J. Marschak, H.S. Houthakker, L. Hurwicz and J.S. Chipman,, among others – were well aware of the ongoing experimental research in the field, and some also contributed to it in more or less direct ways.

In the opposite direction, the paper also investigates the impact of experimental research on the development of consumer demand theory. This inquiry is carried out by tracking references to each experimental study for the twenty years following its publication, primarily in the journals of economics, business and finance listed in JSTOR, but also in other journals and relevant books. It turns out that in the period under consideration the influence of experimental research on the theory of consumer behavior was indeed very limited. Yet, the present study shows that, at least in the United States, this was not due to economists' ignorance of the experimental work.

A final hint about the general meaning in which the term 'experiment' is used in this paper may not be superfluous. The studies examined here are experiments in the sense that the data to derive the indifference curves or to test the transitivity assumption were generated in a controlled environment, that is, in "the laboratory" (which usually and less rhetorically was a classroom). From this viewpoint, the experimental approach is distinct from the statistical or econometric approach. The latter employs real market data, which generally have been collected in a previous moment and with autonomous procedures by an external agency. Therefore, the stream of research studied in this paper differs from, for example, the statistical estimation of demand curves performed in the 1910s by H.L. Moore or later by his student H. Schultz. Similarly, G. Stigler's analysis of English family-budget data for the period 1889-1904 and his claim that these data do not support the existence of Giffen goods, belonged to the statistical rather than to the experimental approach to consumer demand theory (cf. Stigler 1947).

The structure of the paper is as follows. Sections 2-5 are devoted to Thurstone's experiment on indifference curves, its background, and its reception among economists of the period. Sections 6-7 concern the second experiment on indifference curves, performed at the beginning of the 1950s when experimental research in economics began to rise. Sections 8-10 discuss the experiments testing the transitivity hypothesis on preferences, and their influence on the debate regarding the status of this hypothesis. Sections 11-12 outline the 1950s-1960s developments in consumer theory and experimental research, and discuss the last experiment on the indifference curves, whose results were published in 1969. Section 13 concludes.

2. Measuring utility: 1871-1930

Consumer demand theory originated in the marginalist value theories of C. Menger, W.S. Jevons and L. Walras, and became an autonomous body of doctrine with A. Marshall's *Principles*. According to the marginalists, the economic value of a commodity depends on the evaluation the subjects give to its marginal units. This evaluation was called "final degree of utility" by Jevons (1871), who first wrote down total utility as a cardinal function $u(x)$ of the quantity of the good, and marginal utility as the function obtained by differentiating total utility. However, it was not at all clear what the values that the function $u(x)$ associates with the commodity quantities x mean, how these values could be measured, and, above all, what the unit of measure of utility is.

To the question, What is a unit of utility? F.Y. Edgeworth answered that it is "the just perceivable increment of pleasure, the minimum sensible" (Edgeworth 1881: 99). In his *Mathematical Investigations*, I. Fisher criticized this psychological conception of utility and proposed to shift the analysis away from the individual's sensations or motivations, towards the decisions humans actually make. If the individual chooses x_1 over x_2, this means that the utility of x_1 is greater than the utility of x_2. Whether the ground for such a choice "is 'pleasure' or [...] 'duty' or 'fear' concerns a phenomenon in the second remove from the economic act of choice and is completely within the realm of psychology" (Fisher [1892] 1925: 11). However, Fisher's method for measuring utility based on economic choices only works under the restrictive assumption that the utilities of the commodities are independent. Therefore, it did not solve the general problem of having an objective utility measure[3].

Fisher's proposals were taken up by V. Pareto, who after 1900 considered it possible to construct economic theory just by examining consumers' choices. The analytical tool Pareto brought into play in order to obtain this result was Edgeworth's indifference curve. For Edgeworth the bundles on the same indifference line were those that gave the same utility to the subject, whereas for Pareto they were simply those the subject was unable to choose from, and in this sense, was "indifferent." Pareto, in fact, imagined putting a subject – a sort of Buridan's ass – between two bundles x_1 and x_2, then changing the composition of x_2 up to the point where the subject is no longer able to choose between x_1 and x_2, thus determining that x_2 belongs to the same indifference line as x_1 (cf. Pareto 1900: 217, and Pareto 1906: 164-65).

However, the experiment through which Pareto expected to determine the indifference lines is a fictitious one: there is actually no subject and no bundle. Pareto's test is in fact

just a thought experiment. The first actual experimental attempt to determine the indifference curves of an individual was carried out in 1930 by L.L. Thurstone. The results appeared in a paper published in the *Journal of Social Psychology* a year later (Thurstone 1931).

3. Thurstone's experiment

Thurstone's main concern throughout his career was the measurement of mental attitudes and abilities[4]. He was an advocate of "psychology as a quantitative rational science" based on description and experimentation (Thurstone 1959: 3 ff.), and he contributed to the founding of the Psychometric Society. From 1924 on he was professor at the University of Chicago, where he established the Psychometric Laboratory. In 1952, Thurstone moved his Psychometric Laboratory to the University of North Carolina at Chapel Hill, where he spent the rest of his academic career.

A colleague of Thurstone's at Chicago was H. Schultz, who also had a special concern for the quantitative estimation of economic variables, especially demand. Besides, Schultz was one of the major experts on and an admirer of Pareto, as well as an early member of the Econometric Society that was founded in 1930 by Fisher, R. Frisch and others. As Thurstone recalled, it was Schultz who suggested that he apply the experimental methods used in the psychological research to estimate an individual's indifference curve:

> The formulation of this problem is due to numerous conversations about psychophysics with my friend Professor Henry Schultz [...]. It was at his suggestion that experimental methods were applied to this problem in economic theory (Thurstone 1931: 139).

At that time, Schultz appears to have been Thurstone's only source of knowledge of economic theory. In fact, at the beginning of his paper Thurstone cautiously attributed to Schultz the opinion that the indifference curves had never before been subjected to experimental study, and pinpointed that "the writer dares not venture far into economic theory" (Thurstone 1931: 139). Apart from Schultz, Thurstone cited no other economist in his article, and the only explicit hint to the economic literature was the listing of Fisher's *Mathematical Investigations* in the references of the paper. I found no reference to economics in the works Thurstone had published before 1931 (nor, indeed, in those published afterwards).

Unfortunately, the question concerning Thurstone's knowledge of economic theory

can not be illuminated by his unpublished papers or correspondence, since these have not been collected[5]. In this state of affairs, the opinion of Lyle Jones, first a postdoctoral student of Thurstone's, later a collaborator and coauthor of his, and then for many years the director of the Psychometric Laboratory at Chapel Hill, is particularly worth quoting:

> My guess is that the 1930 experiment developed from discussions between Schultz and Thurstone at the Quadrangle Club, the faculty club on the campus at the University of Chicago. Thurstone often referred to the importance he placed on his luncheon talks there with faculty friends from other disciplines, and how important some of them were to his research activities. [...] It would come as a complete surprise to me were I to learn that he [Thurstone] had knowledge of economic theory at that time, beyond what may have been introduced to him by Schultz (e-mail to the author, November 3, 2004).

The supposition that Thurstone was not acquainted with economic theory fits well with the circumstance that Thurstone's experiment basically remained a psychometric one. Thurstone's purposes were in fact only in part similar to the anti-psychological ones of Fisher or Pareto. On the one hand, as a scientist with an empiricist concern, Thurstone sought to provide an experimental basis for the concept of the indifference curve. On the other hand, as a psychologist he was not interested in abandoning the mental variables "utility" and "marginal utility," for which he even used the psychologically-tinged expressions "satisfaction" and "motivation." Let us now examine Thurstone's experiment more in detail.

Thurstone considered three different commodities (hats, shoes and overcoats) and assumed that the satisfaction associated with each commodity was positive (in economic terms, $u(x_i) > 0$ for each good i). He also postulated that the motivation to acquire each commodity was positive but decreasing ($u'(x_i) > 0$ and $u''(x_i) < 0$). Moreover, he posited that the motivation conformed to Fechner's law, which stipulates that the satisfaction associated with a good is proportional to the logarithm of its quantity ($u(x_i) = k_i \log(x_i)$, where k_i is a constant characterizing each commodity i). In effect, this last postulate was an ex post rationalization of the experimental data, in the sense that Thurstone noted that Fecner's logarithmic law fitted his findings better than any other psychological formula. Fechner's law implies that the indifference curves are hyperbolas asymptotic to the axes, with the equation $\bar{u} = k_1 \log(x_1) + k_2 \log(x_2)$, where $\bar{u}$ is a fixed level of satisfaction or utility.

In the experiment, a female subject was presented with a bundle containing a certain number of commodities, e.g. 8 hats and 8 pairs of shoes[6]. Each bundle was plotted on a

Cartesian plane with the commodity quantities on the axes. Subsequently, in random order the subject was presented with other combinations (more than 200) of the same commodities, and asked to choose the preferred one. In this way, Thurstone built up two fields in the plane: one containing the bundles preferred to the initial one (these bundles were labeled with the sign plus), and the other field containing the bundles to which the initial bundle was preferred (these were labeled with the sign minus). Finally, Thurstone drew the indifference curve so that, as far as possible, all the preferred bundles lay above the curve and all the less preferred bundles lay below it. The as-far-as-possible clause means that in fact Thurstone discounted some anomalous preferences in order to draw a regular curve through the plus and the minus fields. According to him, the regular curve best fitting the recorded preferences was the hyperbolic curve $\bar{u} = k_1 \log(x_1) + k_2 \log(x_2)$.

Following the described procedure, Thurstone constructed four indifference hyperbolic curves for hats and shoes. Similarly, he constructed four curves for hats and overcoats. By using the method of averages, he estimated from the data the values of the parameters k_i for hats, shoes and overcoats. Lastly, Thurstone checked the values of the parameters derived from the first two sets of comparisons (hats-shoes and hats-overcoats) by means of the third set of comparisons (shoes-overcoats). He constructed four indifference curves for overcoats and shoes, and found that the agreement between the parameters predicted from the first two sets of comparisons and the parameters empirically derived from the third set of comparisons were quite satisfactory. Thurstone concluded that

> It is possible to reduce the indifference function to experimental treatment and that it is possible to write a rational equation for the indifference function which is based on plausible psychological postulates (Thurstone 1931: 165).

Some comments concerning the relationship between Pareto's thought experiment and Thurstone's are here in order. First, with respect to Pareto's mental exercise, in Thurstone's experiment there is indeed a real subject, but still no bundles: the subjects choose only imaginary objects. Second, the method for identifying the indifference locus is different. Pareto had imagined finding out directly the indifferent bundles by adjusting their composition; Thurstone instead looked for preferred and less preferred bundles, and drew the indifference curve as the frontier between the two groups. Lastly, Thurstone was not concerned with the Paretian issues regarding the ordinal interpretation of utility, and assumed quite straightforwardly that satisfaction and motivation were quantitatively measurable.

6

4. Consumer theory and the reception of Thurstone's paper in the 1930s

In order to appraise the reception of Thurstone's experiment it is convenient to outline the main developments of consumer theory during the 1930s. In that period, the starting point was Pareto's *Manual of Political Economy* (1906). In this work two approaches to demand analysis can be found: the behaviorist approach, according to which consumer theory should be based only on the observable choice behavior; and the ordinalist approach, backing the idea that consumption theory can be based on utility as a ranking index. However, as observed by many authors from the 1930s on, Pareto's theory was defective on both lines of attack[7].

In their celebrated 1934 paper, J.R. Hicks and R.G.D. Allen built up Pareto's analysis adopting a behaviorist approach. They replaced utility with the marginal rate of substitution and decomposed the effect of a price change on demand into income and substitution effects (Hicks-Allen 1934). Schultz (1935) and Allen (1936) rediscovered the paper in which E. Slutsky (1915) had anticipated the Hicks-Allen results, though within a cardinal-utility framework[8]. In particular, Allen (1936) pointed out that Slutsky's results could be obtained entirely within an ordinalist framework. This paved the way for the ordinalist restatement of Slutsky's findings, and for the formulation of the standard ordinalist consumer theory which was worked out by Hicks in *Value and Capital* (1939). The establishment of the Pareto-Slutsky-Allen-Hicks (PSAH) ordinal theory as the new orthodoxy was somehow contrasted by Samuelson's revealed preference theory. Samuelson (1938) showed that almost all the empirical restrictions on demand that derive from the constrained maximization of an ordinal utility function can also be obtained starting from a simple coherence assumption about consumer behavior. This assumption is what was later termed the Weak Axiom of Revealed Preference.

To return to Thurstone, despite the fact that his paper was published in a psychological journal, it was familiar to a number of prominent American demand theorists and to some European demand theorists working in the United States in that period. However, references to it were quite cursory and mainly critical.

In June 1932, Thurstone presented his experimental study at the meeting of the Econometric Society held in Syracuse, New York, also attended by Frisch and H. Hotelling. The Norwegian Frisch, who was in the United States conducting research, had explored the possibility of measuring the marginal utility of money using statistical demand data in two innovative works (Frisch 1926 and 1932). Hotelling had just wrapped up his important paper on the nature of supply and demand functions (Hotelling 1932). In the

discussion following Thurstone's presentation, the applied economist and statistician M. Ezekiel pointed out critically the difference between the answers the individuals give to questions or questionnaires and their actual behavior. Another participant in the session, the economist and methodologist J. Mayer, called into question Thurstone's assumption regarding the cardinal measurability of utility (cf. Mayer 1933: 97-8). In his 1936 paper on the theory of consumer behavior, the Romanian N. Georgescu-Roegen (at Harvard between 1934 and 1936) cited Thurstone's experiment in a footnote, but only to observe that it could not "be relevant to a theory concerned with an actual choice" (Georgescu-Roegen 1936: 585, n. 3). Basically, Georgescu-Roegen also criticized the hypothetical nature of the subject's decisions in Thurstone's study[9]. Hotelling cited Thurstone's experiment just once, writing only that the psychologist "succeeded in mapping out in a tentative manner the indifference loci" (Hotelling 1938: 248). Schultz himself recalled Thurstone's experiment only casually in two footnotes (Schultz 1931: 78, n. 5, and Schultz 1938: 15, n. 18). In the writings of Samuelson, Allen and Hicks, I found no reference to Thurstone's experiment.

In conclusion, Thurstone's paper had a negligible impact on the 1930s developments of demand theory both in the United States and in Europe. At least in the United States, this does not seem to be due to ignorance of his work. More crucial is the fact that Thurstone's experiment was somehow at odds both with the ordinalist tenets and the behaviorist ideas of the period. From an ordinalist viewpoint, the experiment was undermined by Thurstone's cardinal conception of utility. For the behaviorist, the experiment simply did not bear on the actual choices of individuals.

5. The Wallis-Friedman critique of Thurstone's experiment

Thurstone's experiment is not cited in the economic literature until 1942. In that year, W. Allen Wallis and Milton Friedman discussed it critically in a paper questioning the empirical usefulness of the standard PSAH demand theory based on the indifference curves. Their article appeared in the important volume of essays published in memory of Schultz, who had died in a car accident in 1938 (Wallis–Friedman 1942)[10]. Wallis and Friedman had both studied at Chicago and their work in fact belongs to a series of contributions in which Chicago scholars called into question the empirical relevance of the PSAH theory. In order to assess the Wallis-Friedman paper, an outline of its Chicago background is appropriate.

Wallis studied psychology as an undergraduate at the University of Minnesota and

then graduated in economics at the University of Chicago. Friedman took a masters degree in economics at Chicago in 1932-33 and wrote his MA thesis under Schultz's supervision. In 1934-35 Friedman returned to Chicago as Schultz's research assistant, helping his mentor in the completion of a paper on demand and of the book *The Theory and Measurement of Demand* (Schultz 1935 and 1938)[11]. In these two works, Schultz applied the PSAH theory to the analysis of statistical data regarding the demand for beef, pork and mutton. One empirical implication of the theory is that for an individual the compensated variation of the demand for a good when the price of another commodity varies (that is, the cross-price substitution effect) is symmetrical. Schultz found instead that the cross-price substitution effects were asymmetrical in magnitude and, for pork and mutton, even different in sign. Schultz was aware that the symmetry of the individual cross-price substitution effect generally does not carry over the aggregate demand, and hence that his findings represented no falsification of the PSAH theory (Schultz 1935: 474-75, and Schultz 1938: 630-31). However, such findings certainly did not corroborate the PSAH framework and indeed cast doubt on its usefulness for the analysis of statistical data[12].

In a paper published in 1939, G. Stigler, who had received his PhD at Chicago in 1938 and was a close friend of Wallis's and Friedman's, stressed more directly the difficulties in squaring the statistical demand curves with those of the PSAH theory. On the one hand, Stigler criticized the PSAH demand curves because they are so general as to become generic and useless. On the other hand, they appeared too narrow to Stigler, in the sense that they exclude variables of great importance in empirical demand analysis, like the changing value of money, the modifying effect of time, and the fact that the preference scale of an individual depends on his income (Stigler 1939: 471 ff.).

Wallis and Friedman elaborated on Schultz's statistical findings and Stigler's criticisms. They claimed that it is not possible to give a quantitative content to the PSAH indifference curves, neither by means of an experimental approach nor through statistical techniques. The difficulties are basically due to the fact that the PSAH indifference analysis separates the "taste factors" (preferences) from the "opportunity factors" (prices, income) while they are inextricably interwoven (cf. Wallis-Friedman 1942: 188). The impossibility of giving quantitative content to the indifference curves would render the PSAH apparatus useless for practical purposes. Wallis and Friedman therefore supported an approach to consumption analysis that isolates the factors correlative with consumer demand (income, wealth, prices, family type, occupation, age, nationality, etc.) and de-

termines statistically their specific influence. The Wallis-Friedman discussion of Thurstone's experiment fitted in this wide-ranging confrontation with the PSAH approach.

First, Wallis and Friedman pointed out some flaws in Thurstone's statistical techniques that mar his estimation of the k_i parameters. More basically, and similarly to Ezekiel and Georgescu-Roegen, they criticized Thurstone's experiment for the hypothetical nature of the subject's choices:

> For a satisfactory experiment it is essential that the subject give actual reactions to actual stimuli [...]. Questionnaires or other devices based on conjectural responses to hypothetical stimuli do not satisfy this requirement (Wallis-Friedman 1942: 179-80).

The conjectural nature of Thurstone's test, however, does not preclude the possibility of experiments with real stimuli. To discuss this point, Wallis and Friedman imagined a different experiment, in which a child can choose among different combinations of candy and ice cream, but has to consume the chosen one. If different combinations of candy and ice cream are offered to the child, "it would clearly be necessary to offer the same combinations day after day, to give the subject an opportunity to experiment" (Wallis-Friedman 1942: 180). The attractiveness of variety can lead to alterations in the pattern of choices, so that it may be necessary to wait some time between offering different combinations. At the end of this process, only a single observation would be obtained, and it would be necessary to repeat the procedure in order to get additional data for the estimation of the child's indifference curve. Wallis and Friedman argued that in the course of such a prolonged process of data recording, the individual preferences would change so that the entire estimating procedure would fail[13]. The conclusion of their discussion regarding the experimental derivation of the indifference curves is skeptical:

> These are more than technical or practical obstacles and indicate that it is probably not possible to design a satisfactory experiment for deriving indifference curves from economic stimuli (Wallis-Friedman 1942: 181).

This negative statement is somehow in contrast with the circumstance that, as we will see below, around 1950 Wallis and Friedman contributed as supportive discussants to an important experiment on decision making under uncertainty. Moreover, in 1952 Friedman was further involved in the experimental research as a discussant of the two first experiments on transitivity. About this apparent two-sided attitude towards the experimental research, Friedman notes:

I am sure my opinion of experimental research for economics in the 1950s was a mixture of skepticism and optimism. As the article with Allen Wallis showed, we thought the existing attempts at experiments were not very well done and that they would not be easy to carry out. On the other hand, I doubt very much that I was a complete opponent of experimental research. If we had been, we would not have written the article. So I suspect but cannot be sure that the attitude was what it has been all along: open to persuasion but somewhat of a skeptic (e-mail to the author, November 11, 2004).

The paper of Wallis and Friedman did not prevent the PSAH approach from remaining the standard one in neoclassical consumption analysis, but was influential as an early methodological discussion on the factors that can undermine the value of an experiment. The Wallis-Friedman appraisal of Thurstone's experiment had pointed out the importance of using real stimuli that can actually induce the subject to reveal her preference, and the necessity of evaluating carefully all the factors that can modify the experimental situation (the elapse of time, modifications of the external environment, preference for variety and other effects induced by the experimental procedure itself). As we will see below, in this methodological guise the Wallis-Friedman paper became a reference point for some of the economists performing experiments on individual choice during the 1950s and 1960s.

6. Experimental research in the late 1940s and early 1950s

Whereas until the end of the 1940s Thurstone's experiment remained an isolated exception in economics, the late 1940s and early 1950s witnessed an increasing amount of experimental research[14]. E.H. Chamberlin (1948) explored the performance of a market by reproducing it in the classroom. He induced demand and supply among his students at Harvard by assigning to each of them a reservation price as a buyer or a seller of a commodity, and found that the average price and the quantities exchanged were different from those predicted by perfect competition theory. Still, Chamberlin's market experiment had no follow-up until the late 1950s.

The publication of *Theory of Games and Economic Behavior* (von Neumann-Morgenstern 1944) stimulated the two most flourishing streams of experimental research of the late 1940s and the early 1950s. The first stream concerned interactive behavior. The early experiments on gaming and bargaining were performed at the RAND Corporation in Santa Monica, California, by M.M. Flood, a former student of J. von Neumann's. In particular, Flood and his collaborator M. Dresher constructed and studied experimentally the game that subsequently came to be known as the prisoner's dilemma (cf. Flood

1951-52, 1952a, 1952b).

The other stream of research concerned the individual choice under uncertainty and the testing of expected utility (EU) theory. Between 1948 and 1949 the statistician F. Mosteller (PhD in mathematics from Princeton, 1946) and the psychologist P. Nogee conducted a series of experiments to estimate the utility functions of different individuals on the basis of their choices among simple gambles. By using such utility functions, Mosteller and Nogee tested the predictions of the EU theory on more complicated gambles and the inflection-point hypothesis in the utility curve suggested by Friedman and L.J. Savage (cf. Friedman-Savage 1948: 293 ff.). Mosteller and Nogee concluded that the EU predictions "are not so good as might be hoped, but their general direction is correct" (Mosteller-Nogee 1951: 399). For our narrative it is worth pointing out that, as anticipated in the previous Section, the plans for the Mosteller-Nogee experiment "grew out directly of discussions with Friedman and Savage" and that "W. Allen Wallis also contributed to the discussions" (Mosteller-Nogee 1951: 372, n. 7).

In this very early phase of experimental research, the most important event was probably the eight-week seminar on "The Design of Experiments in Decision Processes" that was held in the summer of 1952 in Santa Monica. The forty or so participants were mathematicians, psychologists and economists, and among them were von Neumann, O. Morgenstern and J. Nash (Princeton), Marschak and T. Koopmans (respectively, the former and the then current director of the Cowles Commission, Chicago), Flood (RAND), Mosteller (Harvard) and H. Simon (Carnegie Institute of Technology). In fact, the conference proved to be focused mainly on theoretical themes related to the application of mathematics to the social and behavioral sciences, rather than on the design of experiments. The volume collecting the conference papers and edited in 1954 by R.M. Thrall, C.H. Coombs and R.L. Davis contained in fact just five experimental studies[15]. According to Vernon Smith, the success of the conference in stimulating experimental research "must be judged much more modest than in furthering theory"; despite this, its influence "on subsequent experimentation, and upon fundamental concepts, [...] was very significant and lasting" (Smith 1992: 261).

7. The Rousseas-Hart composite indifference map

At the beginning of the 1950s the second experimental study in consumer theory also took place. It was performed by Stephen W. Rousseas and Albert G. Hart of Columbia University, who attempted to empirically derive indifference curves. Rousseas and Hart

devised their experiment as an outgrowth of Thurstone's and as a response to the Wallis-Friedman critique. Their paper appeared in the *Journal of Political Economy* in 1951.

Hart was also a scholar with a Chicago background. He had been a graduate student along with Friedman and Stigler at the University of Chicago and completed his PhD there in 1936[16]. In 1942 he contributed with a paper on uncertainty to the volume in memory of Schultz, where the Wallis-Friedman articled also appeared (Hart 1942). From 1946 Hart was professor of economics at Columbia. As regards the first-listed author of the paper, in 1951 the Greek-American Rousseas was a PhD student at Columbia and was preparing a dissertation on the logical and empirical foundations of the subjective theory of value[17]. The 1951 paper was the only contribution of Rousseas and Hart to experimental research and afterwards they both moved to the study of macroeconomic themes.

The basic aim of Rousseas and Hart was to show that the PSAH theory of consumer demand was not a framework of purely analytical propositions but had instead an effective empirical basis. They noted that denials of the empirical content of the PSHA had come from very different quarters, e.g. the Marxist economist Maurice Dobb, the eclectic John Maurice Clark (at that time at Columbia), and from Wallis and Friedman. Rousseas and Hart were mainly concerned with the Wallis-Friedman attack and aimed to rebut it by showing that the experimental procedure was indeed suited for empirically identifying the indifference curves.

Rousseas and Hart (1951: 290) judged that the results of Thurstone's experiment were "quite poor". They deemed that this was mainly due to Thurstone's method of eliciting the indifference curves by asking a single subject to evaluate too large a number of commodity combinations: "This inquiry would be more likely to generate a nervous breakdown than a body of valid evidence" (Rousseas-Hart 1951: 290). Moreover, as Wallis and Friedman had pointed out, actual tests on a single individual would have taken so long a time that his preferences would have changed in the meanwhile. Rousseas and Hart claimed that both these problems could be circumvented by asking a group of individuals with homogeneous tastes to make few, very simple choices within a short period of time. The crucial hypothesis of homogeneity of tastes was posited by Rousseas and Hart as a preliminary assumption to be tested in the course of the experiment.

The experiment involved a class of graduate sociology students at Columbia University, probably in 1950[18]. It consisted of two runs at an interval of a month. In the first run 67 students participated, whereas in the second 54 students of the initial group partici-

pated, plus 10 new students. Each student received a card listing three different combinations of scrambled eggs and bacon strips, was asked to rank them, and then apparently had to eat the top ranked combination.

Actually, Rousseas and Hart were very vague on this last point and wrote only that a "further condition was that each individual was obliged to eat all of what he chose – i.e. he could not save any part of the offerings for a future time" (Rousseas-Hart 1951: 291). It is not clear whether this "further condition" was just a theoretical constraint the students had to bear in mind while ranking the combinations, or if the experiment contemplated a final consumption phase. Since Rousseas and Hart analyzed only the rankings indicated by the students on the cards, and their study contains neither a description of a consumption phase nor an evaluation of the possible problems related to it, I deem that in their experiment the subjects' choices were still just conjectural[19].

Based on the ranking of the combinations recorded by each individual, Rousseas and Hart tried to draw a piece of each person's indifference map. The procedure was in essence based on the idea that higher indifference curves pass through preferred combinations, and on the use of the vector connecting the second and the third choices as an approximation of an indifference curve. After this first step, Rousseas and Hart put together the pieces of each person's indifference map, and drew what they called a "composite indifference map". This composite map should be "roughly representative of the [...] indifference map of each individual" (Rousseas-Hart 1951: 318). Unlike Thurstone, Rousseas and Hart did not try to give mathematical expression to the indifference graph they had obtained experimentally.

The entire procedure was repeated in the second run of the experiment, which served to refine the observations made in the first one. According to Rousseas and Hart, the overall data confirmed the initial assumption of taste homogeneity. Moreover, since the experimental observations were collected in a short time and the individual preferences appeared to remain constant during this period, Rousseas and Hart claimed that the Wallis-Friedman objections to Thurstone's study "do not apply to this [i.e., their] experiment" (Rousseas-Hart 1951: 318). In conclusion, Rousseas and Hart contended to have shown that it is possible to get an empirical counterpart for the indifference curves and that, consequently, the propositions of the PSAH demand theory could not be charged with being merely analytical.

However, the Rousseas-Hart experiment has two main flaws that mar this conclusion. First, the hypothesis of taste homogeneity is actually highly questionable: some subjects

were presented with the same bacon-eggs combinations and their rankings proved to be divergent. Second, the way the individual indifference maps were plotted and then put together in order to draw a well-behaved composite map appears quite arbitrary. Probably because of these drawbacks, the Rousseas-Hart experiment seems to have had no influence on subsequent economic research. During the twenty years following its publication, in all the journals of economics, business and finance listed in JSTOR the Rousseas-Hart paper was cited only twice: the first instance in an occasional footnote by J. Rothenberg (1953: 258, n. 2), and the second one by K.R. MacCrimmon and M. Toda (1969: 435-36) in their experimental study on indifference curves.

8. May's experiment on the transitivity of preference patterns

Another byproduct of von Neumann and Morgenstern's book was the spread into economics of the treatment of preference as a formal relation, whose properties can be stated axiomatically. The axiomatic handling of preferences was adopted by Marschak (1950), employed by Houthakker in his important article on revealed preference (1950), and perfected by K. Arrow in *Social Choice and Individual Values* (1951). One of the key axioms on the preference relation is that it is transitive, which means that if x is preferred to y and y is preferred to z, then x must be preferred to z. During the 1950s, the testing of the transitivity assumption became a new line of investigation for experimental research.

The meeting of the Econometric Society held in Chicago at the end of December 1952 is probably better known in the history of economics because Arrow and G. Debreu as well as L. McKenzie there read their papers on the existence of general equilibrium (cf. Weintraub 1983). For our narrative that meeting is more significant because in the session on "Individual Preference Functions," the two first experimental studies on transitivity were presented: Kenneth O. May's and Andreas G. Papandreou's. The general discussants of the session were Houthakker, the experimental psychologist W. Edwards, and, as anticipated in Section 5, Friedman[20]. The results of the two experiments appeared antithetical. May claimed that his experimental findings indicated that "rational individuals do exhibit patterns that contain circularities [i.e. intransitive patterns]" (cf. May: 1953: 476). Papandreou reported instead that "no evidence has been revealed which contradicts [...] the transitivity axiom" (Papandreou 1953: 477). Actually, May's and Papandreou's findings were not at all incompatible. Let us see why, beginning with May's experiment.

May was a mathematician specialized in mathematical economics (PhD at Berkley

under G.C. Evans in 1946) and taught at Carleton College in Northfield, Minnesota. From 1946 to 1947 he also served as a research consultant for the Cowles Commission. In the late 1940s, May had become interested in the theory of majority decisions and was one of the first reviewers of Arrow's 1951 book (May 1952)[21]. The paper he presented in Chicago appeared eventually in *Econometrica* in 1954. Notably, in the opening footnote of the article May thanked Rousseas for suggestions and comments.

In the first part of the paper, May clarified the crucial role of the transitivity assumption in utility theory. The ordinalist-utilitarian representation of preferences initiated by Pareto states that higher utility numbers must be assigned to preferred alternatives. When transitivity does not hold, cyclical preference patterns can occur, in the sense that it can happen that *x* is preferred to *y*, *y* is preferred to *z* and *z* is preferred to *x*. According to the utilitarian representation of preferences, this implies that $u(x) > u(y) > u(z) > u(x)$, which is, however, a contradiction. May argued that this result could be used to test the legitimacy of using a utility function in the analysis of individual decision making:

> This result means that transitivity is necessary for utility. It makes it possible to test the hypothesis of the existence of a utility function by experimental tests of transitivity. Where transitivity does not hold, the use of utility is not justified (May 1954: 3-4).

May in fact claimed that circular patterns in individual preference do occur. He first recalled the well known paradox of majority voting. In this case, the intransitivity at the level of collective choice arises from the aggregation of transitive individual patterns. May's idea was to apply the paradox of voting to individual choice, and to check experimentally if intransitive preference patterns may arise when the individuals order the alternatives according to conflicting criteria.

His experiment was simple and pretty informal. The subjects were 62 college students, presumably at Carleton College (May provided no information regarding the precise venue and date of the experiment). The alternatives proposed to the students were three hypothetical marriage partners, *x*, *y*, and *z*, which were characterized in terms of intelligence, beauty, and wealth. In intelligence, *x* was described as superior to *y*, and *y* as superior to *z*; *y* was portrayed as more beautiful that *z*, and *z* as more beautiful than *x*; finally, the subjects were told that *z* was richer than *x*, who was in turn richer than *y*. The individuals were confronted at different times only with pairs of hypothetical marriage partners, i.e. never with all three alternatives at once, and had to choose the preferred one. The indifference option was ruled out.

16

The result of the experiment was that 17 of 62 students, i.e. more than 27 percent, exhibited intransitive preference patterns. In particular, they chose *x* over *y*, *y* over *z* and then *z* over *x*. May explained these cycles "as the result of choosing the alternative that is superior in two out of three criteria" (May 1954: 7), that is, as the outcome of the majority rule applied to the conflicting features characterizing the marriage partners.

Concerning the relevance of these results to economics, May pointed out that the commodity bundles considered by economic theory are in fact vectors containing different goods. If these goods are ranked according to conflicting criteria, circular preference patterns cannot be excluded. May's general conclusion was that his experiment suggested that individual preferences can easily be intransitive, so that:

> The question is no longer 'Are preferences transitive?' but rather 'Under what conditions does transitivity fail?' (May 1954: 8).

9. Papandreou's test on probabilistic transitivity

The other experimental study read at the Chicago meeting was Papandreou's. Outside economics, Papandreou is best known as the founder of the Greek socialist party (Pasok) and the former prime minister of Greece (1981-89 and 1993-96). In 1940 he left Greece for the United States for political reasons, and enrolled at Harvard, where he graduated with a PhD in economics in 1943. Until 1955 he taught at the University of Minnesota, and from 1955 until 1963 at Berkley. In that year, he went back to Greece and to political activity.

Apparently, Papandreou encountered some difficulties in publishing the experimental study presented at Chicago. In September 1954 he read a second version of it at the summer meeting of the Econometric Society held in Montreal, Canada. This second version was indicated as a joint work of a team composed of Papandreou and (in the following order) Owen H. Sauerlander, Oswald H. Brownlee, Leonid Hurwicz and William Franklin, all of the University of Minnesota (Papandreou and others 1955)[22]. The most prominent individual in this group was Hurwicz, who had been research associate at the Cowles Commission since 1942 and in the middle of the 1950s was working on linear and non-linear programming. The paper was published only in 1957 under the title "A Test of a Stochastic Theory of Choice" in the minor periodical *University of California Publications in Economics*. In this final version, the paper appeared as one of Papandreou's, with the collaboration of the previously listed individuals (Papandreou and others 1957).

The expression "stochastic" in the title of the paper does not mean that Papandreou studied risky choices, i.e. choices among alternatives with uncertain outcomes like bets or lotteries. Rather, the expression "stochastic" refers to a probabilistic statement of the theory of choice among sure outcomes, originally proposed by Georgescu-Roegen (1936: 568 ff.). Georgescu-Roegen's basic idea was to define preference in terms of frequency of choice rather than as an invariable relation, in the sense that an individual can be said to prefer x to y if, when confronted with these two alternatives more than once, she chooses x over y more frequently than y over x. More formally, x is said to be "stochastically" preferred to y, if the probability that the individual chooses x over y is greater than 0.5. This probabilistic or stochastic theory of choice seemed suited to analyse the not so uncommon situations where individuals make different choices when confronted with the same alternatives.

In the 1950s, a considerable stream of theoretical research was devoted to develop this probabilistic approach, to give it an axiomatic treatment, and to analyse the utilitarian representation of stochastic preferences[23]. The Papandreou experiment fits in this line of research and was set out to test the following stochastic version of the transitivity axiom: if the probability that an individual chooses x over y is greater than 0.5, and the probability that she chooses y over z is greater than 0.5, then the probability that she chooses x over z is greater than the probability that she chooses y over z (all the inequalities are meant to be weak). In others words, this axiom states that if x is stochastically preferred to y, and y is stochastically preferred to z, then it is more likely to choose x over z than to choose y over z[24].

The Papandreou experiment was conducted with students from an elementary sociology class and a graduate class in education. In all probability they were students from Minnesota University (Papandreou is not explicit regarding the venue and the date of the experiment). The subjects were confronted with ten hypothetical activities of recreational nature, like "admission to a college basketball game between two nationally ranked teams," or "admission to an evening at the Minneapolis symphony under the direction of a leading conductor." These ten alternatives were combined into all 45 possible pairs. For each pair, Papandreou constructed the triple of bundles containing four admissions to the activities in the pair. For example, for the pair (basketball game, Minneapolis symphony), the three bundles with a total number of four admissions were: x=(3 basketball, 1 symphony), y=(2 basketball, 2 symphony), and z=(1 basketball, 3 symphony).

Between these three bundles, three binary comparisons are possible, namely x vs. y, x

vs. *z*, and *y* vs. *z*. Papandreou's axiom requires that the choices resulting from these comparisons be transitive in the probabilistic sense described above. 18 students participated in the six sittings of the experiment, which took place over a period of two and a half weeks. In every sitting each student was asked to make $3 \times 45 = 135$ choices. The subjects were faced with such a large number of choices in order to reduce the influence of memory factors in inducing transitivity over the decisions.

Since the definition of the transitivity axiom is probabilistic, Papandreou's test was a statistical one. The binary choices were made over 45 triples and for 18 individuals, so that Papandreou's sample contained $45 \times 18 = 810$ observations. The null hypothesis to be tested was that the choices recorded in the experiment had been generated by stochastically transitive preferences. Unfortunately, since the statistical properties of the test were unknown, it was impossible to define precisely the appropriate critical region of rejection for the null hypothesis. In spite of this, Papandreou claimed that this flaw was not really relevant because the experimental evidence supported the transitivity assumption beyond any possible doubt:

> There are at least 704 cases [of a total of 810, that is 87 percent] in which the hypothesis [of transitivity] could not be rejected, and at most 101 cases [12 percent] in which it could be rejected. [...] The evidence provided by the data is [...] overwhelming by any criterion (Papandreou and others 1957: 7-8)[25].

Comparing the findings of May's and Papandreou's experiments, it comes out that they were not at all incompatible. Basically, the two scholars tested different axioms: May tested the standard, deterministic transitivity axiom, whereas Papandreou examined a probabilistic version of transitivity. Yet, both experiments suggested that intransitive patterns (deterministic as well as probabilistic) cannot be excluded from choice theory.

Before discussing the impact of these findings on the following developments of consumer theory, I briefly review the experiment the sociologist Arnold Rose performed on transitivity. Rose's study was in direct connection with May's and Papandreou's and its results were published in the *Journal of Political Economy* in 1957.

10. Rose's test and the impact of the transitivity experiments

Like Papandreou, Hurwicz and the other scholars who performed the experiment on stochastic transitivity, Rose taught at the University of Minnesota. Indeed, in the opening footnote of his article, Rose expressed his appreciation to Hurwicz "for aid in formulating the problem" (Rose 1957: 394, n. 1).

Rose had studied sociology (BA in 1938 and PhD in 1946) and economics (BA in 1939) at the University of Chicago. He was in Minneapolis beginning in 1948. His research interests ranged from the study of racial integration, to the analysis of social perception of crime, to questions related to sociological method. His experiment aimed to suggest that the intransitive behavior observed in common experience or recorded in experiments like May's and Papandreou's should mostly be considered as a case of "apparent intransitivity". By this term, Rose meant choice circularities that are not due to underlying intransitive preferences, but to disturbing factors that enter in the passage from subjective preferences to observable choice behavior.

Rose presented 74 students in a sociology course at the University of Minnesota with 13 possible types of crimes. In a preliminary phase, the students had to rank the 13 crimes according to their seriousness. This preliminary ranking was realized according to the so-called method of comparative judgments introduced by Thurstone (1927a, 1927b). Subsequently, each student was presented in random order with every possible pair of crimes and had to indicate the more serious of the two, without the option of judging them indifferent. The experiment was replicated two months later with the same subjects. As in the studies of May and Papandreou, the choices the subjects made were in effect hypothetical.

Rose recorded 341 intransitive patterns of choice in the first run of the experiment and 336 in the second run. Only three students made no circularities whatsoever. Rose then attempted to show that most of the recorded intransitivity was apparent. He first noted that more than 90 percent of the circularities made in the first run were different from those made in the second run, so they could not be attributed to underlying, consistently intransitive preferences. According to Rose, the cause of such circularities was in fact careless judgment. As for the remaining 10 percent of repeated circularities, they concerned crimes that the subjects had judged very close in the initial ranking. Rose claimed that in these cases the subjects were in effect indifferent about crimes and had selected one just because the instructions of the experiment told them to make a choice. Most likely, such choice was made in a random way and this was the source of the observed intransitivity. In sum, for Rose the intransitivity recorded was always apparent intransitivity:

> The evidence from this study is that true intransitivity in making choices [...] appears to be a very rare phenomenon, if it exists at all. Most of what appears 'irrationality' can plausibly be explained by [...] carelessness [...] or random error (Rose 1957: 401).

What was the impact of the experimental studies of May, Papandreou and Rose on the developments of consumer theory in the 1950s and in the following years? Basically, they were influential in raising the debate on the status of the transitivity assumption. This methodological question had already emerged in the discussion at the Chicago meeting. In commenting on the experiments of May and Papandreou, Edwards observed:

> Transitivity [...] can be considered an axiom to be used as the basis of a theory, or it can be regarded as an observable property of choices [...]. From the latter point of view, there is no reason to suppose that all choices will be transitive (Edwards 1953a: 477).

The debate on whether the transitive assumption is a descriptive hypothesis, a normative postulate of rational behavior, or an assumption necessary for constructing a systematic theory of choice and demand, began in the 1950s and lasts until today[27]. For the present study, it is important to point out that the experiments of May and Papandreou stimulated this debate, and that the discussion regarding transitivity always referred to at least one of them, especially in the 1950s but also later. In particular, May's experiment was mainly referred to as a clear indication that preferences can be intransitive[28]. Other authors, instead, attributed May's evidence against transitivity to the fact that he had excluded the indifference option[29]. Quite singularly, in his influential survey on the state of consumption theory, Houthakker (1961: 730) judged the results of May's experiment "encouraging" for the transitivity assumption. Scholars chiefly cited Papandreou's experiment in contributions on probabilistic choice theory to support the idea of a stochastic version of the transitivity axiom[30]. In particular Debreu (1958), Arrow (1959) and Chipman (1960) referred to May's and Papandreou's tests. Rose's experiment was far less influential than the other two and is seldom cited in the literature[31].

On the whole, however, the discussion of the transitivity axiom and the attempts to build a theory of choice based on some weaker version of it remained (and has remained until today) a secondary research subject in demand analysis. In the 1950s, the major contributions in the field were in fact devoted to clarifying the relationship between the behaviorist and the ordinalist approach to consumer demand, and to complete the axiomatic foundation of the PSAH theory.

11. Consumer theory and experimental economics in the 1950s and 1960s

In Samuelson's revealed preference theory, the only restriction on demand functions that derives from the utility maximization hypothesis but not from the Weak Axiom is the

symmetry of the cross-price substitution effects. This property is important since it is necessary for the existence of a utility function whose maximization generates the consumption choices expressed by a given demand function. The problem of defining under which conditions such a generating utility function exists, is the so-called "integrability problem"[32]. For our purposes, it suffices to say that in 1950 Houthakker introduced the Strong Axiom of Revealed Preference (an iteration of the Weak Axiom that rules out the possibility of cyclical choices) and showed that it is sufficient for the integrability of demand functions. On the basis of Houthakker's contribution, Samuelson (1950) proved that the PSAH framework and the Samuelson-Houthakker one led in effect to identical restrictions on demand. With respect to the ordinalist approach, Debreu (1959) carried out the axiomatic analysis of preferences and determined the assumptions that are sufficient to derive the continuous quasi-concave utility function required by the PSAH theory. In this way, Debreu provided the received ordinal demand theory with a sound axiomatic foundation and brought it to its current standard form.

From the 1950s on the pace of new contributions in the field of experimental research quickened[33]. Edwards (1953b, 1954b, 1954c), D. Davidson and P. Suppes (1957), Davidson and Marschak (1959), and D. Ellsberg (1961) conducted experiments on individual behavior under uncertainty and tested the EU theory. T. Schelling (1957), J. Stone (1958) and others after them studied gaming, bargaining and coordination problems. Chamberlin's idea of exploring the behavior of a market through a controlled experiment was put forward by A. Hoggatt (1959), S. Siegel and L. Fouraker (1960) and V. Smith (1962) in the United States, and by H. Sauermann and R. Selten (1959) in Germany[34].

In this period, also a group of experiments related to consumer choice were performed, and one of the scholars involved in them was, again, Thurstone. In a conference of the American Meat Institute held at the University of Chicago in 1952, he presented the results of an experimental study based on the preferences expressed by 250 subjects on forty foods. In the first phase of the experiment, the subjects were faced with one food at time and asked to rate it by means of a nine interval scale ranging from "extreme dislike" to "extreme like". The subjects were then given menus containing groups of the same foods, and were asked to rank the foods within each menu. Thurstone checked whether the number of subjects choosing a good from a certain menu could be predicted by the number of subjects that in the separate evaluations had rated the food better than the other foods on the menu. He obtained a high degree of correlation between the two types of ranking (cf. Thurstone [1952] 1959). In another paper written with L. Jones and

published in the *Journal of the American Statistical Association*, Thurstone attempted to locate experimentally the stimulus that gives zero utility to the subject (eliciting the zero utility point would be crucial to make statements about the ratio of utilities). However, Thurstone and Jones ascertained the zero point under two peculiar assumptions: *i)* that the utilities of different subjects could be compared, so that the subjective values at issue were in fact those of "the average subject" (Thurstone-Jones 1957: 458); *ii)* that the utility function of the average subject was additive, so that the utility of a combination of two goods could be considered equal to the sum of their separate utilities. In a paper published in the *Journal of Applied Psychology* in 1959, Jones built on Thurstone's 1952 experiment. He attempted to predict the menu choices of the members of the Chicago University faculty club by means of the preferences they had expressed on the single menu items in a previous questionnaire. His model of choice prediction was based on the aggregation of individual preferences, that is, on "group utility rather than individual utility" (Jones 1959: 337)[35].

In 1955 the *Journal of Applied Psychology* had published the results of another experiment on meal choices conducted by the sociologist P.H. Benson. Benson presented 263 students with a list of appetizers, a list of entrees, and a list of desserts. The students were asked to imagine themselves as invited guests and rank their preferences from 1 to 10 for appetizers, entrees, and desserts. By attributing cardinal meaning to such rank order, by assuming that the tastes of the students were homogeneous, and by supposing that the utilities of the goods were additive, Benson constructed a model of choice through which he predicted the average amount of money spent on appetizers, entrees, and desserts "with a mean error of six cents" (cf. Benson 1955: 381). In 1960 the marketing scholar E.A. Pessemier published in the *Journal of Business* an experimental study designed to elicit the demand schedules for 14 different branded goods belonging to four commodity classes (tooth paste, cigarettes, toilet soap and headache remedies). Yet, the demand schedules derived by Pessemier were not the standard ones. Even if their upper part was obtained in the expected way, that is by facing the subjects with increases in the price of the brand under consideration, their lower part was peculiarly elicited by reducing the prices of all but that brand (cf. Pessemier 1960 and also Pessemier 1963, where the results of an almost identical experiment are presented).

The experimental studies of Thurstone, Jones, Benson and Pessemier went unobserved in the economics literature[36]. In part, this seems to be due to the circumstance that these studies had not been published in economics journals. Yet, much more critical ap-

23

pears the fact that the assumptions behind them (homogeneity of tastes, cardinally measurable and additive utility) as well as the issues investigated (comparison between one-by-one and menu goods rankings, elicitation of peculiar demand schedules) were remote from the assumptions and the issue proper to the consumer demand theory of the period.

In hindsight, it is surprising that no experimental study was conducted on the axioms of revealed preference. Samuelson and Houthakker made clear that the response to changes in prices or income of a utility maximizer must fulfill the Strong Axiom. This result makes the testing of the Strong Axiom especially worthwhile, since it means that the utility maximization hypothesis has to be rejected if the Strong Axiom is violated. As Wallis and Friedman would have pointed out, the basic problem for such an experimental test is that the Strong Axiom restricts the individual's choices only if any modification of tastes, income effect, or other perturbing factors can be ruled out. Maybe because of this difficulty, in the period under consideration nobody tested the Strong Axiom. The first experiment on the revealed preference axioms was in fact conducted only in the 1990s by R. Sippel (1997).

12. Drawing the indifference curves: MacCrimmon's and Toda's experiment

The October 1969 issue of the *Review of Economic Studies* was almost entirely devoted to experimental studies. One of the six papers collected in the journal under the general title "Symposium on experimental economics" was that of the Canadian Kenneth R. MacCrimmon's and the Japanese Masanao Toda's. Like Thurstone, Rousseas, and Hart before them, also MacCrimmon and Toda elaborated a procedure to derive experimentally the indifference curves. In particular, they were interested in testing to what extent the curves satisfy the properties the theory commonly attributes to them, that is, non-intersection, negative slope, and convexity. MacCrimmon and Toda reviewed the studies of Thurstone, Rousseas, and Hart, and recognized the importance of the methodological issues brought up by Wallis and Friedman. Yet, they claimed to have solved the problems raised by the two Chicago scholars.

For our story it is important to note that MacCrimmon and Toda (1969: 433, n. 1) thanked Marschak for his helpful comments on their paper. Actually, Marschak had not only commented on their work but was at the very origin of their collaboration. McCrimmon had studied at the Western Management Science Institute at the University of California, Los Angeles, and had followed interdisciplinary BS and PhD programs comprehending economics, business, mathematics, and behavioral science. From the be-

ginning of the 1960s, he also collaborated with the RAND Corporation, working on operation research projects. In 1965 MacCrimmon finished his PhD under Marschak (at UCLA since 1960) with an experimental study on the decision making behavior of business executives. From 1964, MacCrimmon taught at Carnegie-Mellon University.

Toda's career brought him to Marschak by a different route[37]. After studies in psychics Toda moved to psychology, and in 1952 he graduated from the Graduate School of Psychology at Tokyo University. He then taught at Hokkaido University. In 1962 Toda visited the Psychometric Laboratory at Chapel Hill, at that time directed by Jones (Thurstone had died in 1955). In the same year, Toda published an often cited article in the psychological and cognitive literature that presented a behavioral model for a pseudo-human agent, the "Fungus-Eater Robot." In his paper, Toda thanked Marschak for helpful suggestions (Toda 1962: 164). In 1966-67 Toda was a visiting professor at UCLA on Marschak's invitation. It was during this period, more precisely in May 1967, that the experiment under consideration took place. When asked about the original idea of the experiment, Toda noted:

> It is [...] obscure when I got the idea of the experiment; it could quite well be when I was at the Psychometric Lab [at Chapel Hill], but there is no proof. [...] What I can tell [...] is that the basic idea of the experiment is mine, neither Marschak's nor MacCrimmon's (e-mail to the author, November 19, 2004).

MacCrimmon and Toda administered the experiment to 7 UCLA undergraduates. In two experimental session held on two consecutive days, the students were asked to draw their indifference curves for money and pens, and then for money and French pastries. At the very end of the experiment, the students received a bundle of money and pastries and had to eat the pastries on the spot. Afterwards, they received a bundle containing money and pens. Therefore, this time the subject's choices were not merely conjectural but had actual consequences.

Like in Thurstone's experiment, each student was first presented with a reference bundle (labeled P_0) and asked to compare it with other bundles. Let P_A be a combination the subject prefers to P_0. If more is preferred to less for both goods, the subject should prefer not only P_A, but also all the bundles which yield no lesser amount of each good. In geometrical terms, not only P_A but also all bundles above and to the right of P_A belong to what Thurstone had called the "plus region" and MacCrimmon and Toda termed the "accept region". Symmetrically, if a bundle P_R is rejected in favor of P_0, also all the bundles below and to the left of P_R are in the "reject region". By asking the subject to compare

with P_0 new bundles not yet in the accept or reject regions, MacCrimmon and Toda were able to put tighter and tighter constraint on the locus where the indifference curve could lie, and this with a reduced number of comparisons. When a satiation point existed for a commodity, as in the case of pastries, a less powerful but similar method was applied.

After this comparison phase, each subject had to draw his own indifference curve passing through P_0 and lying in the area between the accept region and the reject region. In order to motivate the individual to draw the indifference curve as carefully as he could to mirror his true preferences, MacCrimmon and Toda used a payoff procedure similar to that employed by G.G. Becker, M.H. DeGroot, and Marschak (1964). The subject was told that after he had drawn the indifference curve through P_0, a bundle in the commodity space would be selected at random. If the selected bundle would lie in the accept region delimitated by the plotted curve, the subject would receive it. If the selected bundle would be in the reject region, the subject would receive P_0. Clearly this procedure induces a utility maximizer to draw his true indifference curve, and this independently from his attitude toward risk.

As mentioned above, the subjects actually received pens and money and had to eat the pastries, so MacCrimmon and Toda claimed that the Wallis-Friedman objection against experiments with merely conjectural stimuli did not apply to their study. Furthermore, MacCrimmon and Toda deferred the determination and distribution of payoffs after the second session of the experiment, so that the income as well as the preference-for-variety effects on which Wallis and Friedman had called attention, could be excluded. Finally, the Wallis-Friedman concern about the long time and the large number of observations necessary to elicit the indifference curves, was handled by exploiting the non-satiation principle in the comparison phase.

Each student drew seven indifference curves for money and pens, and four indifference curves for money and pastries. The money-pens curves for each student did not intersect and were downward sloping. Although the curves were also generally convex, concave parts were recorded for two subjects. With respect to the money-pastries curves, they did not show a general negative slope, and the beginning of the positive slope clearly indicated a point of satiation for pastries. Even though no money-pastries curves crossed, the curves of three subjects joined up at their right end, so violating transitivity. The money-pastries curves were in general convex, but the curves of two subjects exhibited some concave parts.

In conclusion, MacCrimmon and Toda felt that they had "satisfactorily performed an

experiment constructing indifference curves" (MacCrimmon-Toda 1969: 436). They basically attributed the concavities recorded to the subjects' "careless in considering the choices", and the joining-up of the curves to the subjects' "inability to represent graphically the minute differences in [their] preference" (MacCrimmon-Toda 1969: 446 and 447, respectively). Therefore, MacCrimmon-Toda suggested that the indifference curves satisfy the properties commonly postulated by consumer demand theory, at least for the small group of individuals and the particular commodities they considered.

The first element that strikes the reader in the MacCrimmon-Toda paper is its careful methodological discussion of the factors that can mar the experiment's value. This is a major difference not only with respect to Thurstone's 1930 pioneering study, but also in comparison with all the experiments so far considered. The difference shows that between the 1950s and the end of the 1960s methodological awareness in the design of experiments developed significantly, and that experimental economics was emerging as a specific sub-discipline with its own scientific standards.

As regards the MacCrimmon-Toda findings, actually they suggest that violations of convexity and non-intersection are not so rare. Furthermore, the anomalies recorded by MacCrimmon and Toda probably would have been more numerous if they had not so carefully explained to the students the contradictory implications of curve intersections, and had not allowed them to adjust the plotted curves upon ex post reflection.

In any event, the experiment performed by MacCrimmon and Toda had little impact on the subsequent theoretical work in demand analysis. In the period 1969-89, the MacCrimmon-Toda paper was cited only eleven times, mainly with reference to specific applicative problems, often in general surveys, and always in a cursory way.

13. Conclusions

The present study reconstructs a significant part of the early history of experimental economics, namely that concerning the research on consumer demand behavior in the period 1930-70. The experiments performed in this early phase basically regarded the derivation of indifference curves and the evaluation of the transitivity hypothesis. Thereafter, experimental tests dealt with different issues, such as the negativity of the substitution effect, the existence of Giffen goods or the validity of the revealed preference axioms. Experimentation on animal choice behavior was also introduced, but no further experiment on indifference curves or transitivity took place. After 1970, then, a new phase in the history of experimental research in consumer theory began. This phase could be the subject

matter of a future study.

Regarding the period under consideration here, the influence of the experimental studies on the developments in consumer theory deserves a last consideration. As we have seen, such influence was very circumscribed. The fact that the methods for deriving experimentally the indifference curves appeared highly questionable did not prevent the PSAH approach from remaining the standard one. The experimental evidence that preference can easily be intransitive may have provoked a discussion on the status of the transitivity hypothesis, but it certainly did not cause the abandonment of it. Finally, the insight that convexity and non-intersection of the indifference curves cannot be taken for granted has deserved not even a footnote in microeconomics textbooks.

From an epistemological viewpoint, the failure of experimental research to influence the theoretical may be rationalized in different ways[38]. Without entering here into these methodological discussions, I claim that the present study has shown that, at least in America, the failure of experimental research to affect consumer demand theory was not due to unawareness on the part of the economics profession. Maybe unexpectedly, many prominent economists were somehow involved in this stream of research, and almost all the major theorists of the period – Arrow, Chipman, Debreu, Friedman, Frisch, Georgescu-Roegen, Hotelling, Houthakker, Hurwicz, Marschak, Schultz – were aware of at least one of the early experiments in consumer theory. The main absence in this narrative is that of Samuelson, and this appears somewhat surprising given that the axioms of revealed preference, according to their operational character, seem to call quite naturally for an experimental test.

Footnotes

[1]   See, among others, Smith 1992, Roth 1993, Leonard 1994, Innocenti 2000, Mirowski 2002, and Lee-Mirowski 2003.

[2]   In his valuable review of the early history of experimental economics, A. Roth devoted only two and half pages to the subject (cf. Roth 1993: 186-88; reproduced in Roth 1995: 5-6). Partial accounts of the early experiments in consumer theory can also be found in Edwards 1954a, Davis 1958, and Sauermann-Selten 1967.

[3]   In 1927 Fisher proposed a method for measuring the marginal utility of income by means of statistical data on family purchases. However, Fisher's method was still based on the assumption of independent utilities; cf. Fisher 1927: 175 ff.

[4]   More on Thurstone's life and scientific vision in Thurstone 1952 and Jones 1998.

[5]   I thank Lyle Jones for looking over the historic files at the Psychometric Laboratory at Chapel Hill, and Robert Thurstone, Louis Leon's son, for checking the family files. Not even the libraries of the University of Chicago and of the University of North Carolina at Chapel Hill have files containing Thurstone's unpublished papers.

[6]   In a letter to the experimental psychologist W. Edwards dated December 7, 1953, Thurstone specified the following: "I selected as subject a research assistant in my laboratory who knew nothing about psychophysics [...]. She had a very even disposition, and I instructed her to take an even motivational attitude on the successive occasions" (quoted in Edwards 1954a: 387).

[7]   See Hicks–Allen 1934, Lange 1934, Georgescu-Roegen 1936, Samuelson 1938 and Hicks 1939: 18-9.

[8]   A widespread interpretation puts the Hicks–Allen and Slutsky papers in the ordinalist camp. I make my case for their behaviorist and cardinal nature, respectively, in Moscati 2004.

[9]   In Georgescu-Roegen's Papers collected at Duke University, Rare Book, Manuscript, and Special Collections Library, I found no material related to experimental economics.

[10] In the same year, Thurstone's experiment was mentioned in two other papers – Tintner 1942 and Staehle 1942 – but in a very generic way.

[11] More on Friedman's early years and his friendship with Wallis in Hammond 2001.

[12] More on the problems related to Schultz's findings in Hands–Mirowski 1998.

[13] Wallis and Friedman also considered the possibility that psychophysical experiments regarding color and shape of objects or experiments with chimpanzees could be used for economic purposes. However, the application of this type of experimental research to economics appeared them quite problematic; cf. Wallis-Friedman 1942: 181-83.

[14] More on what follows in Edwards 1954a, Smith 1992, Roth 1995: 6-19, Innocenti 2000, and Mirowski 2002: 353-59.

[15] Flood 1954a and Kalnish-Milnor-Nash-Nering 1954 were on games; Coombs-Beardslee 1954 and Flood 1954b were on decision making, and Hoffman-Festinger-Lawrence 1954 was on bargaining.

[16] See Friedman-Friedman 1998: 39-40, and Stigler 1988: 25.

[17] Cf. American Economic Association, Biographical Listings of Members 1970: 375.

[18] Rousseas and Hart did not indicate any exact date. However, they reported that the experiment evolved out of discussions held in a Columbia seminar on the "Content and Methods of the Social Sciences" in 1949-50. Since the paper was published in August 1951, the year 1950 appears the most plausible for their test.

[19] This opinion has already been expressed by A. Roth (1995: 87, n. 6).

[20] Cf. the Report of the Chicago Meeting (1953: 476-77), which contains the abstracts of May's and Papandreou's papers as well as the abstract of Edward's discussion. The abstracts of discussion by

Friedman and Hotelling are not available. About his discussion of May's and Papandreou's papers, Friedman notes: "I am sorry to say I have no recollection of what I thought about the May and Papandreou experiments. So far as I know, I do not have in my papers a copy of my comments at the Chicago meeting" (e-mail to the author, November 11, 2004).

[21] More on May and his adventurous early life in Jones-Enros-Tropp 1984 and Bartlow 1996.

[22] Also Papandreou 1954 presents a summary of the second version of the paper.

[23] The main contributions to this research program were: Georgescu-Roegen 1950 and 1958, Marschak 1955, Quandt 1956, Luce-Raiffa 1957: 371-84, Luce 1958, Debreu 1958, Davidson-Marschak 1959.

[24] Papandreou's is a strong notion of transitivity, which can be used to compare differences of stochastic preferences and hence to construct a cardinal representation of them; cf. Papandreou and others 1957: 4-6.

[25] Five cases (one percent) were unclassified because of incomplete response to test.

[27] See e.g. Sen 1997 and Mongin 2000 for two recent discussions on transitivity.

[28] Cf. Edwards 1954a, Pfouts 1955, Luce 1956, Wagner 1956a, Luce-Raiffa 1957: 25, Wagner 1958, Simon 1959, Wagner 1959, Edwards 1961, Michalos 1967, Fishburn 1968, Tversky 1969, Fishburn 1970, Sonnenschein 1971, Behling-Starke 1973, Fishburn 1974.

[29] Cf. Wagner 1956b, Tullock 1964, Weinstein 1968. In 1958, the psychiatrist J.M. Davis duplicated May's experiment and, similarly to Rose, interpreted the observed circularities as a result of choices among indifferent objects in which a random element enters (cf. Davis 1958: 30). I found no reference to Davis' experiment in the economic literature.

[30] Cf. Marschak 1955, Quandt 1956, Georgescu-Roegen 1958, Davidson-Marschak 1959, Marschak 1959, Block-Marschak 1960, Edwards 1961, Luce-Suppes 1965, Fishburn 1968, Behling-Starke 1973.

[31] Cf. Wagner 1959, Simon 1959, Weinstein 1968, Fishburn 1970.

[32] On the integrability problem see the historical reconstruction contained in Hurwicz 1971.

[33] More on what follows can be found in Edwards 1954a and 1961, Rapoport-Orwant 1962, Friedman 1969, Castro-Weingarten 1970, Roth 1995.

[34] It is of interest that in 1951 V. Smith wrote his MA thesis at Kansas University on "The Integrability Condition in the Pure Theory of Consumer's Behavior". In discussing the empirical content of integrability, Smith touched on the problems connected with directly testing the existence of indifference curves and referred in a footnote to the Wallis-Friedman paper (Smith 1951: 145, n. 77). Smith did not cite Thurstone's experiment. He appeared quite skeptical about the usefulness of the experimental methods for testing integrability, and claimed that "for the present at least, it is necessary for the theorist to satisfy himself with arguments for or against integrability that stem from so-called mental experiments". Smith's decisive contact with the experimental research occurred a year later, in 1952, when he participated as a graduate student at Harvard in one of Chamberlin's market experiments (cf. Smith 1992: 241-46). In Smith's papers collected at Duke University, Rare Book, Manuscript, and Special Collections Library, I found no material relevant for the present paper. For more on Smith's MA thesis see Lee-Mirowski 2003.

[35] I thank Lyle Jones for calling my attention to this part of the experimental literature.

[36] In the economics journals listed in JSTOR, none of these experimental studies is cited in the twenty years following its publication.

[37] Toda's biography is outlined by Crombag 1982.

[38] I proposed my explication in Moscati 2004.

References

Allen, R. G. D. 1936. Professor Slutsky's Theory of Consumers' Choice. *Review of Economic Studies* 3.2: 120-29.

American Economic Association. 1970. Biographical Listings of Members. *American Economic Review* 59.6: 1-492

Arrow, K. J. 1951. *Social Choice and Individual Values*. New York: Wiley.

—. 1959. Rational Choice Functions and Orderings. *Economica* [N.S.] 26: 121-27.

Bartlow, T. L. 1996. Kenneth O. May and the Theory of Social Choice. *Proceedings of the Canadian Society for History and Philosophy of Mathematics* 9: 73-78.

Becker G. M., M. H. DeGroot, and J. Marschak. 1964. Measuring Utility by a Single-response Sequential Method. *Behavioral Science* 9: 226-32.

Behling, O., and F. A. Starke. 1973. The Postulates of Expectancy Theory. *Academy of Management Journal* 16.3: 373-88.

Benson, P. H. 1955. A Model for the Analysis of Consumer Preference and an Exploratory Test. *Journal of Applied Psychology* 39.5: 375-81.

Block, H. D., and J. Marschak. 1960. Random Orderings and Stochastic Theories of Responses. In *Contributions to Probability and Statistics*, edited by I. Olkin and others. Stanford: Stanford University Press.

Castro, B.; and K. Weingarten. 1970. Toward Experimental Economics. *Journal of Political Economy* 78.3: 598-607.

Chamberlin, E. H. 1948. An Experimental Imperfect Market. *Journal of Political Economy* 56.2: 95-108.

Chipman, J. S. 1960. Stochastic Choice and Subjective Probability. In vol. 1 of *Decisions, Values, and Groups*, edited by Dorothy Willner, 70-95. New York: Pergamon Press.

Coombs, C. H., and D. Beardslee. 1954. On Decision-Making Under Uncertainty". In *Decision Processes*, edited by R. M. Thrall, C. H. Coombs, and R. L. Davis, 255-85. New York: Wiley.

Crombag, H. F. M. 1982. Masanao Toda: The Man and His Model of Man. In *Masanao Toda, Man, Robot, and Society*, xi-xviii. Boston-The Hague- London: Nijoff.

Davidson, D., and P. Suppes. 1957. *Decision Making: An Experimental Approach*. Stanford: Stanford University Press.

—, and J. Marschak. 1959. Experimental Tests of a Stochastic Decision Theory. In *Measurement: Definitions and Theories*, edited by C. W. Churchman and P. Ratoosh, 233-69. New York: Wiley.

Davis, J. M. 1958. The Transitivity of Preferences. *Behavioral Science* 3: 26-33.

Debreu, G. 1958. Stochastic Choice and Cardinal Utility. *Econometrica* 26.3: 440-44.

—. 1959. *Theory of Value*. New York: Wiley.

Edgeworth, F.Y. 1881. *Mathematical Physhics*. London: Kegan Paul.

Edwards, W. 1953a. Discussion. *Econometrica* 21.3: 476 (abstract).

—. 1953b. Probability-Preferences in Gambling. *American Journal of Psychology* 66: 349-64.

—. 1954a. The Theory of Decision Making. *Psychological Bulletin* 51.4: 380-417.

—. 1954b. Probability-Preferences among Bets with Differing Expected Values. *American Journal of Psychology* 67: 56-67.

—. 1954c. The Reliability of Probability-Preferences. *American Journal of Psychology* 67: 68-95.

—. 1961. Behavioral Decision Theory. *Annual Review of Psychology* 12: 473-98.

Ellsberg, D. 1961. Risk, Ambiguity and the Savage Axioms. *Quarterly Journal of Economics* 75.4: 643-69.

Fishburn, P. C. 1968. Utility Theory. *Management Science* 14.5: 335-78.

—. 1970. Intransitive Individual Indifference and Transitive Majorities. *Econometrica* 38.3: 482-89.

—. 1974. Lexicographic Orders, Utilities and Decision Rules: A Survey. *Management Science* 20.11: 1442-71.

Fisher, I. [1892] 1925. *Mathematical Investigations in the Theory of Value and Prices*. New Haven: Yale University Press.

—. 1927. A Statistical Method for Measuring "Marginal Utility" and Testing the Justice of a progressive Tax. In *Economic Essays Contributed in Honor of John Bates Clark*, edited by J. H. Hollander, 157-93. New York: Macmillan.

Flood, M. M. 1951-52. A Preference Experiment. Research Memorandums P-256, P-258, and P-263, RAND Corporation.

—. 1952a. On Game-Learning Theory and Some Decision-Making Experiments. Research Memorandum P-346, RAND Corporation.

—. 1952b. Some Experimental Games. Research Memorandum RM-789, RAND Corporation.

—. 1954a. On Game-Learning Theory and some Decision-Making Experiments. In *Decision Processes*, edited by R. M. Thrall, C. H. Coombs, and R. L. Davis, 139-58. New York: Wiley.

—. 1954b. Environmental Non-Stationarity in a Sequential Decision-Making Experiment. In *Decision Processes*, edited by R. M. Thrall, C. H. Coombs, and R. L. Davis, 287-99. New York: Wiley.

Friedman, J. W. 1969. On Experimental Research in Oligopoly. *Review of Economic Studies* 36.4: 399-415.

Friedman, M., and L. J. Savage. 1948. The Utility Analysis of Choices Involving Risk. *Journal of Political Economy* 56.4: 279-304.

—, and R. D. Friedman. 1998. *Two Lucky People*. Chicago: University of Chicago Press.

Frisch, R. 1926. Sur un problème d'economie pure, *Norsk Matematisk Forenings Skrifter* [1] 16: 1-40.

—. 1932. *New Methods of Measuring Marginal Utility*. Tübingen: Mohr.

Georgescu-Roegen, N. 1936. The Pure Theory of Consumer's Behavior. *Quarterly Journal of Economics* 50.4: 545-93.

—. 1950. The Theory of Choice and the Constancy of Economic Laws. *Quarterly Journal of Economics* 64.1: 125-38.

—. 1958. Threshold in Choice and the Theory of Demand. *Econometrica*: 26.1: 157-68.

Hammond, J. D. 2001. Columbia Roots of the Chicago School: The Case of Milton Friedman. Paper presented at the ASSA Meeting, New Orleans.

Hands, D. W., and P. Mirowski. 1998. Harold Hotelling and the Neoclassical Dream. In *Economics and Methodology*, edited by R. Backhouse and others, 322-97. London-New York: Macmillan.

Hart, A. G. 1942. Risk, Uncertainty and the Unprofitability of Compounding Probabilities. In *Studies in Mathematical Economics and Econometrics*, edited by O. Lange and others, 110-18. Chicago: University of Chicago Press.

Hicks, J. R. 1939. *Value and Capital*. Oxford: Claredon Press.

—, and R. G. D. Allen. 1934. A Reconsideration of the Theory of Value, *Economica* [N.S.] 1.2: 52-76, 196-219.

Hoffman, P. J., L. Festinger, and D. H. Lawrence. 1954. Tendencies Toward Group Comparability in Competitive Bargaining. In *Decision Processes*, edited by R. M Thrall, C. H. Coombs, and R. L. Davis, 231-53. New York: Wiley.

Hoggatt, A. C. 1959. An Experimental Business Game. *Behavioral Science* 4: 192-203.

Hotelling, H. 1932. Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions, *Journal of Political Economy*, 40.5: 577-616.

—. 1938. The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates. *Econometrica* 6.3: 242-69.

Houthakker, H. S. 1950. Revealed Preference and Utility Function, *Economica* [N.S.], 17: 159-74.

—. 1961. The Present State of Consumption Theory. A Survey Article. *Econometrica* 29.4: 704-40.

Hurwicz, L. 1971. On the Problem of Integrability of Demand Functions. In *Preference, Utility and Demand*, edited by J. S. Chipman, L. Hurwicz and others, 174-214. New York: Harcourt Brace Jovanovich.

Innocenti, A. 2000. The Early Developments of Experimental Economics: the Influence of Game theory. Working paper, Università di Siena.

Jevons, W. S. 1871. *The Theory of Political Economy*. London-New York: Macmillan.

Jones, C. V., P. C. Enros, and H. S. Tropp. 1984. Kenneth O. May, 1915-1917. His Early Life to 1946. *Historia Mathematica*:11: 359-79.

Jones, L. V. 1959. Prediction of Consumer Purchase and the Utility of Money. *Journal of Applied Psychology* 48.5: 334-337.

—. 1998. L.L. Thurstone's Vision of Psychology as a Quantitative Rational Science. In vol. 3 of *Portraits of pioneers in Psychology*, edited by G. A. Kimble and M. Wertheimer: 85-102. Mahwah, NJ: Erlbaum.

Kalisch, G. K., J.W. Milnor, J. F. Nash, and E. D. Nering. 1954. Some Experimental n-Person Games. In *Decision Processes*, edited by R. M. Thrall, C. H. Coombs, and R. L. Davis, 301-27. New York: Wiley.

Lange, O. 1934. The Determinateness of the Utility Function. *Review of Economic Studies* 1.3: 218-25.

Lee, S. L., and P. Mirowski. 2003. The Purest Form of Rationality is the one Imposed from Without: Vernon Smith, Experimentation, Integrability and the Induced Value Theory. Paper presented at the HES Meeting, Duke University.

Leonard, R. 1994. Laboratory Strife: Higgling as Experimental Science in Economics and Social Psychology. In *Higgling*, edited by N. De Marchi and M. S. Morgan, 343-69. Durham: Duke University Press.

Luce, R. D. 1956. Semiorders and a Theory of Utility Discrimination. *Econometrica* 24.2: 178-91.

—. 1958. A Probabilistic Theory of Utility. *Econometrica* 26.2: 193-224.

—, and H. Raiffa. 1957. *Games and Decisions*. New York: Wiley.

—, and P. Suppes. 1965. Preference, Utility, and Subjective Probability. In vol. 3 of *Handbook of Mathematical Psychology*, edited by R. D. Luce, R. R. Bush and E. Galanter, 249-410. Wiley: New York.

32

MacCrimmon, K. R., and M. Toda. 1969. The Experimental Determination of Indifference Curves. *Review of Economic Studies* 36.4: 433-51.

Marschak, J. 1950. Rational Behavior, Uncertain Prospects, and Measurable Utility. *Econometrica* 18.2: 111-41.

—. 1955. Norms and Habits of Decision Making Under Uncertainty. In *Mathematical Models of Human Behavior*, 45-53. Stamford, Conn.: Dunlap and Associates.

—. 1959. Binary Choice Constraints and Random Utility Indicators. *Cowles Foundation Papers* 155.

May, K. O. 1952. Review of "Social Choice and Individual Values" by Kenneth J. Arrow. *Science and Society* 16.2: 181.

—. 1953. The Intransitivity of Individual Preferences. *Econometrica* 21.3: 476 (abstract).

—. 1954. Intransitivity, Utility, and the Aggregation of Preference Patterns, *Econometrica*, 22.1: 1-13.

Mayer, J. 1933. The Meeting of the Econometric Society in Syracuse, New York, June, 1932. *Econometrica* 1.1: 94-104.

Michalos, A. C. 1967. Postulates of Rational Preference. *Philosophy of Science* 34.1: 18-22.

Mirowski, P. 2002. *Machine Dreams*. Cambridge: Cambridge University Press.

Mongin, P. 2000. Does Optimization Imply Rationality? *Synthese* 124: 73-111.

Moscati, I. 2004. History of al Consumer Demand Theory: A Neo-Kantian Rational Reconstruction. Working Paper, Università Bocconi.

Mosteller, F., and P. Nogee, 1951. An Experimental Measurement of Utility. *Journal of Political Economy* 59.5: 371-404.

Papandreou, A. G. 1953. An Experimental Test of an Axiom in the Theory of Choice. *Econometrica* 21.3: 477 (abstract).

—. 1954. An Experimental Test of a Proposition in the Theory of Choice, *Consumer Behavior* 1: 90-91 (abstract).

—, and others. 1955. A Test of a Proposition in the Theory of Choice (with O. H. Sauerlander, O. H. Brownlee, L. Hurwicz and W. Franklin). *Econometrica* 23.3: 333-34 (abstract).

—, and others. 1957. A Test of a Stochastic Theory of Choice (with the collaboration of O. H. Sauerlander, O. H. Brownlee, L. Hurwicz and W. Franklin). *University of California Publications in Economics* 16: 1-18.

Pareto, V. 1900. Sunto di alcuni capitoli di un nuovo trattato di economia pura del prof. Pareto. *Giornale degli Economisti* [2] 20: 216-35, 511-49.

—. 1906. *Manuale di Economia Politica*. Milano: Società Editrice Libraria.

Pessemier, E. A. 1960. An Experimental Method for Estimating Demand. *Journal of Business* 33.4: 373-83.

—. 1963. *Experimental Methods of Analyzing Demand for Branded Consumer Goods*. Pullman, WA: Washington State University Press.

Pfouts, R. W. 1955. Prolegomena to the Testing of Utility Theory. *Southern Economic Journal*: 22: 178-88.

Quandt, R. E. 1956. A Probabilistic Theory of Consumer Behavior. *Quarterly Journal of Economics* 70.4: 507-36.

Rapoport, A., and C. Orwant. 1962. Experimental Games: A Review. *Behavioral Science* 7: 1-37.

Report of the Chicago Meeting, December 27-29, 1952. 1953. *Econometrica* 21.3: 463-490.

Rose, A. M. 1957. A Study of Irrational Judgments. *Journal of Political Economy* 65.5: 394-402.

Roth, A. E. 1993.The Early History of Experimental Economics. *Journal of the History of Economic Thought* 15: 184-209.

—. 1995. Introduction to Experimental Economics. In *Handbook of Experimental Economics*, edited by J. H. Kagel and A. E. Roth, 3-109. Princeton: Princeton University Press.

Rothenberg, J. 1953. Marginal Preference and the Theory of Welfare. *Oxford Economic Papers* [N.S.] 5.3: 1953: 248-263.

Rousseas, S. W., and A. G. Hart. 1951. Experimental Verification of a Composite Indifference Map. *Journal of Political Economy* 59.4: 288-318.

Samuelson, P. 1938. A Note on the Pure Theory of Consumer's Behaviour. *Economica* [N.S.] 5.17: 61-71, 353-54.

—. 1950. The Problem of Integrability in Utility Theory, *Economica* [N.S.] 17.68: 355-85.

Sauermann, H. and R. Selten. 1959. Ein Oligolpolexperiment. *Zeitschrift für die Gesamte Staatswissenschaft* 115: 427-71.

—. 1967. Zur Entwicklung der Experimentellen Wirtschaftsforschung. In *Beiträge zur Experimentellen Wirtschaftsforschung*, edited by H. Sauermann, 1-8. Tübingen: Mohr.

Schelling, T. C. 1957. Bargaining, Communication, and Limited War. *Journal of Conflict Resolution* 1: 19-36.

Schultz, H. 1931. The Italian School of Mathematical Economics, *Journal of Political Economy*, 39.1: 76-85.

—. 1935. Interrelations of Demand, Price, and Income, *Journal of Political Economy*, 43.4: 433-481.

—. 1938. *The Theory and Measurement of Demand*. Chicago: University of Chicago Press.

Sen, A. 1997. Maximization and the Act of Choice. *Econometrica* 65.4: 745-79.

Siegel, S., and L. E. Fouraker. 1960. *Bargaining and Group Decision Making: Experiments in Bilateral Monopoly*. New York: McGraw-Hill.

Simon, H. A. 1959. Theories of Decision-Making in Economics and Behavioral Science. *American Economic Review* 49.3: 253-83.

Sippel, R. 1997. An Experiment on the Pure Theory of Consumer's Behavior, *Economic Journal*, 107.444: 1431-44.

Slutsky, E. 1915. Sulla teoria del bilancio del consumatore. *Giornale degli Economisti* [3] 51: 1-26.

Smith, V. L. 1951. *The Integrability Condition in the Pure Theory of Consumer's Behavior*. Unpublished MA thesis, University of Kansas, currently at Anschutz Library, University of Kansas.

—. 1962. An Experimental Study of Competitive Market Behavior. *Journal of Political Economy* 70: 111-37.

—. 1992. Game Theory and Experimental Economics. Beginning and Early Influences. In *Toward a History of Game Theory*, edited by E. R. Weintraub, 241-282. Durham: Duke University Press.

Sonnenschein, H. F. 1971. Demand Theory without Transitive Preferences, with Applications to the Theory of Competitive Equilibrium. In *Preference, Utility and Demand*, edited by J. S. Chipman, L. Hurwicz and others, 215-23. New York: Harcourt Brace Jovanovich.

Staehle, H. 1942. The Measurement of Statistical Cost Functions: An Appraisal of Some Recent Contributions. *American Economic Review* 32.2: 321-33.

Stigler, G. J. 1939. The Limitations of Statistical Demand Curves. *Journal of the American Statistical Association* 34.207: 469-81.

—. 1947. Notes on the History of the Giffen Paradox. *Journal of Political Economy* 55: 152-56.

—. 1988. Memoirs of an Unregulated Economist. New York: Basic Books.

Stone, J. J. 1958. An Experiment in Bargaining Games. *Econometrica* 26: 286-97.

Thrall, R. M., C. H. Coombs, and R. L. Davis (Eds.). 1954. *Decision Processes*. New York: Wiley.

Thurstone, L. L. 1927a. A Law of Comparative Judgment. *Psychological Review* 34: 273-86.

—. 1927b. The Method of Paired Comparisons for Social Values. *Journal of Abnormal and Social Psychology* 21: 384-400.

—. 1931. The Indifference Function. *Journal of Social Psychology* 2: 139-67.

—. [1952] 1959. An Experiment in the Prediction of Choice. In L.L. Thurstone, *The Measurement of Values*, 161-69. Chicago: University of Chicago Press.

—. 1952. L.L. Thurstone. In vol. 4 of *A History of Psychology in Autobiography*, edited by E.G. Boring and others: 295-321. Worcester, MA: Clark University Press.

—. 1959. *The Measurement of Values*. Chicago: University of Chicago Press.

—, and L. V. Jones. 1957. The Rational Origin for Measuring Subjective Values. *Journal of the American Statistical Association* 52.280: 458-71.

Tintner, G. 1942. A Contribution to the Non-Static Theory of Choice. *Quarterly Journal of Economics* 56.2: 274-306.

Toda, M. 1962. The Design of a Fungus-Eater: A Model of Human Behavior in an Unsophisticated Environment. *Behavioral Science* 7: 164-83.

Tullock, G. 1964. The Irrationality of Intransitivity. *Oxford Economic Papers* [NS] 16.3: 401-06.

Tversky, A. 1969. Intransitivity of Preferences. *Psychological Review*. 76: 31-48.

von Neumann, J., and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Wagner, H. M. 1956a. An Eclectic Approach to the Pure Theory of Consumer Behavior. *Econometrica* 24.4: 451-66.

—. 1956b. A Note on Testing the Transitivity Axiom. *Southern Economic Journal*: 22: 493-494.

—. 1958. Advances in Game Theory. *American Economic Review* 48.3: 368-87.

—. 1959. The Case for "Revealed Preference". *Review of Economic Studies* 26.3: 178-89.

Wallis W. A., and M. Friedman. 1942. The Empirical Derivation of Indifference Functions. In *Studies in Mathematical Economics and Econometrics*, edited by O. Lange and others, 175-89. Chicago: University of Chicago Press.

Weinstein, A. A. 1968. Individual Preference Intransitivity. *Southern Economic Journal* 34: 335-43.

Weintraub, E. R. 1983. On the Existence of a Competitive Equilibrium: 1930-1954. *Journal of Economic Literature* 21.1: 1-39.