Dagsvik J. K., S. Strøm og Z. Jia (2006): Utility of income as a random function: Behavioral characterization and empirical evidence, Mathematical Social Sciences, Volum 51, Issue 1, pp 23-57

Title: **Utility of income as a random function: Behavioral characterization and empirical evicence**

Author: John K. Dagsvik
Steinar Strøm
Zhiyang Jia

Publisher: Elsevier
DOI: http://dx.doi.org/10.1016/j.mathsocsci.2005.07.006

Source: **Mathematical Social Sciences**

5 July 2005

Utility of Income as a Random Function: Behavioral Characterization and Empirical Evidence

John K. Dagsvik[a]*, Steinar Strøm[b], Zhiyang Jia[a]

[a] Research Department, Statistics Norway, P.O. Box 8131, Dep., 0033, Oslo Norway

[b] Department of Economics, University of Oslo, P.O. Box 1095 Blindern, 0317, Oslo Norway

Abstract

The paper proposes a particular approach to model the utility of income. We develop a theoretical framework that restricts the class of admissible functional forms and distributions of the random components of the model. The theoretical approach is based on theories of probabilistic choice and ideas that are used in modern psychophysical research. From our theoretical framework, we obtain the empirical model and the corresponding likelihood function. The empirical analysis is based on a "Stated Preference" survey. The model fits the data quite well. Finally, we discuss the concept of cardinality and the implications for consumer demand relations.

Keywords: Utility of income, Random utility, Invariance principles

JEL classification: B21, D19

* Thanks to Anne Skoglund for technical assistance and correction of errors. We have benefited greatly from the reports from two anonymous referees, which significantly improved the paper. We also thank Olav Bjerkholt for helpful comments. Corresponding author: John K. Dagsvik,. Tel: +47 21094871, Fax: +47 21090040 Email: john.dagsvik@ssb.no.

1. Introduction

Utility of income, marginal utility of income and the elasticity of the marginal utility of income are widely used concepts in economics. For example, in analysis of welfare, game theory, choice under uncertainty and dynamic choice, models are formulated in terms of (time independent) utility of income. The utility of income is of course also basic within the theory of consumer behavior since it is equivalent to the indirect utility—as a function of income (when prices are kept fixed). Despite the central role utility of income plays in economics, "direct" empirical studies of how utility varies with income are rare.

In this paper we develop a stochastic model for the utility of income. By this we understand that the utility function depends on random error terms. The motivation for introducing random error components is that, (i) these terms represent unobservables as viewed by the researcher, (ii) the errors may be random even to the decision-maker himself in the sense that he may make different choices in replications of identical choice settings, cf. Quandt (1956) and Thurstone (1927). This notion of individual randomness is consistent with psychological experiments and the explanation is that the agent may find it difficult to assess a fixed utility level once and for all to the respective alternatives. The agent's assessments will typically vary according to his moods and whims. Another reason for non-anticipating fluctuations in the agent's tastes may be due to uncertainty: As new information appears, the agent will update his tastes accordingly.

A common problem with most quantitative economic models is the lack of theoretical justification for the choice of functional form and the distribution of unobservables. The tradition in economics has been to employ ad hoc assumptions with regards to functional form and the distribution of unobservables; alternatively to rely on non-parametric approaches[1]. In this paper we propose an alternative strategy, namely an axiomatic approach to justify the choice of functional form of the utility function and the distribution of unobservables. In this approach, we have adopted and modified ideas and principles from the literature of psychophysical measurement. Within psychophysical measurement there is a tradition that addresses the problem of scale representations of the relation between physical stimuli and sensory response. A central part of this literature is concerned with the interpretation and implications of specifications and laws that are invariant under admissible transformations of the input variables. Typically, these transformations are scale- or affine transformations. In fact we demonstrate how the application of invariance principles similarly to the ones employed in psychophysics, combined with a version of the "Independence from Irrelevant Alternatives" axiom, lead to explicit characterizations of functional form and the distribution of the

[1] Simon (1986):"Contemporary neoclassical economics provides no theoretical basis for specifying the shape and content of the utility function, and this gap is very inadequately filled by empirical research using econometric techniques. The gap is important because many conclusions that have been drawn in the literature about the way in which the economy operates depend on assumptions about consumers' utility function."

random terms of the utility function. We consider these invariance principles to be intuitive and plausible as a theoretical rationale for restricting the class of admissible specifications, as we shall discuss below.

Our empirical analysis is based on interview data from a “Stated Preference" (SP) survey. We consider this type of survey data to be a promising avenue to advance beyond conventional econometric analysis based on market data. Recall that market data yield only *one* observation for each individual at each point in time. In contrast, SP data are generated through experiments in which the participants are exposed to several trials. Thus, with the SP approach the researcher can acquire *several* observations for each individual. In some cases this has enabled the researcher to formulate behavioral models that are estimated separately for each individual. The empirical model we specify and estimate is based on the corresponding theoretical model we have developed in this paper. A particular goodness of fit measure shows that the model fits the data rather well.

The paper is organized as follows. Section 2 contains a discussion of the literature and the relationship between psychophysics and the measurement of utility. In section 3, we present the theoretical model and in section 4 we extend the model to allow for heterogeneity in preferences. Section 5 discusses the expenditure function that corresponds to the utility model and its distributional properties are examined. Section 7 and 8 present the empirical specification, estimation results and a specification test. Section 9 discusses the concept of cardinal utility and in section 10 we derive some implications for the structure of demand relations.

2. The measurement of utility and the link to psychophysics: a review

Here we shall briefly discuss some selected works that analyze theoretical and empirical issues related to the measurement of sensation in general and the measurement of utility in particular. We refer to Ellingsen (1994) for an excellent survey of the attempts to measure utility and its variation with income.

One of the first to specify a statistical method for measuring utility was Fisher (1892, 1918, 1927). However, to the best of our knowledge, the first one to estimate the marginal utility of money was Frisch (1926, 1932). Frisch (1926) introduced certain behavioral or choice axioms. The choice axioms Frisch referred to are of two types. The "Axioms of the first kind", also called "Axioms related to a given position", are preference ordering axioms concerning completeness, transitivity and regularity and imply an ordinal utility representation. The "Axioms of the second kind", also called "Axioms related to different positions", give restrictions on the ordering of changes from one position in the commodity space to another and imply a cardinal utility representation, which means that individuals are able to compare and rank changes in the commodity space, given the reference points. Later, in his lectures in the 1940s, Frisch called the axioms of the second kind "Inter-local Choice axioms". On several occasions Frisch expressed the view that inter-local choice axioms are highly plausible because most of the individuals' daily actions imply that they are in fact able to make inter-local comparisons. Despite this strong belief in the existence of a cardinal utility function, derived from the axioms mentioned above, Frisch never carried out surveys where the respondents were asked to rank utility differences. Instead he assumed an additively separable utility function and used the cardinal property (cross-derivatives of the utility function are identically zero) of the utility function to estimate cardinal utility concepts like the marginal utility of income and the elasticity of marginal utility of income with respect to income, cf. Frisch (1926, 1959), and Johansen (1960). In his earliest work referred to above, he assumed that there exists at least one good with the property that its marginal utility of consumption is independent of the consumption of other goods. The additive assumption was never tested against market or survey data.

In an attempt to revitalize the cardinal utility concept and to employ utility functions to describe consumer behavior van Praag and numerous co-authors (hereafter called the Leyden school) carried out large scale surveys, see for instance van Praag (1968, 1971, 1991,1994), van Herwaarden, Kapteyn and van Praag (1977), van Herwaarden and Kapteyn (1981), Kapteyn and Wansbeck (1985). A discussion and critique of their approach is given in Seidl (1994) to which van Praag and Kapteyn responded in van Praag and Kapteyn (1994). The data they have used are typically collected through Income Evaluation Questions (IEQ). This means that each respondent was asked to indicate (under his current conditions), what income level y_6 or above, of net household income per year would he consider to be excellent, what income interval (y_5, y_6) would he consider to be good, what income interval (y_4, y_5) is more than sufficient, what income interval (y_3, y_4) would he consider to be sufficient, what income interval (y_2, y_3) would he consider to be insufficient, what income interval (y_1, y_2) would he considered to be bad and what income level y_1 or below would he consider to be very bad[2]. The $\{y_j\}$ represent the respective income boundaries reported by the respondents. It is assumed that the respondents evaluate income on the basis of the utility that they derive from income. Thus, the answers may be used to recover an underlying utility function. It is not implied in the IEQ scheme that the respondents rank utility differences and therefore additional assumptions have to be introduced in order to interpret the answers as yielding information about a *cardinal* utility function. In the Leyden school approach it is *assumed* that the so called Equal Quantile Assumption (EQA) holds. The EQA states that the respondents maximize informational content by letting the perceived difference in utility between two adjacent labels be equal. In a study by Buyze (1982), it was concluded that EQA provides a reasonable approximation to reality. To proceed with numerical estimates of the parameters of the utility function one has to specify the functional form of the utility function. In the Leyden school approach the utility function is assumed to have the same functional form as a log-normal

[2] Although most questionnaires of the Leyden School used a six-level question, some studies used eight or nine levels.

distribution function, see van Praag (1968, 1971) and Seidl (1994) for more details. Van Herwaarden and Kapteyn (1981) reported the outcome of tests on 13 different functional form specifications, which implied that a logarithmic utility function gave a better fit than a log-normal utility function. However, the authors still preferred the log-normal form. The Leyden School approach has been criticized by Seidl (1994) who argues that key features of the employed model, such as the log-normal functional form of the utility function, are based on ad hoc assumptions rather than on principles derived from convincing axioms. Seidl (1994) concludes that instead one should apply Weber-Fechnerian laws or Stevens' power law in the measurement of the utility income, cf. Stevens (1975) and Gescheider (1997).

With the exception of Fisher (1892), Frisch (1929, 1932) and the Leyden School and their followers, economists traditionally express considerable uneasiness when confronted with the issue of how to measure utility. In contrast, psychologists have for a long time been concerned with both theoretical and empirical aspects of measuring sensory response as a function of physical stimuli such as intensity of sound, light, and money amounts. Within psychophysics the study of mathematical laws for the relation between physical stimuli (money) and sensory response (utility) seem to have started with Fechner (1860/1966), Thurstone (1927) and Stevens (1946, 1951). After Fechner introduced his psychophysical methodology in 1860 a vivid debate took place. For a summary of this debate, see Heidelberger (2004), ch. 6.4. According to Heidelberger, the debate centred initially on three issues; (i) whether Fechner's measurement method and mathematical law for the link between stimuli and response are correct, (ii) whether Fechner's law is a relation between the external stimulus and inner psychophysical excitation, or between sensitivity and awareness of sensation, or (as Fechner thought) a relation between the sensation and psychophysical activity (i.e., neural excitation), and (iii) if it is at all meaningful to deal with the measurement of sensations and psychological magnitudes in general. Mach (1886/1996) contributed to clarifying some fundamental issues. His notion of measurement is, in short, as follows:

> "I measure a phenomenon that I experience, meaning that I have a sensation of it as one of its features, by numerically representing the behavior of an external observational element serving the purpose of being a feature of my sensation, and this happens in such a way that the order inherent in the external feature correlates isotonically with the order within the sensation: If the sensation becomes stronger, the external feature also increase."

From a passage in Fechner (1882/1965), where he replies to one of his critics, it seems clear that Mach's point of views are completely consistent with Fechner's interpretation, see Heidelberger (2004), p. 240. Our essential point here is that the Fechnerian school does *not* claim that the measurement of sensation leads to more precise knowledge of the "true" or "real essence" of sensation or to have identified objectivity in subjectivity. It is only claimed that in *one particular aspect* one has found a way to understand the relation between sensation and the exterior world.

The arguments of Fechner and Mach are also valid for the measurement of utility. Certainly, our utility concept is not meant to represent some sort of psychological happiness or states of fulfilment in a deep existential sense, or notions along such lines. Similarly to Fechner's and Stevens' psychophysical law, utility is only meant to represent peoples judgments about ordering sensations of stimuli on an ordinal scale (quantities of goods), or similarly, representing sensations of changes in stimuli on a ratio or interval scale (changes of quantities of goods).

In contrast to Fechner's logarithmic law, Stevens (1957, 1975) proposed the "power law", which is claimed to represent the link between stimulus and sensation. To substantiate this claim Stevens has presented both theoretical arguments as well as an impressive amount of empirical results from laboratory type experiments. See also Gescheider and Bolanowski (1991a,b).

There are several different types of survey questionnaires applied by Stevens and his followers to obtain SP data, cf. Falmagne (1985) and Gescheider (1997). One frequently applied method is called *Magnitude estimation*. In a typical magnitude estimation experiment, questions such as the following are asked: "Suppose you are given 1000 US dollars. How much more money will you need to increase your utility by 20 per cent?" The initial stimulus 1000 $ is called the *standard modulus,* or simply a *standard*. In another version of the method *no standard is provided*. The subject is simply asked to assign to any stimulus presented any number that seem suitable as an estimate of the sensation magnitude. Yet another method is labeled *Production and matching*. Here the subject is requested to react to stimuli (money) by "producing" a value of a sensory variable, for example, by turning a dial. There are several versions of this method. In a version called *Magnitude production* the procedure used in magnitude estimation is reversed. Thus, the subject is given a number and asked to produce a matching intensity of the stimulus. In a second version called *Ratio production* the subject is instructed to adjust the intensity of the stimulus in such a manner that it appears to be a particular multiple or fraction of a standard. For example, the subject may be asked to produce a tone intensity appearing one third as loud as the standard tone of the same frequency. A third version is called *Cross-modality matching*. In this method two experiments based on magnitude estimation are conducted first. For example, the two sensory continua may be loudness and brightness. Second, the subject is requested to directly match the values from one sensory continuum to the other.

At first glance such methods may seem strange and ill suited to obtain sensible results in economics. The reason why is the convention, established purely by habit, that agents are only able to make ordinal rank orderings and that only observations of actual choices, that is, market data can be trusted. Many economists typically believe that agents reveal their "true preferences" only under market like conditions; i.e., when financial incentives matter[3]. A good illustration of the skepticism

[3] Smith (1990): "Although replication using cash payoffs (where this has not been done) is certainly needed, I think it is mistake to assume that the economist's paradigm will somehow be rescued in the context of these experimental designs, if experimenters would just pay money."

among economists as regards laboratory type SP- experiments based on questionnaires is reported in Sen (1982, p. 9):

> "One reason for the tendency in economics to concentrate only on "revealed preference" relations is a methodological suspicion regarding introspective concepts. Choice is seen as information, whereas introspection is not open to observation. ... Even as behaviorism this is particular limited since verbal behavior (or writing behavior, including response to questionnaires) should not lie outside the scope of the behaviorist approach."

In a large number of experiments Stevens and his followers have demonstrated that their data, which are consistent with the power law and different experimental methods, such as the ones described above, yield consistent results. Perhaps the most startling result is that in the cross-modality matching method subjects are not only capable of performing the task requested in such experiments without much difficulty, but they also produce reasonably regular data. It seems however, to have been overlooked by Stevens and his followers that results obtained by methods that depend on a standard, such as Magnitude estimation, not necessarily will be independent of the chosen standard. Morover, one cannot be sure that the data obtained are independent of the order in which stimuli are presented (commutativity property). These problems were pointed out by Narens (1996), and tests were carried out by Ellermeier and Faulhammer (2000) who found that the commutativity property seems to hold but that results do indeed depend on the standard.

Clearly, the IEQ approach of the Leyden School is a version of Magnitude production, where instead of numbers, the subject is given questions that are supposed to represent utility levels. Hence, in this case no standard is used, and consequently Narens' critique should not be relevant for analyses based on questionaires such as the one we use in this study. However, results *may* depend indirectly on a "reference standard", namely current income and possibly other conditions of the household. This is indeed confirmed by the Leyden School and it is also found in the empirical analysis of this paper.

With Stevens' results as a point of departure, Luce (1959b) took an important step towards formulating a suitable formal theory from which laws such as the power function can be shown to follow. In the last four decades several authors have been following up this line of research and there exists now a considerable body of literature where explicit functional form characterization- and restrictions are obtained from surprisingly general invariance principles, such as for example the requirement of scale invariance as argued by Stevens, see Stevens (1946,1951,1957,1975). A good reference source to this kind of theories is the book by Falmagne (1985).

In the context of random utility models, an early contribution within the tradition of Fechnerian psychophysics is Thurstone (1927). Thurstone conducted choice- and psychophysical experiments among students and often found that some students would make different choices when choice experiments where replicated. To account for the variability of responses in choice experiments, Thurstone proposed a model based on the idea that a stimulus induces a "psychological state", which is a realization of a random variable. From this idea he was led to formulate a random utility model, that is, to represent preferences over alternatives by random variables (random utilities), in which the individual decision-maker would choose the alternative with the highest value of the random variable. He assumed that these random utilities were normally distributed. In the binary choice setting this setup yields the so-called Probit choice model.

Further important contributions to the theory of stochastic choice models were made in the 1950s and subsequent decades. Luce (1959a) introduced his famous choice axiom, "Independence from Irrelevant Alternatives" (IIA), and demonstrated that this axiom is equivalent to a model that determines choice probabilities in a multinomial setting as a simple function of the choice set (set of feasible alternatives) and of alternative-specific response strengths (representative utilities); the so-called Luce model. Later Holman and Marley (see Luce and Suppes, 1965) demonstrated that the Luce model is indeed consistent with an additive random utility model in which the representative utilities in the Luce model can be interpreted as the respective deterministic parts of the corresponding random utilities and where the (additive) error terms are i.i. extreme value distributed.

Among economists, an early contribution in the tradition of Fechnerian psychophysics is due to Debreu (1958). Without relying on a random utility formulation he considered a stochastic choice setting with binary choice probabilities that were assumed to satisfy certain conditions. Given these conditions he and other researchers (cf. Falmagne, 1985, and Suppes et al., 1989) demonstrated that they imply a (deterministic) cardinal utility representation of the choice probabilities in the sense that the binary choice probabilities can be expressed as a monotone function (cumulative distribution function) of the utility difference. This cumulative distribution function is unique apart from a scale transformation of the argument[4]. McFadden (1973), Yellott (1977) and Strauss (1979) provided important characterizations of random utility models satisfying IIA. In particular, they showed that under different regularity conditions, the additive random utility model with independent random terms is consistent with IIA only when the error terms are extreme value distributed. McFadden (1978, 1981) extended the Luce model by introducing the Generalized Extreme Value model (GEV), which contains the nested logit models as a special case. The motivation for this extension is that IIA may not hold when the same latent aspects characterize several alternatives. The GEV model is derived from an additive random utility model where the joint distribution of the error terms is a multivariate extreme value distribution. Dagsvik (1994,1995) has demonstrated that any random utility model can be approximated arbitrarily closely by GEV models. Thus the general GEV allows a very rich pattern of correlation between the error terms. Consequently, the GEV framework can be applied in choice setting where the IIA property is questionable. Similarly, the Probit model of Thurstone has been extended to the multinomial choice setting in which the utilities have multinormally distributed error

[4] Debreu (1958) only proved the existence of the cardinal utility function and did not discuss the c.d.f. linking the utility function to the binary choice probabilities. The relationship between the cardinal utility function and the binary choice probabilities was established by Falmagne (1985).

terms. Through the development of appropriate econometric theory tailor-made for this type of models and through a large number of applications (cf. McFadden, 2001), McFadden and others have demonstrated that the random utility framework is very useful for econometric analysis. We refer to Anderson et al. (1992) and Fishburn (1998) for more detailed reviews of stochastic choice models.

3. The model

We consider a general choice setting in which the consumer makes choices of quantities of consumption bundles as well as discrete choices among variants of differentiated products and other qualitative alternatives (such as type of work, schooling and transportation modes). The discrete alternatives are indexed by *(j, r),* where $r \in C_j, j \in \Omega$, and $\{C_j\}$ are disjoint sets. Thus, the sets $\{C_j, j \in \Omega\}$ represent a partition of the universal set of discrete alternatives. The sets $\{C_j\}$ and Ω are possibly infinite. A good example of this structure of the set of alternatives is the case of differentiated products. In this example the index set Ω represents an enumeration of products while the subset C_j is the set of variants of product *j*. We assume that the degree of similarity between product variants of different products is independent of which products are compared. However, we make no assumptions about the degree of similarity between alternatives within each set C_j. Let y denote the agent's income, and let $\tilde{U}(j, y)$ denote the *conditional indirect utility* given that the discrete alternative belongs to C_j, $j \in \Omega$. Thus, $\tilde{U}(j, y)$ is the utility of the most preferred consumption bundle and product variant, given product type *j* and given income y and prices. For notational simplicity we have suppressed the price vector in the notation of $\tilde{U}(j, y)$. We shall assume that the utility function is random and $y \ge \gamma$, where γ is interpreted as a subsistence level that may be specific to the agent. Let B be the family of all finite subsets of Ω.

3.1. The individual utility function as a stochastic process

In this section we shall introduce behavioral assumptions that will enable us to characterize the stochastic properties of the utility function. To this end we shall distinguish between "Conditions" and "Axioms". By an axiom we understand assumptions that can be supported by a clear behavioral intuition, in contrast to regularity conditions that do not necessarily have a behavioral interpretation.

Axiom 1

The conditional indirect utility processes $\{\tilde{U}(j, y), y \ge \gamma\}$, $j \in \Omega$, *are independent max-stable processes (with y as parameter) with standard Fréchet marginals.*

Recall that a max-stable process has finite dimensional distributions which are multivariate extreme value distributions. The Fréchet distributions (also known as the type I extreme value distributions), defined on R_+, have one-dimensional marginal distributions equal to $\exp\left(-bx^{-a}\right)$ for $x>0$ and $a>0$ and $b>0$. When $a=b=1$ we call it the *standard Fréchet distribution*.

The reason why we call the assumption above an axiom is because it is motivated by the "Independence from Irrelevant Alternatives" assumption (IIA). Specifically, McFadden (1973) and Yellott (1977) showed that IIA is equivalent to a choice model that can be represented by additive independent random utilities with type III extreme value distributed random terms[5]. Recall also that Dagsvik (1995) has demonstrated that in the absence of state dependence effects and transaction costs there is no loss of generality in restricting the utility processes to max-stable processes. This is so because the "multiperiod" random utility model (with income y as parameter) can be approximated arbitrarily closely by random utility models generated from max-stable processes. It is therefore the requirement of independence in Axiom 1 that yields the essential restriction. For a summary of the properties of multivariate extreme value distribution functions, we refer to Resnick (1987).

Axiom 1 implies that one can, for each given *y*, write

$$\tilde{U}(j,y)=\tilde{v}(j,y)\tilde{\varepsilon}(j,y) \tag{1}$$

for $j\in\Omega$, where $\tilde{v}(\cdot)$ is a positive deterministic mapping from $\Omega\times[\gamma,\infty)$ to R_+ and $\tilde{\varepsilon}(j,y)$ is a random variable that is standard Fréchet distributed.

For a given income y and a given choice set $B=\{1,2,...,m\}\in\mathrm{B}$, let $J_B(y)$ denote the index of the preferred attribute in *B*, i.e.,

$$J_B(y)=j\Leftrightarrow\tilde{U}\left(j,y\right)=\max_{r\in B}\tilde{U}\left(r,y\right).$$

Axiom 2 (DIM)

For $B\in B$

$$P\left(\max_{r\in B}\tilde{U}\left(r,y\right)\le u\,\middle|\,J_B\left(x\right),x\le y\right)=P\left(\max_{r\in B}\tilde{U}\left(r,y\right)\le u\right).$$

Axiom 2 states that the conditional distribution of the indirect utility at income *y*, given the index of the preferred alternative at any income *x*, $x\le y$, equals the unconditional distribution of the indirect utility.

[5] Note that a multiplicative random utility model with Fréchet distributed error terms is equivalent to a corresponding additive random utility model with type III extreme value distributed error terms.

For $x = y$, Axiom 2 is a version of the DIM property (*Distribution is Invariant of which variable attains the Maximum*), proposed by Strauss (1979). He did, however, not produce any behavioral motivation to support it. Note that it is understood here that preferences are *exogenous*, as conventionally assumed in economics. This means that utilities are *not* affected by previous choice experience. For the case $x = y$, our motivation for DIM is as follows: The values of the alternatives are fully captured by the corresponding utilities and the indirect utility is the utility of the chosen alternative. Once the highest utility has been attained the information about which alternative that yields maximum utility does not represent additional information that is relevant for the value of the indirect utility. Moreover, Axiom 2 states that preferred alternatives under income less than *y* should be irrelevant for the evaluation of the highest utility at income *y*. This is so because the alternatives available at income *x* also are available at income *y* when $x \le y$. Consequently, the "information" about the preferences over consumption possibilities that are restricted by income *y* includes the corresponding information when income is less than *y*.

DIM represents of course an idealization that cannot be expected to hold exactly in many real life situations. For example, it is clear that in many real life choice settings preferences may indeed be influenced by choice experience. Note moreover, that, a priori, it is *not* evident that there exist stochastic utility processes which satisfy DIM. However, if we can find utility processes that satisfy DIM, then this will be useful for obtaining a representation of preferences in idealized choice settings (under DIM).

Axiom 3

The conditional indirect utility processes $\{\tilde{U}(j,y), y \ge \gamma\}$, *j∈Ω, are non-decreasing in y with probability one. The probability that the utility process is constant in any given income interval is positive.*

Axiom 3 means that there is a positive probability (possibly rather small) that the agent's utility of alternative *j* will remain constant even if income increases. Thus, if we consider two incomes y_2 and y_1, where $y_2 > y_1$, there is a positive probability that $U(j, y_2) = U(j, y_1)$. This property is consistent with the famous notion in psychophysics called "just noticeable differences". Within economics an early discussion on this is found in Quandt (1956):

> "…that the consumer is often ignorant of the exact state of his preferences and he is frequently insensitive to small changes or differences in stimuli. As a result, a small movement in any direction from any initial position may leave the consumer as well off as before. It might be suggested that we deal with this problem by considering an indifference map consisting not of indifference curves but of indifference bands…"

The intuition for the property is that, in an observationally homogenous population, an increase of income from y_1 to y_2 (say) may not make everybody better off. This is because this income increase may, for some consumers, not be sufficient for them to switch to a new commodity group, or be able to buy another indivisible consumer good that makes them better off. (See Patel and Subrahmanyam, 1978, for a similar argument). It is important to realize that the notion of randomness and indifferences with respect to small changes in income is meant to represent a consumer's *typical* behavior in choice situations. For example, when asked if one dollar more a day is better than status quo in a SP questionnaire, most persons will probably answer yes. However, a question like that will be misleading because it is not put in the appropriate context, namely in *typical* daily life choice settings. Quandt's point is that in daily life behavior, few persons may care about having a few dollars more or less. Note that our stochastic framework also allows for the following interpretation, on the individual level: an individual that participates in a replication of a choice experiment may in some cases be indifferent between y_1 and y_2, and in other cases strictly prefer y_2 over y_1.

The next condition is a mathematical regularity condition.

Condition 1

The conditional indirect utility processes $\{\tilde{U}(j,y), y \geq \gamma\}$, $j \in \Omega$, *are separable and continuous in probability.*

The separability requirement is very weak and does not represent any essential restriction. We refer to textbooks in probability theory for a definition of this concept. The continuity requirement means that the probability that

$$\left|\tilde{U}(j,y_1) - \tilde{U}(j,y_2)\right| > \eta,$$

where $\eta > 0$ is an arbitrarily given number, decreases toward zero when $y_2 - y_1$ tends to zero. This means that when y_1 is close to y_2, there are "very few" sample paths where $\tilde{U}(j,y_1) - \tilde{U}(j,y_2)$ is "large".

Theorem 1

If condition 1 and Axioms 1 to 3 hold, then $\{\tilde{U}(j,y), y \geq \gamma\}$ *is an extremal process, that is, it can be represented as*

$$\tilde{U}(j,y_2) = \max\left(\tilde{U}(j,y_1), V_j(y_1,y_2)\right) \tag{2}$$

with $\tilde{U}(j,\gamma) = 0$, *where* $\{V_j(y_1,y_2), y_1 \leq y_2\}$ *are random variables such that* $V_j(y_1,y_2)$ *is independent of* $V_j(y_1',y_2')$ *if* $[y_1,y_2] \cap [y_1',y_2'] = \varnothing$, $V_j(y_1,y_1) = 0$ *and*

$$P\left(V_j\left(y_1,y_2\right)\le u\right)=\exp\left(-\frac{\left(\tilde{v}\left(j,y_2\right)-\tilde{v}\left(j,y_1\right)\right)}{u}\right) \tag{3}$$

for $u>0$ *and* $y_2\ge y_1$.

The proof of Theorem 1 is given in the Appendix. Note that if $\tilde{v}(j,y)=0$ for $y<\gamma$, then also $\tilde{U}(j,y)=0$.

The class of extremal processes is well known in the statistical literature and has been studied extensively by many authors, see for example Resnick (1987). At first glance, the result of Theorem 1 may seem strange. However, as demonstrated in the example below, the class of extremal utility processes can be given an intuitive behavioral interpretation.

Example:

In this example, we discuss a *direct utility* representation that is consistent with the result of Theorem 1. To this end let x denote a vector of quantities of a consumption bundle and let $U^*\left(x,j\right)$ denote the corresponding direct utility of (x,j) where j represents the set C_j, as discussed before. In modern markets and stores consumption bundles are usually made available as combinations of "packages" with given quantities. Thus, in this case the set of available quantities is countable. Let x_r, $r=1,2,...,$ be an arbitrary enumeration of the consumption bundles, and suppose that

$$U^*\left(x_r,j\right)=m\left(x_r\right)b(j)\varepsilon\left(r,j\right)$$

where *m(·)* and *b(·)* are positive deterministic functions and $\left\{\varepsilon\left(r,j\right)\right\}$ are i.i.d. random error terms with standard Fréchet cumulative distribution function (c.d.f.). The budget constraint is given by

$$px_r+q_j\le y$$

where *p* is a vector of goods prices and q_j is the cost of alternative *j*. As a result, the conditional indirect utility $\tilde{U}\left(j,y\right)$ can be written as

$$\tilde{U}\left(j,y\right)=\max_{r,\; px_r\le y-q_j} U^*\left(x_r,j\right)=b(j)\max_{r,\; px_r\le y-q_j}\left(m\left(x_r\right)\varepsilon\left(r,j\right)\right). \tag{4}$$

Define

$$V_j\left(y_1,y_2\right)=\max_{r,\; y_1<px_r+c_j\le y_2} U^*\left(x_r,j\right). \tag{5}$$

It follows immediately from (4) and (5) that

$$\tilde{U}(j, y_2) = \max\left(\tilde{U}(j, y_1), V_j(y_1, y_2)\right) \tag{6}$$

for $y_1 \leq y_2$. Furthermore, our distributional assumptions imply that $\tilde{U}(j, y_1)$ is independent of $V_j(y_1, y_2)$ and that $V_j(y_1, y_2)$ has Fréchet c.d.f. Thus, we have shown that extremal utility processes have an intuitive interpretation.

As above, let *B* be the agent's choice set which we assume belongs to $\mathcal{B}$, and let

$$U(y) = \max_{c_j \leq y,\, j \in B} \tilde{U}(j, y).$$

The process $\{U(y),\, y \geq \gamma\}$ is the utility-of-income (indirect utility) process. Although *U(y)* depends on *B*, we drop *B* in the notation for simplicity.

The next result is immediate.

Corollary 1

The utility of income is an extremal process that can be expressed as

$$U(y_2) = \max\left(U(y_1), V(y_1, y_2)\right) \tag{7}$$

for $y_1 \leq y_2$, *where*

$$V(y_1, y_2) = \max_{y_1 \leq c_j \leq y_2,\, j \in B} V_j(y_1, y_2).$$

The c.d.f. of $V(y_1, y_2)$ *is given by*

$$P\left(V(y_1, y_2) \leq u\right) = \exp\left(-\frac{\left(v(y_2) - v(y_1)\right)}{u}\right)$$

for $u > 0$, *where*

$$v(y) = \sum_{c_j < y,\, j \in B} \tilde{v}(j, y).$$

Since $v(\gamma) = 0$, it follows from (2) and (3) that

$$P\left(U(y) \leq u\right) = \exp\left(-\frac{v(y)}{u}\right)$$

for $u > 0$, which means that we can write

$$U(y) = v(y)\varepsilon(y)$$

where $\varepsilon(y)$ is standard Fréchet distributed.

3.2. Functional form of the deterministic part of the utility function

In this section we postulate an axiom that enables us to derive important restrictions on the functional form of the deterministic part of the utility of income.

Axiom 4

Suppose that y_1, y_2, y_1^*, y_2^* *are equal to or greater than* γ *and such that*

$$P\left(U\left(y_2\right) > U\left(y_1\right)\right) < P\left(U\left(y_2^*\right) > U\left(y_1^*\right)\right).$$

Then for all $\lambda > 0$

$$P\left(U\left(\lambda\left(y_2-\gamma\right)+\gamma\right) > U\left(\lambda\left(y_1-\gamma\right)+\gamma\right)\right) < P\left(U\left(\lambda\left(y_2^*-\gamma\right)+\gamma\right) > U\left(\lambda\left(y_1^*-\gamma\right)+\gamma\right)\right).$$

The interpretation of Axiom 4 is that if the fraction of consumers that strictly prefers y_2 to y_1 is less than the fraction of consumers that strictly prefers y_2^* to y_1^*, then this inequality does not change when all incomes beyond the subsistence level are multiplied by an arbitrary positive constant λ. The intuition is as follows: associate the different income levels y_1, y_2, y_1^*, y_2^* with consumption profiles 1, 2, 1^*, 2^* (when prices are given) and suppose the fraction of individuals that prefer consumption profile 2 over 1 is less than the fraction of individual that prefer 2^* over 1^*. To the consumers income beyond subsistence matters to some extent in the sense that a scale transformation of the respective incomes beyond subsistence will affect utility levels, but not in such a way that the fraction of consumers that prefer consumption profile 2 over profile 1 will be greater than the fraction of consumers that prefer consumption profile 2^* over profile 1^*. Recall that in our setup the probability that $U\left(y_2\right) > U\left(y_1\right)$, for $y_2 > y_1$, is *not* equal to one because there is a positive probability that $U\left(y_2\right) = U\left(y_1\right)$. We realize that if satiation can happen then evidently Axiom 4 may not hold. Dagsvik and Røine (2005) have carried out tests of Axioms 3 and 4 based on SP data and found that these axioms are supported by the data.

Theorem 2

Assume that Condition 1 and Axioms 1 to 4 hold and that v(·) *defined in Corollary 1 is continuous and strictly increasing in y, $y \geq \gamma$. Then* v(y) *has the structure*

$$\mathrm{v}(y) = \kappa exp\left(\delta\left(\frac{(y-\gamma)^{\tau}-1}{\tau}\right)\right) \tag{8}$$

for $y \geq \gamma$, where τ and $\delta > 0, \kappa > 0$ are constants.

The proof of Theorem 2 is given in the Appendix. Note that the parameter τ is allowed to be negative.

Axiom 5

For any $y_2 > y_1$, and $\lambda > 0$

$$P\big(U(y_2) > U(y_1)\big) = P\big(U(\lambda(y_2-\gamma)+\gamma) > U(\lambda(y_1-\gamma)+\gamma)\big).$$

Axiom 5 is stronger than Axiom 4, and it means that income beyond subsistence level is perceived in a strict relative sense, that is, the fraction of consumers that are better off when incomes beyond subsistence is increased from $y_1 - \gamma$ to $\lambda(y_1 - \gamma)$ and $y_2 - \gamma$ to $\lambda(y_2 - \gamma)$ is independent of λ. Note that this property is *not* implied by Axiom 4.

Theorem 3

Assume that Condition 1 and Axioms 1 to 3 and 5 hold and that v(·) *is continuous and strictly increasing in y, $y \geq \gamma$. Then* v(*y*) *has the structure*

$$v(y) = \kappa(y-\gamma)^{\delta} \tag{9}$$

for $y \geq \gamma$, and $\delta > 0, \kappa > 0$

A proof of Theorem 3 is given in the Appendix.

We note that (9) is obtained as a special case of (8) when $\tau \to 0$.

4. Heterogeneity in preferences

In the empirical specification to be described in Section 6 below, we shall introduce observed covariates that may affect the individual's evaluation of income. These observed covariates may capture some of the heterogeneity in the population, but obviously not all. To account for the remaining unobserved heterogeneity, we will introduce an individual specific effect, known to the agent but not to the analyst. Specifically, we shall assume that the systematic part of the utility function contains a positive multiplicative component that is a constant for each individual agent but varies across the population according to some probability distribution (random effect). Thus, the utility function, modified to include this random effect becomes

$$U(y) = Wv(y)\varepsilon(y), \tag{10}$$

where W is the random effect. Note that the way we include *W* is analogous to allowing for an additive constant term in an additive separable utility representation (which is seen by taking logarithm in (10)). This multiplicative random effect is motivated by the functional form given in (8), with κ containing a multiplicative random effect. Recall that W is irrelevant for individual choice behavior since it cancels out in utility comparisons. However, it matters in our context in which data are generated by the Leyden School type of SP data that yield information about utility evaluations across individuals.

In this section, we shall propose a theoretical justification for the distribution of *W*. For the sake of notational precision in the following Axiom, let us introduce individual specific notation, i.e,. let $U_i(y_i) = W_i v(y_i)\varepsilon_i(y_i)$, be the utility of agent *i*.

Axiom 6

Let the incomes of every individual in the population S be given. Then

$$P\left(U_i(y_i) = \max_{r\in S} U_r(y_r) \middle| \max_{r\in A} U_r(y_r) = \max_{r\in S} U_r(y_r)\right) = P\left(U_i(y_i) = \max_{r\in A} U_r(y_r)\right)$$

for $A \subset S$.

The statement in Axiom 6 says that the probability that individual *i* has the highest utility in *S*, given that this individual belongs to a subset *A* that contains the individual with the highest utility, is equal to the probability that *i* has the highest utility within *A*. In other words, given that the highest ranked individual belongs to *A*, information about the ranking of the individuals within *S\A* is irrelevant for assessing who is the highest ranked individual in *A*. We recognize Axiom 6 as a particular version of IIA. We note that Axiom 6 requires that individual utilities can be compared and ranked.

Theorem 4

Assume that Condition 1 and Axioms 1 to 4 and 6 hold, and that W_i *and* $\varepsilon_i(y_i)$ *are independent,* $i \in S$. *Then the distribution of* W_i *is strictly* α*-stable and totally skew to the right with* $\alpha < 1$.

The proof of Theorem 4 is given in Appendix.

Recall that the family of α-stable distributions, often denoted by $\{S_\alpha(c, \beta, \mu)\}$, is characterized by four parameters, namely (α, c, β, μ), where $0 < \alpha \le 2$ represents the tail thickness and is called the *characteristic exponent*, $c > 0$, is a scale parameter, $\beta \in [-1,1]$ is a skewness parameter and μ is a location parameter. When $\alpha = 2$, one obtains as a special case the normal distribution. It is strictly α-stable when $\mu = 0$ and totally skew to the right when $\beta = 1$. When $\alpha \le 1$, neither the variance nor the mean of the stable random variable exist. When $\mu = 0, \alpha < 1$ and $\beta = 1$ the probability that the stable random variable attains non-positive values is zero. (See Samorodnitsky and Taqqu, 1994).

As mentioned in Section 3, the choice among characteristics will, under Axiom 1, satisfy IIA. This is still true if (10) holds because the random effect *W* vanishes in utility comparisons. In our context, IIA does not seem overly restrictive since the characteristics have not been given an explicit empirical content. It is however possible to motivate more general representations of unobserved heterogeneity. This extension consists in assuming that the utility representation has the form

$$\tilde{U}(j,y) = \tilde{W}(j)\tilde{v}(j,y)\tilde{\varepsilon}(j,y), \tag{11}$$

where $\{\tilde{W}(j), j \in \Omega\}$ are strictly stable processes that are totally skew to the right with $\alpha < 1$. It can be demonstrated that (11) implies that the choice of characteristics model will have a Generalized Extreme Value (GEV) structure. According to Dagsvik (1994, 1995) the GEV model represents in practice no restriction on the general random utility model. Our conjecture is that Axiom 6 implies that $\{\tilde{W}(j), j \in \Omega\}$ is a stable process. However, we have so far only been able to prove that the one-dimensional marginal distributions of this process are stable. For simplicity, in this paper we have chosen to base our empirical model on the special case (10) rather than on (11).

5. The random Expenditure Function

Let the random expenditure functions $\{Y(u), u > 0\}$ be defined by

$$Y(u) = \min\{y : U(y) \ge u\}.$$

Due to the fact that the indirect utility function *U(y)* is a stochastic process with parameter *y* defined in Theorem 1, we have the following results.

Theorem 5

Assume that (2) and (3) hold. For $0 < u_1 \le u_2 \le \ldots \le u_m$, *and* $\gamma \le y_1 \le y_2 \le \ldots \le y_m$, *we have*

$$G_m(y_1, y_2, \ldots, y_m) \equiv P\big(Y(u_1) > y_1, Y(u_2) > y_2, \ldots, Y(u_m) > y_m\big) = E\, exp\big(-WH(y_1, y_2, \ldots, y_m)\big)$$

where

$$H(y_1, y_2, ..., y_m) = v(y_m)u_m^{-1} + \sum_{j=1}^{m-1} v(y_j)(u_j^{-1} - u_{j+1}^{-1}).$$

For $y_1 < y_2 < ... < y_m$ *and* $u_1 \le u_2 \le \cdots \le u_m$ *, the corresponding joint density of* $(Y(u_1), Y(u_2), ..., Y(u_m))$ *exists and is equal to*

$$g_m(y_1, y_2, ..., y_m) = v'(y_m)u_m^{-1} \prod_{j=1}^{m-1} (u_j^{-1} - u_{j+1}^{-1}) v'(y_j) \, E(W^m exp(-WH(y_1, y_2, ..., y_m))).$$

A proof of Theorem 5 is given in Appendix.

It turns out to be convenient to normalize the scale parameter *c* in the stable distribution of the random effect *W* such that

$$c^\alpha = 1/\cos(\alpha\pi/2). \tag{12}$$

This is purely a matter of convenience and represents no loss of generality since the scale parameter in the distribution of *W* cannot be identified.

The next lemma is essential for calculating $E(W^m \exp(-WH(y_1, y_2, \cdots, y_m)))$ when $m=6$.

Lemma 1

Let W be α-stable, $S_\alpha(c,1,0)$ *with* $\alpha < 1$ *and c given in* (12). *Let* $\psi(\lambda;\alpha)$ *be defined as*

$$\psi(\lambda;\alpha) \equiv E(W^6 exp(-\lambda W))$$

for $\lambda \ge 0$. *Then for* $\lambda > 0$,

$$\begin{aligned}\psi(\lambda;\alpha) = E(W^6 exp(-\lambda W)) = \Big[& (\alpha\lambda^{\alpha-1})^6 + 15(1-\alpha)\alpha^5\lambda^{5\alpha-6} \\ & +5(13\alpha^2 - 10\alpha + 17)\alpha^4\lambda^{4\alpha-6} + 15(-6\alpha^3 + 25\alpha^2 - 36\alpha + 15)\alpha^3\lambda^{3\alpha-6} \\ & +(31\alpha^4 - 225\alpha^3 + 595\alpha^2 - 675\alpha + 274)\alpha^2\lambda^{2\alpha-6} \\ & +(-\alpha^5 + 15\alpha^5 - 85\alpha^5 + 225\alpha^5 - 274\alpha^5 + 120)\alpha\lambda^{\alpha-6} \Big] exp(-\lambda^\alpha).\end{aligned} \tag{13}$$

Proof:

From the properties of α-stable distributions it follows that

$$E\exp(-\lambda W) = \exp(-\lambda^\alpha). \tag{14}$$

By differentiating (14) six times with respect to λ we get (13).

Q.E.D.

Corollary 2

The structure of $G_2(y_1, y_2)$ *implies that we can write*

$$Y(u_1) = \min\left(Y(u_2), v^{-1}\left(\frac{u_1 u_2\, \eta(u_1, u_2)}{(u_2 - u_1)W} \right) \right) \tag{15}$$

for $u_2 > u_1$*, where* $\eta(u_1, u_2)$ *is a random variable, which is exponentially distributed with parameter equal to one, and is independent of W and* $Y(u_2)$.

The proof of Corollary 2 follows readily since any finite dimensional marginal distribution functions of the process $\{Y(u), u > 0\}$, given by (15), are the same as the ones given by Theorem 5.

Corollary 3

The structure of $G_1(y)$ implies that we can write

$$v(Y(u)) = u\eta(u)^{1/\alpha}$$

for $u > 0$*, where* $\eta(u)$ *is a random variable that is exponentially distributed with parameter equal to one.*

The proof of Corollary 3 is given in the Appendix.

6. Empirical specification

Consistent with the result of Theorem 5, let $W \sim S_\alpha(c,1,0)$, where c is given in (12). In the following it will be convenient to reparametrize the model by introducing $\{a_j\}$ defined by

$$a_j = -\log\left(\frac{1}{u_j} - \frac{1}{u_{j+1}} \right) \tag{16}$$

for $j \le 5$ and $a_6 = \log u_6$, where $\{u_j\}$ are unknown utility threshold levels associated with the ordered structure of the income questionnaire we are using. We suspect that households may have different threshold levels. We allowed initially a_j to depend on selected household specific characteristics such as income, debt and family size, etc. However, the estimation results indicated that only income seemed to have a significant effect on these threshold levels. Motivated by these preliminary results, we assume that $a_j = d_j + t\log(I)$, where *I* denotes the current household income level. From Theorem 5, (8) (with $\delta = 1/\sigma$) and Lemma 1 we get that

$$G_6\left(y_1,y_2,...,y_6 \mid \{a_j\}\right) \equiv P\left(Y(u_1)\geq y_1, Y(u_2)\geq y_2,...,Y(u_6)\geq y_6\right)$$

$$= \exp\left[-\left\{\sum_{j=1}^{6}\exp\left(\frac{(y_j-\gamma)^{\tau}-1}{\tau\sigma}-a_j\right)\right\}^{\alpha}\right], \tag{17}$$

and

$$g_6\left(y_1,y_2,...,y_6 \mid \{a_j\}\right)$$

$$=\prod_{j=1}^{6}\left(\exp\left(\frac{(y_j-\gamma)^{\tau}-1}{\tau\sigma}-a_j\right)\frac{(y_j-\gamma)^{\tau-1}}{\sigma}\right)\cdot\psi\left(\sum_{j=1}^{6}\exp\left(\frac{(y_j-\gamma)^{\tau}-1}{\tau\sigma}-a_j\right);\alpha\right), \tag{18}$$

where $\gamma < y_1 < ... < y_6$. Unfortunately, the functional form of the density function in (18) implies that standard conditions for maximum likelihood estimation are not fulfilled. The difficulty here is that the support of the density in (18) depends on parameters of the specification of the subsistence level γ. In the simple case where subsistence level is a constant, the global maximum of the likelihood function is achieved at $\hat{\gamma} = \min(Y_i(u_1))$ and $\tau = 0$. However, there may be an additional local maximum. Barnard (1965) has given reasons for ignoring the singular solution and settle for the local one (if it exists). In the statistical literature the estimation of densities like (18) when γ is unknown does not seem to have been analyzed. Zanakis and Kyparisis (1986), and Smith (1994) only discuss special cases of (18). In the estimation procedure below we have therefore chosen to specify the subsistence level a priori. In particular, we assume that the subsistence level γ for the household is defined as $\gamma = \sqrt{N}\cdot 30000$, where N is the household size. Here, 30000 NOK (as of 1995, one USD is approximately equal to 7.20 Norwegian Kroner) is assumed to be the subsistence level for a single individual, and $\sqrt{N}$ is used as the household equivalent scale. This type of household equivalence scale is used in many countries.

7. Data and parameter estimates

In September 1995 a questionnaire was distributed to 569 employees at Statistics Norway and the staff at the Department of Economics, University of Oslo. It contained questions concerning the social background of the respondent, including income and wealth, and the income evaluation question similarly to the ones of the Leyden school quoted in Section 2. Let $\{y_j\}$ denote the individual's answers of the questionnaire, that is (y_1, y_2), (y_2, y_3), …, (y_5, y_6) are the income intervals that correspond to the IEQ questionnaire discussed in section 2[6]. Consistent to the theoretical setup above, we assume that the observed income levels are related to the utility levels through the expenditure function as

$y_j = Y\left(u_j\right)$ where u_j is the underlying utility level that corresponds to income y_j. That is, y_j is the lowest disposable income needed to achieve utility level u_j. The response rate was slightly above 50 per cent. 250 of those who responded were able to fill in answers on all the income intervals in the income evaluation questionnaire, with $y_{j+1} > y_j$ for all *j* and positive reported household income.

Table 1 around here.

Obviously, this sample is not representative for the Norwegian population. The majority of the respondents are individuals with high education. In addition they work in similar public institutions and therefore have similar incomes. Table 2 gives a summary picture of some basic characteristics of the sample. The loglikelihood function is obtained from (18) by inserting the respective observed individual levels $\mathrm{Y}(\mathrm{u}_\mathrm{j})$, *j*= 1, 2,…,6, for each individual in the sample. The estimation results are reported in Table 3. We observe that the parameters are sharply determined. Although the Box-Cox exponent τ is significantly different from zero, the estimate of the exponent is only slightly above zero. Thus, the deterministic part of the utility function is found to be approximately a power function, given the family of functions specified in (8). Our results thus are consistent with the assumption of Theorem 3. The characteristic exponent α is significantly below 1 as expected and thus neither the mean nor the variance exists in the distribution of *W*. As to be expected the constants $\{d_j\}$ are increasing with utility levels. The coefficient t is estimated to be positive and significantly different from zero, which means that the threshold levels are increasing with the actual income of the household.

Table 2 around here.

As mentioned in Section 2, the Leyden School approach is based on the assumption that individuals partition the income range according to equal quantiles of the utility function, referred to as EQA. Buyze (1982) tested this assumption empirically and concluded that the assumption holds approximately. Within our framework it is also possible to test EQA. Specifically, we have used the likelihood ratio test to test whether or not EQA is rejected for our dataset. Note that in our setting, our utility is measured in ratio scale so the equal quantile assumption implies:

$$\frac{U(y_{j+1})}{U(y_j)} = \frac{u_{j+1}}{u_j} = \frac{U(y_{j+2})}{U(y_{j+1})} = \frac{u_{j+2}}{u_{j+1}} = \rho, \quad \text{for j} = 1, \cdots, 4$$

From Table 3 we see that twice the difference between the a priori loglikelihood function and the loglikelihood function under EQA equals 3.4. The corresponding critical value of the Chi-square

---

[6] We have dropped the indexation for the individuals for notation simplicity.

distribution with four degrees of freedom equals 9.5. Thus we cannot reject EQA. We observe that the estimates of the key parameters $\{\tau, \sigma, \alpha\}$ are the same in the two models, the general model and the model under EQA.

Table 3 around here.

It is interesting to compare our results with the evidence from the Leyden School. To this end, it is useful to use the result of Corollary 3. Given that $\tau \to 0$, Corollary 3 implies that

$$\log(Y(u_j) - \gamma) = \sigma \log u_j - \frac{\sigma}{\alpha}\theta_j = \sigma s_j + \sigma t \log I - \frac{\sigma}{\alpha}\theta_j, \tag{19}$$

where $\theta_j = \log \eta(u_j)$ and

$$s_j = -\log(\sum_{k=j}^{6} \exp(-d_k)).$$

Moreover, Corollary 3 implies that

$$P(\theta_j \le x) = \exp(-\exp(-x))$$

for $x \in R$. It is well known that in this case, $E\theta_{ij} = 0.5772$ (Euler's constant) and $\operatorname{var}\theta_{ij} = \pi^2 / 6$. The relation (19) is similar to the corresponding empirical relation in the Leyden School approach apart from the subsistence level γ and the distributional properties of the error term. Our distributional assumptions also constrain the expenditure function $\{Y(u)\}$ to be nondecreasing with probability one. Moreover, the random effect W (cf. Corollary 2) implies that the respective increments in expenditures, for a given individual become correlated. From the estimates in Table 3, we find that the coefficient $\sigma t / \alpha$, associated with *logI* is estimated to be 0.252 with standard error approximately equal to 0.03. In the Leyden School approach estimates of this coefficient range from 0.53 to 0.68; see van Hervaarden and Kapteyn (1977), Table 4. Thus, our estimate of the preference drift parameter is significantly lower than the ones obtained by the Leyden School.

8. Testable properties of the model

Let

$$F(y_1, y_2, \cdots, y_6) = -\log G_6((y_1 \cdot I^{t\sigma} + \gamma), (y_2 \cdot I^{t\sigma} + \gamma), \cdots, (y_6 \cdot I^{t\sigma} + \gamma) \mid \{a_j\}),$$

where G_6 is given in (17) with $\tau = 0$. It follows that

$$F(y_1, y_2, \cdots, y_6) = (\sum_{j=1}^{6} y_j^{1/\sigma} \exp(-d_j))^{\alpha}. \tag{20}$$

We note that F is independent of I and $\{a_j\}$, and hence becomes equal for all households. Furthermore, define

$$F_j(y) = F(0,\cdots,0,\underset{j_{th}\text{ component}}{y},y,\cdots,y),$$

for $1 \le j \le 6$. The function $F_j(y)$ has the interpretation

$$F_j(y) = -\log P(\frac{Y(u_j)}{I^{t\sigma}} - \gamma > y).$$

It follows from (20) that

$$F(y_1, y_2, \cdots y_6)^{\frac{1}{\alpha}} = \sum_{j=1}^{6} A_j y_j^{\frac{1}{\sigma}}, \tag{21}$$

and

$$F_j(y)^{\frac{1}{\alpha}} = y^{\frac{1}{\sigma}} \sum_{k \ge j} A_k, \tag{22}$$

where $A_j = \exp(-d_j)$. From (22) we get that

$$\sum_{j=1}^{6} A_j y^{\frac{1}{\sigma}} = \sum_{j=1}^{6} (F_{j+1}(y_j)^{\frac{1}{\alpha}} - F_j(y_j)^{\frac{1}{\alpha}}) = \sum_{j=2}^{6} F_j(y_{j-1})^{\frac{1}{\alpha}} - \sum_{j=1}^{6} F_j(y_j)^{\frac{1}{\alpha}},$$

which combined with (21) yields

$$F(y_1, y_2, \cdots y_6)^{\frac{1}{\alpha}} = \sum_{j=2}^{6} F_j(y_{j-1})^{\frac{1}{\alpha}} - \sum_{j=1}^{6} F_j(y_j)^{\frac{1}{\alpha}}. \tag{23}$$

Recall that when s is known then $F(y_1, y_2, \cdots y_6)$ and $F_j(y)$ are *observable*. In this case, (23)implies a testable property, namely that (23) must hold for some positive constant $\alpha \le 1$. When $\{a_j\}$, σ and α have been estimated one can also use (21) to test the model. Analogous to (23), (21) implies that $F(y_1, y_2, \cdots y_6)^{1/\alpha}$ is linear and additive separable in $A_j y_j^{\frac{1}{\sigma}}$, j=1,2,…,6. Furthermore we get from (22)

$$\log(F_j(y)) = \frac{\alpha}{\sigma}\log(y) + a\log(\sum_{k=j}^{6} A_k). \tag{24}$$

It also follows that

$$\log(\sum_{j=1}^{6} F_j(y)) = \frac{\alpha}{\sigma}\log(y) + Q, \tag{25}$$

where $Q = \log(\sum_{j=1}^{6}(\sum_{k=j}^{6} A_k)^{\alpha})$.

Figure 1 around here.

Motivated by (25), a check of the model fit can be obtained by plotting

$$K(y_r) \equiv \log(\sum_{j=1}^{6} F_j(y_r))$$

against $\log(y_r)$, for suitable y_r, r=1,2,$\cdots$M. We have plotted $K(y)$ against log(y), which is shown in Figure 1. As we can see, the plot does not deviate much from a linear relationship between $K(y)$ and $\log(y)$, which means that our model fits the data quite well.

Since this sample is rather small, we have not tried to construct tests based on (21) and (23).

9. Measuring the utility of changes in income

As mentioned in the introduction and discussed at length in Ellingsen (1994), several authors have discussed the concept of marginal utility of income. This concept is intrinsically linked to the concept of cardinal utility. What seems to have been overlooked in the literature is that economists have implicitly assumed utility to be additive when they discuss cardinal utility. For example, when Frisch postulated his interlocal choice axioms (Frisch, 1926), he assumed implicitly that changes in utility were represented as utility differences. However, there are no a priori theoretical reasons why changes in utility should be represented as utility differences rather than utility ratios. Now consider the setting of this paper. Suppose there are two multiplicatively separable random utility representations $\{\tilde{U}(j,y)\}$ and $\{\tilde{U}^*(j,y)\}$ where $\tilde{U}(j,y) = \tilde{v}(j,y)\tilde{\varepsilon}(j,y)$ and $\tilde{U}^*(j,y) = \tilde{v}^*(j,y)\tilde{\varepsilon}^*(j,y)$, that yield the same choice probabilities. Then it follows from Yellott (1977), Theorem 3, that

$$\tilde{U}(j,y) \stackrel{d}{=} a\tilde{U}^*(j,y)^b \tag{26}$$

where $\stackrel{d}{=}$ means equality in distribution and a and b are arbitrary positive constants[7]. Thus we conclude from (26) that our family of utility functions is represented on a so-called *log interval scale*, see Falmagne (1985) for a definition of different scale types. It is a random scale due to the stochastic error term. From (26) it follows moreover that when aggregating over the discrete goods, the random utility of income {U(y)} is also represented as a random log interval scale. This also implies that the average scale v(y) is also represented as a log interval scale. Now define $U^{**}(y) = \log U(y)$. As mentioned above, the utility function $U^{**}(y)$ is completely equivalent to U(y) in the sense that it can also be used as a scale that measures changes. The only difference is that $U^{**}(y)$ yields an interval

[7] The only difference between Yellott's formulation and ours is that he assumes an additive utility representation while we use a multiplicative one. These two representations are of course completely equivalent since the additive representation is obtained from the multiplicative one by taking logarithm of the multiplicative utility representation

scale representation instead of a log interval scale one. Similarly, the representative scale v(y) is transformed to an interval scale logv(y).

Consider next scale representation of changes in incomes. The following result is useful in the context of interpreting utility ratios.

Corollary 4

Assume that the assumptions of Theorem 2 hold. Then for $y_2 > y_1$,

$$\left(\frac{U(y_2)}{U(y_1)}\right)^b = \max\left(\left(\frac{v(y_2)}{v(y_1)} - 1\right)^b Z(y_1, y_2)^b, 1\right), \tag{27}$$

where $Z(y_1,y_2)$ *is a positive random variable with c.d.f.*

$$P(Z(y_1, y_2) \le z) = z/(1+z)$$

for z>0, *Moreover,*

$$E\left(\frac{U(y_2)}{U(y_1)}\right)^b = \begin{cases} M\left(\frac{v(y_2)}{v(y_1)}\right) & \text{if } b < 1 \\ \infty & \text{if } b \ge 1 \end{cases}, \tag{28}$$

where M(x) is a positive function that is strictly increasing.

Also, we have the following result,

Corollary 5

Suppose $y_j > \gamma$ *, j=1,2,3,4, are incomes that satisfy* $v(y_2)/v(y_1) < v(y_4)/v(y_3)$. Then $U(y_2)/U(y_1)$ *is stochastically dominated by* $U(y_4)/U(y_3)$*, i.e., for q > 1,*

$$P\left(\frac{U(y_2)}{U(y_1)} \le q\right) > P\left(\frac{U(y_4)}{U(y_3)} \le q\right).$$

The proofs of Corollaries 4 and 5 are given in the Appendix.

The result of Corollary 5 shows that when $v(y_2)/v(y_1) < v(y_4)/v(y_3)$, a population of consumers will, on average, assign higher value to $U(y_4)/U(y_3)$ than to $U(y_2)/U(y_1)$. The results of Corollary 4 and 5 demonstrate that the c.d.f. of the utility ratio depends on (y_1, y_2) through $v(y_2)/v(y_1)$. If we take the logarithm transformation of the corresponding utility ratios we obtain results that are completely analogous to the results of Corollaries 4 and 5, where the relevant c.d.f. in this case depends on the logarithm of the ratios of the representative utilities. Thus, on the individual level, the utility ratios can be transformed to perfectly equivalent utility differences. Similarly, on the aggregate level the ratios of the representative utilities transforms to corresponding representative utility differences. Thus, we have demonstrated that Theorem 1 yields either an interval scale utility representation or an equivalent log interval scale representation. Torgerson (1961, pp.202-203) and Narens (1996, pp. 117-118) have reached a similar conclusion. In fact Narens (1996) provides a theoretical basis for the empirical findings of Torgerson (1961). When utility is represented either on a log interval scale or on an interval scale we shall say that we have a *weak cardinal representation*. When utility is solely represented on an interval scale we shall say that we have a *strong (interval scale) cardinal representation*. When utility is solely represented on a ratio scale we shall say that we have a *strong (ratio scale) representation*. In the present context, weak cardinality is actually all we need because it enables us to rank both *levels* and *differences*, in contrast to an ordinal scale. However, in the context of choice under uncertainty and expected utility theory we need strong (interval scale) cardinality.

As demonstrated above, our weak cardinal scale allows both an individual random scale as well as a deterministic average scale representation that both are weakly cardinal. Hence, this setting allows us to rank levels and differences both on the individual as well as on the aggregate level.

In psychophysics, as in economics, the concept of cardinality seems to be controversial. For example, the school of Stevens claims that the power function representation is the appropriate strong (ratio scale) cardinal psychophysical law, of which utility of income is a special case. In the literature, several researchers have disagreed with Stevens on this matter; see for example Shepard (1981). Recall that in the typical experimental settings described in section 2, such as for example Magnitude estimation, the subjects have proven to be able to "produce" numbers on a ratio scale that matches *changes* in intensities of stimuli. To us it therefore seems plausible that Fechner's logarithm law and Steven's power law can at least allow the interpretation of scales that can measure utility of changes. This is possible if we interpret these laws as weakly cardinal.

As regards to a marginal utility concept one cannot apply the conventional definition to the random utility function of income simply because U(y) is not differentiable with respect to y. This is seen immediately from (7). One can, however, define the corresponding aggregate marginal utility. From (28) and (A.56) in the Appendix, if follows that

$$\lim_{y_2\to y_1} E\left(\frac{U(y_2)^b - U(y_1)^b}{(y_2-y_1)U(y_1)^b}\right) = \frac{v'(y_1)}{v(y_1)}\cdot \lim_{x\to 1} M'(x) = \frac{v'(y_1)}{v(y_1)}\cdot\frac{b}{1-b} \quad . \tag{29}$$

Alternatively, we can choose to use the equivalent interval scale representation in which case one get:

$$\lim_{y_2\to y_1}\frac{bE\log U(y_2) - bE\log U(y_1)}{(y_2-y_1)} = \lim_{y_2\to y_1}\frac{b\log v(y_2) - b\log v(y_1)}{y_2-y_1} = b\frac{v'(y_1)}{v(y_1)}. \tag{30}$$

From (29) and (30) we see that whether we use the log interval or the interval scale representation the aggregate marginal utility concept introduced above is only determined up to a multiplicative constant (b/(1-b) or b). Moreover, we notice that (29) and (30) are equivalent (equal apart from a multiplicative constant). Let $\breve{\omega}$ (y) denote the elasticity of the marginal utility of income with respect to income, which Frisch (1959) called the money flexibility. An immediate consequence of (29) and (30) is that the elasticities of the two versions of the aggregate marginal utility of income are equal.

Corollary 6

Assume a weakly cardinal representation, i.e., a log interval or an interval random scale representation consistent with Theorem 1. Then the aggregate marginal utility function is uniquely defined by

$$\breve{\omega}(y) = \frac{d \log(v'(y)/v(y))}{d \log y} \tag{31}$$

for $y > \gamma$.

From (31) and (8), we obtain that

$$\breve{\omega}(y) = -\frac{1-\tau}{1-\gamma y^{-1}} \tag{32}$$

for $y > \gamma$. Because $\tau \le 1$ and $\gamma \le y$, the money flexibility is negative, it approaches $-\infty$ when $y \to \gamma$ and $(\tau - 1)$ when $y \to \infty$. Recall that our estimation results suggest that τ is close to zero.

From (32) we observe that $|\breve{\omega}|$ declines with income, as Frisch (1959) suggested. Hence $|\breve{\omega}|$ becomes infinitely large when y approaches γ (the subsistence level) and it approaches 1 (τ=0) when y increases toward large values.

In his well known article on consumer demand Frisch (1959) presented a complete scheme for computing all direct and cross demand elasticities. He employed a deterministic additive separable utility function, where each element gave the utility of a good. Based on this separability assumption, (by Frisch called want-independence), he demonstrated that all elasticities with respect to price could be deduced from the knowledge of budget proportions and Engel (income) elasticities. The money flexibility, $\breve{\omega}$, had an essential role in the formulas for these elasticities.

Johansen (1960) provides the first example of a computable general equilibrium model (CGE) in economics. It was estimated and calibrated on Norwegian data. Based on demand data for different goods he used the approach of Frisch (1959) and obtained very similar results for the different goods[8].

---

[8] The formula for $\breve{\omega}$ used in Frisch (1959) was $\breve{\omega} = E_r\left(1-\alpha_r E_r\right)/(e_{rr}+\alpha_r E_r)$ where E_r, α_r and e_{rr} are the income elasticity, budget share and direct price elasticity for commodity *r*, respectively.

The estimates of $\breve{\varpi}$ varied from –1.85 to –2.13 and Johansen concluded that the compromise value should be –1.89, which he then also used in his CGE-model.

Frisch concluded that ϖ was equal to –2 for the median part of the population (the middle class), in absolute values much higher for the poor and very small for the rich.

Table 4 around here

We have simulated the distribution of ϖ (y) based on a large population that is representative for the population of Norway. The summary statistics of the population used is given in Table 4. We find the mean of ϖ is equal to -1.7, which is quite close to the number suggested by Frisch and Johansen. The distribution of the predicted ϖ (with γ=30000 $\sqrt{N}$) is shown in Figure 2.

Figure 2 around here.

10. Relations to Demand Theory

We shall now discuss briefly some implications from the above analysis for consumer demand systems. Recall that since our theory implies a stochastic utility function the corresponding demand system will be stochastic. Conventional methods based on duality theory (Roy's identity) will not work here because the random error term in the utility function depends on income and prices. A more general and rigorous treatment can be made by applying the approach of Dagsvik (1994). This would, however, be beyond the scope of the present paper. Here we shall therefore ignore the random term in the utility function. Provided $\tau \to 0$, the corresponding "representative" utility is, apart from a power transformation, equal to the utility function $\bar{U}(y,p)$ given by

$$\bar{U}(y,p) = (y-\gamma(p))^{1/\sigma}\kappa(p)I^{t(p)} \tag{33}$$

where p denotes a vector of prices, y denotes income and I will be interpreted as the real income lagged one year. As discussed above the variable I implies a "drift" in the utility function. Note that the parameter $\kappa \succeq \kappa(p)$ that appears in (8) and may depend on prices, is absorbed into the constants a_j in the likelihood function. In principal, the parameter σ may also depend on prices. Below we only consider the case with constant σ . Note that $\kappa(p)$ and $I^{t(p)}$ must be homogenous of degree $-\sigma$ in prices. Given that σ is a constant, we observe from (33) that our indirect utility function belongs to the class of functional forms called "Gorman-Polar" form (Gorman (1953)). Applying Roy's identity $\bar{U}(y,p)$ we get the following demand system

$$x_k = \gamma_k(p) - \sigma(y-\gamma(p))\left(\frac{\kappa_k(p)}{\kappa(p)} + \log I \cdot t_k(p)\right); \quad k = 1,\cdots K, \text{ for } y > \gamma$$

where x_k is the consumption of good k, $k = 1,\cdots K$, and $\kappa_k(p)$ and $\gamma_k(p)$ are the derivative of $\kappa(p)$ and $\gamma(p)$ with respect to the price of good k. It is beyond the scope of the present article to discuss how $\kappa(p)$ and $\gamma(p)$ vary with prices. We notice that the Engel equations implied by (33) are linear in income y. However, due to the effect of preference drift, represented by the logarithm of income lagged one year, one may falsely interpret empirical evidence from panel data as an indication of nonlinear Engel functions because lagged income typically is highly correlated with current income. The derivation of the demand relations in the more general case when σ depends on prices is similar. In particular, when $\gamma_k(p)$=0 and σ, κ have suitable functional forms we obtain the AIDS model.

11. Conclusion

Utility theory represents a fundamental part of microeconomic theory. Yet, few researchers address the issue of establishing a theoretical framework for characterizing and measuring utility as a stochastic process in income.

In this paper we have proposed a set of behavioral axioms from which we derived a characterization of the utility of income, viewed as a stochastic process in income. Specifically, it turns out that the implied utility function is an *extremal process*.

Subsequently, we have specified an empirical model for the distribution of the utility of income process based on the theoretical characterization, and we have applied SP data to estimate the unknown parameters of the model. We demonstrate that the estimated model fits the data rather well. Within the framework developed in this paper the empirical results show that the utility function is consistent with the power law established by Stevens (1975). We discussed the concept of cardinality and marginal utility of income in our setting. We also demonstrated how our utility of income model can be employed to yield a consumer demand system.

Appendix

To prove Theorem 1, it will be convenient to prove the following Lemma first.

Lemma 2

Let $F_j(u_1,u_2)$ *be the joint c.d.f. of* $\left(\tilde{U}(j,y_1),\tilde{U}(j,y_2)\right)$ *for income level* $y_1 < y_2$ *, if Condition 1and Axioms 1 to 3 hold, then*

$$F_j(u_1,u_2)=\begin{cases}\exp\left(-v(j,y_1)u_1^{-1}-\left(v(j,y_2)-v(j,y_1)\right)u_2^{-1}\right) & \text{for} \quad u_1 \le u_2 \\ \exp\left(-v(j,y_2)u_2^{-1}\right) & \text{for} \quad u_1 > u_2 .\end{cases}$$

Proof of Lemma 2:

By Axiom 1, F_j is a bivariate extreme value (Fréchet) c.d.f. (See Resnick for a description of multivariate extreme value c.d.f.). Recall that in this case F_j has the property

$$\log F_j(u_1,u_2)=\frac{1}{z}\log F_j\left(\frac{u_1}{z},\frac{u_2}{z}\right) \tag{A.1}$$

for any $z>0, u_1>0, u_2>0$. The marginal c.d.f. of $\tilde{U}(j,y_1)$ and $\tilde{U}(j,y_2)$ are equal to

$$P\left(\tilde{U}(j,y_2)\le u_2\right)=F_j(\infty,u_2)=\exp\left(-v(j,y_2)u_2^{-1}\right), \tag{A.2}$$

and

$$P\left(\tilde{U}(j,y_1)\le u_1\right)=F_j(u_1,\infty)=\exp\left(-v(j,y_1)u_1^{-1}\right). \tag{A.3}$$

By Lemma 1, p. 827, in Dagsvik (2002) the left and right derivatives $\partial_+ F_j(u_1,u_2)/\partial u_k$, , $\partial_- F_j(u_1,u_2)/\partial u_k$, $k=1,2$ exist and are non-decreasing. From now on the notion "derivative" will mean the (first order) right derivative.

Define

$$\varphi_j(u)=-\log F_j\left(\frac{1}{u},1\right) \tag{A.4}$$

for $u \ge 0$. By (A.1) we have, with $z=u_2$,

$$\log F_j\left(u_1,u_2\right)=-u_2^{-1}\varphi_j\left(\frac{u_2}{u_1}\right). \tag{A.5}$$

Let ∂_r denote the partial derivative with respect to component r, $r=1,2$. We get from (A.5) that

$$\partial_1 \log F_j\left(u_1,u_2\right)=u_1^{-2}\varphi_j'\left(\frac{u_2}{u_1}\right) \tag{A.6}$$

and

$$\partial_2 \log F_j\left(u_1,u_2\right)=u_2^{-2}\varphi_j\left(\frac{u_2}{u_1}\right)-u_1^{-1}u_2^{-1}\varphi_j'\left(\frac{u_2}{u_1}\right). \tag{A.7}$$

Let J(y) denote the choice, given income y. We have that

$$\begin{aligned}
&P\left(J\left(y_1\right)=2,\, J\left(y_2\right)=1,\ \max_k \tilde{U}\left(k,y_2\right)\in\left(u,u+\Delta u\right)\right)\\
&=P\left(\tilde{U}(1,y_1)<\tilde{U}(2,y_1),\tilde{U}\left(2,y_2\right)<\tilde{U}\left(1,y_2\right)\in\left(u,u+\Delta u\right)\right)\\
&=\int_x P\left(\tilde{U}(1,y_1)<x,\tilde{U}(2,y_1)\in\left(x,x+dx\right),\tilde{U}\left(2,y_2\right)<\tilde{U}\left(1,y_2\right)\in\left(u,u+\Delta u\right)\right)dx
\end{aligned} \tag{A.8}$$

Due to the assumption of independent utilities across alternatives the above integral reduces to

$$\begin{aligned}
&\int_x P\left(\tilde{U}\left(1,y_1\right)<x,\tilde{U}\left(1,y_2\right)\in\left(u,u+\Delta u\right)\right)P\left(\tilde{U}\left(2,y_1\right)\in\left(x,x+dx\right),\tilde{U}\left(2,y_2\right)<u\right)\\
&=\Delta u\int_0^\infty \partial_2 F_1\left(x,u\right)\partial_1 F_2\left(x,u\right)dx+o(\Delta u).
\end{aligned} \tag{A.9}$$

From (A.5), (A.6) and (A.7) it follows that

$$\begin{aligned}
&\int_0^\infty \partial_2 F_1\left(x,u\right)\partial_1 F_2\left(x,u\right)dx=\int_0^\infty F_1\left(x,u\right)F_2\left(x,u\right)\cdot\partial_2\log F_1\left(x,u\right)\partial_1\log F_2\left(x,u\right)dx\\
&=\int_0^\infty \exp\left(-u^{-1}\left(\varphi_1\left(\frac{u}{x}\right)+\varphi_2\left(\frac{u}{x}\right)\right)\right)\left(u^{-2}\varphi_1\left(\frac{u}{x}\right)-x^{-1}u^{-1}\varphi_1'\left(\frac{u}{x}\right)\right)x^{-2}\varphi_2'\left(\frac{u}{x}\right)dx.
\end{aligned} \tag{A.10}$$

Let

$$\varphi(z)=\varphi_1(z)+\varphi_2(z)\,. \tag{A.11}$$

When we combine (A.9), (A.10) and make the change of variable $x=uw^{-1}$, in the last integral we obstain

$$P\left(J\left(y_1\right)=2, J\left(y_2\right)=1, \max_k \tilde{U}\left(k, y_2\right)\in\left(u, u+\Delta u\right)\right)$$
$$=u^{-3}\Delta u\int_0^\infty \exp\left(-u^{-1}\varphi(w)\right)\varphi_2'(w)\left(\varphi_1(w)-w\varphi_1'(w)\right)dw+o(\Delta u). \tag{A.12}$$

Let $b=\max\{w:\varphi'(w)=0\}$, that is, b is the largest w for which $\varphi'(w)=0$. If there is no such w satisfying the condition $\varphi'(w)=0$, then $b=0$. Since $\varphi_j'(w)$ is non-decreasing, it must be true that $\varphi_j'(w)=0$ for $w\le b$. Hence, $\varphi_j'(0+)=\varphi_j'(b)=0$ so that $\varphi_j(0)=\varphi_j(b)$. Thus (A.12) can be rewritten as

$$P\left(J\left(y_1\right)=2, J\left(y_2\right)=1, \max_k \tilde{U}\left(k, y_2\right)\in\left(u, u+\Delta u\right)\right)$$
$$=u^{-3}\Delta u\int_b^\infty \exp\left(-u^{-1}\varphi(w)\right)\varphi_2'(w)\left(\varphi_1(w)-w\varphi_1'(w)\right)dw+o(\Delta u). \tag{A.13}$$

On the other hand, from Axiom 2 it follows that

$$P\left(J\left(y_1\right)=2, J\left(y_2\right)=1, \max_k \tilde{U}\left(k, y_2\right)\in\left(u, u+\Delta u\right)\right)$$
$$=P\left(J\left(y_1\right)=2, J\left(y_2\right)=1\right)P\left(\max_k \tilde{U}\left(k, y_2\right)\in\left(u, u+\Delta u\right)\right) \tag{A.14}$$

In the following, it turns out to be convenient to define

$$C_{ij}=P\left(J\left(y_1\right)=i, J\left(y_2\right)=j\right)/\varphi(b). \tag{A.15}$$

Note furthermore that

$$P\left(\max_k \tilde{U}\left(k, y_2\right)\in\left(u, u+\Delta u\right)\right)$$
$$=\exp\left(-u^{-1}\left(\varphi_1(0)+\varphi_2(0)\right)\right)\left(\varphi_1(0)+\varphi_2(0)\right)u^{-2}\Delta u+o(\Delta u) \tag{A.16}$$
$$=\exp\left(-u^{-1}\varphi(b)\right)\varphi(b)u^{-2}\Delta u+o(\Delta u).$$

From (A.13), (A.14) and (A.16), when Δu goes to zero, we obtain that,

$$\lambda\int_b^\infty \exp\left(-\lambda\varphi(w)\right)\varphi_2'(w)\left(\varphi_1(w)-w\varphi_1'(w)\right)dw=C_{21}\exp\left(-\lambda\varphi(b)\right) \tag{A.17}$$

where $\lambda=u^{-1}$. Note also that

$$\exp\left(-\lambda\varphi(b)\right)=\lambda\int_b^\infty \exp\left(-\lambda\varphi(w)\right)\varphi'(w)dw. \tag{A.18}$$

Consequently, (A.17) is equivalent to

$$\int_b^\infty \exp\left(-\lambda\varphi(w)\right)\varphi_2'(w)\left(\varphi_1(w)-w\varphi_1'(w)\right)dw = C_{21}\int_b^\infty \exp\left(-\lambda\varphi(w)\right)\varphi'(w)dw \qquad \text{(A.19)}$$

for $\lambda \geq 0$.

Recall furthermore that $\varphi(w)$ is strictly increasing and differentiable for $w > b$. The uniqueness property of the Laplace transform therefore implies that the integrands in both sides of (A.19) must be equal, i.e.,

$$\varphi_2'(w)\left(\varphi_1(w)-w\varphi_1'(w)\right) = \varphi'(w)C_{21}. \qquad \text{(A.20)}$$

Similarly, it follows by symmetry that

$$\varphi_1'(w)\left(\varphi_2(w)-w\varphi_2'(w)\right) = \varphi'(w)C_{12}. \qquad \text{(A.21)}$$

By subtracting (A.20) from (A.21) we get

$$\begin{aligned}&\varphi_2'(w)\varphi_1(w)-w\varphi_1'(w)\varphi_2'(w)-\varphi_1'(w)\varphi_2(w)+w\varphi_1'(w)\varphi_2'(w)\\&=\varphi_2'(w)\varphi_1(w)-\varphi_2(w)\varphi_1'(w)=\varphi'(w)\varphi_1(w)-\varphi_1'(w)\varphi(w)=\varphi'(w)\left(C_{21}-C_{12}\right)\end{aligned}$$

which, when dividing by $\varphi(w)^2$ becomes

$$\frac{\varphi_1'(w)\varphi(w)-\varphi_1(w)\varphi'(w)}{\varphi(w)^2} = \frac{\varphi'(w)\left(C_{12}-C_{21}\right)}{\varphi(w)^2}. \qquad \text{(A.22)}$$

When we integrate both sides of (A.22) we get

$$\frac{\varphi_1(w)}{\varphi(w)} = \frac{C_{21}-C_{12}}{\varphi(w)} + d_1$$

for $w > b$, where d_1 is a constant. Hence

$$\varphi_1(w) = C_{12} - C_{21} + d_1\varphi(w). \qquad \text{(A.23)}$$

Similarly, it follows that

$$\varphi_2(w) = C_{21} - C_{12} + d_2\varphi(w). \qquad \text{(A.24)}$$

By inserting (A.23) into (A.21) we get

$$\varphi'(w)\left(\varphi_2(w)-w\varphi_2'(w)\right)d_1 = \varphi'(w)C_{12} \qquad \text{(A.25)}$$

for $w > b$. Since $\varphi'(w) > 0$ for $w > b$, (A.25) implies that

$$\varphi_2(w) - w\varphi_2'(w) = \frac{C_{12}}{d_1}. \tag{A.26}$$

Similarly, we get that

$$\varphi_1(w) - w\varphi_1'(w) = \frac{C_{21}}{d_2}. \tag{A.27}$$

Equations (A.26) and (A.27) are first order differential equations that have solutions of the form

$$\varphi_j(w) = \alpha_j + \beta_j w \tag{A.28}$$

for $w > b,\ j = 1,2,$ where α_j and β_j are suitable constants. Since $\varphi_j'(w) = 0$ for $w \le b$ and $\varphi_j'(w)$ is continuous it follows from (A.28) that

$$\varphi_j(w) = \alpha_j + \beta_j b \tag{A.29}$$

for $w \le b$. As a consequence (A.5)and (A.29) yields

$$-\log F_j(u_1, u_2) = \begin{cases} \beta_j u_1^{-1} + \alpha_j u_2^{-1} & for \quad bu_1 \le u_2 \\ (\beta_j b + \alpha_j) u_2^{-1} & for \quad bu_1 > u_2. \end{cases} \tag{A.30}$$

We realize that (A.30) implies that

$$P\left(\tilde{U}(j, y_2) \le u_2 \middle| \tilde{U}(j, y_1) = u_1\right) = \begin{cases} \exp\left(-\alpha_j u_2^{-1}\right), & u_2 \ge bu_1 \\ 0 & u_2 < bu_1. \end{cases} \tag{A.31}$$

This means that $\tilde{U}(j, y_2) \ge b\tilde{U}(j, y_1)$ with probability one. So for $\{\tilde{U}(j,y), y \ge \gamma\}$ to be nondecreasing one must have that $b \ge 1$. If $b > 1$ then $\tilde{U}(j, y_2) > \tilde{U}(j, y_1)$ with probability one and then Axiom 3 cannot hold. We therefore conclude that $b = 1$.

With $b = 1$, we have

$$P\left(\tilde{U}(j, y_2) \le u_2\right) = F_j(\infty, u_2) = \exp(-(\beta_j + \alpha_j)u_2^{-1}) = \exp\left(-v(j, y_2)u_2^{-1}\right),$$

and

$$P\left(\tilde{U}(j, y_1) \le u_1\right) = F_j(u_1, \infty) = \exp(-\beta_j u_1^{-1}) = \exp\left(-v(j, y_1)u_1^{-1}\right).$$

This shows that $\alpha_j = v(j, y_2) - v(j, y_1)$, and $\beta_j = v(j, y_1)$. So the joint distribution can be written as follows:

$$F_j(u_1,u_2)=\begin{cases}\exp\left(-v(j,y_1)u_1^{-1}-\left(v(j,y_2)-v(j,y_1)\right)u_2^{-1}\right) & \text{for} \quad u_1\le u_2,\\ \exp\left(-v(j,y_2)u_2^{-1}\right) & \text{for} \quad u_1>u_2.\end{cases}$$

Q.E.D.

Proof of Theorem 1:

Assume that there are only 2 alternatives $j=1,2$. The arguments in the general case with more than 2 alternatives will be similar.

Let $\{V_j(y_1,y_2)\}$ be random variables with c.d.f.

$$\exp\left(-\frac{\left(v(j,y_2)-v(j,y_1)\right)}{u}\right)$$

for $u>0$, and with the property that $V_j(y_1,y_2)$ is independent of $V_j(y_1',y_2')$ for $y_1<y_2$, $y_1'<y_2'$ provided $[y_1,y_2]\cap[y_1',y_2']=\varnothing$. Define

$$U^*(j,y_2)=\max\left(U^*(j,y_1),V_j(y_1,y_2)\right).$$

Then it follows that $\left(U^*(j,y_1),U^*(j,y_2)\right)$ has joint c.d.f. as in Lemma 1 for all $\gamma\le y_1<y_2$. By Kolmogorov's existence theorem there exists two random variables with c.d.f. as in Lemma 1 and they are unique with probability one. A similar result holds for more than 2 random variables indexed by $\gamma\le y_1<y_2<\cdots<y_d$. With the equivalence of finite dimensional distributions, the processes $\{\tilde{U}(j,y),\, y\ge\gamma\}$ and $\{U^*(j,y),\, y\ge\gamma\}$ are equivalent.

Q.E.D.

Proof of Theorem 2:

From Theorem 1 it follows for $y_2>y_1>\gamma$, that

$$P\left(U(y_2)>U(y_1)\right)=P\left(V(y_1,y_2)>U(y_1)\right). \tag{A.32}$$

Since $V(y_1,y_2)$ and $U(y_1)$ are independent and Fréchet distributed, we get from standard results in discrete choice theory that

$$\begin{aligned}P(U(y_2)>U(y_1))&=1-P\left(V(y_1,y_2)<U(y_1)\right)\\&=1-\frac{v(y_1)}{v(y_1)+v(y_2)-v(y_1)}=1-\frac{v(y_1)}{v(y_2)}\end{aligned} \tag{A.33}$$

for $y_2 \geq y_1 \geq \gamma$. When we combine (A.32) and (A.33) we realize that Axiom 4 implies that whenever

$$\frac{v(y_1)}{v(y_2)} > \frac{v(y_1^*)}{v(y_2^*)}$$

then

$$\frac{v(\lambda(y_1-\gamma)+\gamma)}{v(\lambda(y_2-\gamma)+\gamma)} > \frac{v(\lambda(y_1^*-\gamma)+\gamma)}{v(\lambda(y_2^*-\gamma)+\gamma)}$$

for $y_2 \geq y_1 \geq \gamma$, $y_2^* \geq y_1^* \geq \gamma$ and $\lambda > 0$. Now we can apply Theorem 14.19 in Falmagne (1985), p. 338, which yields[9]

$$\frac{v(y_1)}{v(y_2)} = F\left(\frac{\delta_1\left((y_1-\gamma)^\tau - 1\right) - \delta_2\left((y_2-\gamma)^\tau - 1\right)}{\tau}\right) = 1 - P\left(U(y_2) > U(y_1)\right) = P\left(V(y_1, y_2) < U(y_1)\right) \quad \text{(A.34)}$$

for $y_2 \geq y_1 \geq \gamma$, where $\tau, \delta_1 > 0, \delta_2 > 0$ are constants and F(·) is a continuous and strictly increasing mapping. Evidently, F(·) is defined on $(-\infty, 0]$. When $y_2 = y_1$ we obtain that

$$F\left(\frac{(\delta_1-\delta_2)\left((y_1-\gamma)^\tau - 1\right)}{\tau}\right) = \frac{v(y_1)}{v(y_1)} = 1 = F(0)$$

must hold, for all $y_1 \geq \gamma$. This implies that $\delta_1 = \delta_2 = \delta$ (say).

Let

$$x = \frac{\delta\left((y_1-\gamma)^\tau - 1\right) - \delta\left((y_2-\gamma)^\tau - 1\right)}{\tau}, \quad z = \frac{\delta\left((a-\gamma)^\tau - 1\right) - \delta\left((y_1-\gamma)^\tau - 1\right)}{\tau}$$

where $a \in (\gamma, y_1)$ is fixed. From (A.34) we get

$$\frac{v(a)}{v(y_2)} = F(x+z) \quad \text{(A.35)}$$

---

[9] Note that Theorem 14.19, p.338, in Falmagne (1985) can be expressed more compactly as

$$\mathrm{M}(\mathrm{a}, \mathrm{b}) = \mathrm{F}\left(\frac{\delta_1\left(\mathrm{a}^\theta - 1\right) + \delta_2\left(\mathrm{b}^\theta - 1\right)}{\theta}\right)$$

where $(\mathrm{x}^0 - 1)/0$ is defined as $\lim_{\theta \to 0}\left(\mathrm{x}^\theta - 1\right)/\theta = \log \mathrm{x}$.

and

$$\frac{v(a)}{v(y_1)} = F(z). \tag{A.36}$$

When (A.35) and (A.36) are combined with (A.34) we get

$$F(x+z) = F(x)F(z). \tag{A.37}$$

Eq. (A.37) is a Cauchy functional equation which only continuous solution is the exponential function. Consequently, for $y \geq \gamma$,

$$\log v(y) = \frac{\beta\left((y-\gamma)^{\tau} - 1\right)}{\tau}. \tag{A.38}$$

Q.E.D.

Proof of Theorem 3:

From (A.32) and (A.33) we get that for $y_2 > y_1 > \gamma$

$$P\left(U(y_2) > U(y_1)\right) = 1 - \frac{v(y_1)}{v(y_2)}. \tag{A.39}$$

Hence, Axiom 5 implies that

$$\frac{v\left(\lambda(y_2-\gamma)+\gamma\right)}{v\left(\lambda(y_1-\gamma)+\gamma\right)} = \frac{v(y_2)}{v(y_1)} \tag{A.40}$$

for all $\lambda > 0$. For simplicity, let

$$g(x) = \frac{v(x+\gamma)}{v(1+\gamma)}$$

for $x \geq 1$. With $y_1 = \gamma + 1,\ y_2 = x + \gamma$, we get from (A.40) that

$$g(\lambda x) = g(x)g(\lambda) \quad x \geq 1. \tag{A.41}$$

Eq. (A.41) is a functional equation of the Cauchy type which only continuous solution is the power function

$$g(x) = x^{\delta} \tag{A.42}$$

for some constant δ. Since

$$v(y) = g(y-\gamma)v(1+\gamma)$$

the result of Theorem 3 follows.

Q.E.D.

Proof of Theorem 4:

Let $\xi_i = W_i \varepsilon(y_i)$. Then we can write

$$E\left(\frac{W_i v(y_i)}{\sum_{r\in A} W_r v(y_r)}\right) = P\left(v(y_i)\xi_i = \max_{r\in A}\left(v(y_r)\xi_r\right)\right). \tag{A.43}$$

Since the error terms $\{\xi_i\}$ are independent, we know from Yellott (1977) that Axiom 6 (IIA) can only be satisfied if the errors $\{\xi_i\}$ are Fréchet distributed, i.e.,

$$P(\xi_i \le x) = \exp\left(-x^{-a}\right) \tag{A.44}$$

for $x > 0$, where a is a positive constant. But (A.44) implies that

$$P(\xi_i \le x) = P\left(W_i \varepsilon_i(y_i) \le x\right) = E\,P\left(\varepsilon_i(y_i) \le \frac{x}{W_i}\middle| W_i\right) = E\exp\left(-\frac{W_i}{x}\right). \tag{A.45}$$

The last equality in (A.45) follows because by assumption $P(\varepsilon(y) \le x) = \exp(-1/x)$. When we combine (A.44) and (A.45) we obtain that under the normalization in (12)for any $\lambda > 0$

$$E\exp(-\lambda W_i) = \exp\left(-\lambda^{\alpha}\right). \tag{A.46}$$

The left hand side of (A.46) is the Laplace transform of the distribution of W_i. From Samorodnitsky and Taqqu (1994) Proposition 1.2.12, p. 15, it follows that (A.46) holds W_i must be a strictly α-stable random variable that is totally skew to the right and with $\alpha < 1$.

Q.E.D.

Proof of Theorem 5:

For the sake of simplicity consider first the case with $m = 2$ and $W = 1$. In this case it follows that

$$
\begin{aligned}
&G_2(y_1,y_2)=P\Big(U\big(Y(u_1)\big)\ge U(y_1),U\big(Y(u_2)\big)\ge U(y_2)\Big)\\
&=P\big(u_1\ge U(y_1),u_2\ge U(y_2)\big)=P\Big(U(y_1)\le u_1,\max\big(U(y_1),V(y_1,y_2)\big)\le u_2\Big)\\
&=P\big(U(y_1)\le \min(u_1,u_2),V(y_1,y_2)\le u_2\big)=P\big(U(y_1)\le \min(u_1,u_2)\big)P\big(V(y_1,y_2)\le u_2\big).
\end{aligned} \tag{A.47}
$$

From Corollary 1 and (A.47) we obtain that

$$
G_2(y_1,y_2)=\exp\Big(-v(y_1)\max\big(u_1^{-1},u_2^{-1}\big)-\big(v(y_2)-v(y_1)\big)u_2^{-1}\Big). \tag{A.48}
$$

In particular, when $u_2\ge u_1$, (A.48) reduces to

$$
G_2(y_1,y_2)=\exp\Big(-v(y_1)\big(u_1^{-1}-u_2^{-1}\big)-v(y_2)u_2^{-1}\Big). \tag{A.49}
$$

The multivariate case is completely analogous. It follows readily that when $u_1\le u_2\le\ldots\le u_m$, $y_1\le y_2\le\ldots\le y_m$,

$$
G_m(y_1,y_2,\ldots,y_m)=\exp\left(-v(y_m)u_m^{-1}-\sum_{j=1}^{m-1}v(y_j)\big(u_j^{-1}-u_{j+1}^{-1}\big)\right). \tag{A.50}
$$

For $u_1\le u_2\le\ldots\le u_m$ and $y_1<y_2<\ldots<y_m$, the corresponding joint density function equals

$$
g_m(y_1,y_2,\ldots,y_m)=G_m(y_1,y_2,\ldots,y_m)v'(y_m)u_m^{-1}\prod_{j=1}^{m-1}\big(u_j^{-1}-u_{j+1}^{-1}\big)v'(y_j). \tag{A.51}
$$

The general case with random W now follows readily from (A.50) and (A.51).

Hence, the proof is complete.

Q.E.D.

Proof of Corollary 3:

From Theorem 5 and (14) we get that

$$
\begin{aligned}
P(Y(u)>y)&=E\exp(Wv(y)u^{-1})=\exp(-u^{\alpha}v(y)^{\alpha})\\
&=P(\eta(u)>u^{\alpha}v(y)^{\alpha})=P(u\eta(u)^{1/\alpha}>v(y))\\
&=P(v^{-1}(u\eta(u)^{1/\alpha})>y)
\end{aligned}
$$

where $\eta(u)$ is exponentially distributed with parameter one. Hence $v^{-1}(u\eta(u)^{1/\alpha})$ has the same c.d.f. as Y(u) and this implies that v(Y(u)) has the same c.d.f. as $u\eta(u)^{1/\alpha}$.

Q.E.D.

Proof of Corollary 4:

We have from (7) that

$$(\frac{U(y_2)}{U(y_1)})^b = \max(1,(\frac{V(y_1,y_2)}{U(y_1)})^b). \tag{A.52}$$

Note that $V(y_1,y_2)$ and $U(y_1)$ are independent Fréchet distributed $\exp(-kx^{-1})$, x>0 with parameter k equal to $v(y_2)-v(y_1)$ and $v(y_1)$, respectively. Hence, from standard results in discrete choice theory we get that (for $q>0$):

$$P\left(V(y_1,y_2)\le qU(y_1)\right)=\frac{q}{q+K(y_1,y_2)}=\frac{1}{1+q^{-1}K(y_1,y_2)} \tag{A.53}$$

where

$$K(y_1,y_2)=\frac{v(y_2)-v(y_1)}{v(y_1)} \tag{A.54}$$

Hence, (A.53) implies that we can write

$$\left(\frac{V(y_1,y_2)}{U(y_1)}\right)^b = K(y_1,y_2)^b Z(y_1,y_2)^b,$$

where $Z(y_1,y_2)$ is a positive random variable with c.d.f. $P(Z(y_1,y_2)\le z)=\frac{1}{1+z^{-1}}$ for z 0.

Hence ,

$$E\left[\left(\frac{U(y_2)}{U(y_1)}\right)^b\right]=E\max(K(y_1,y_2)Z(y_1,y_2)^b,1)=1+\int_1^\infty P\left(K(y_1,y_2)^b Z(y_1,y_2)^b>q\right)dq. \tag{A.55}$$

By using the change of variable $q\to K(y_1,y_2)^b x^b$, and using (A.54) we get

$$E\left[\left(\frac{U(y_2)}{U(y_1)}\right)^b\right]=1+(\frac{v(y_2)}{v(y_1)}-1)^b\int_{K(y_1,y_2)^{-1}}^\infty b\,x^{b-1}P(Z(y_1,y_2)>x)\,dx. \tag{A.56}$$

Clearly, the right hand side of (A.56) is strictly increasing in $v(y_2)/v(y_1)$ and it exists when b<1 and is equal to infinity for b 1.

Q.E.D.

Proof of Corollary 5:

From Corollary 4 and (A.53), we get that when $q\ge 1$

$$P\left(\left(\frac{U(y_j)}{U(y_i)}\right)^b \le q\right) = P\left(\left(\frac{V(y_i,y_j)}{U(y_i)}\right)^b \le q\right) = \frac{1}{1+K(y_i,y_j)q^{-1/b}} . \tag{A.57}$$

Since $K(y_i,y_j)$ is increasing as a function of $v(y_j)/v(y_i)$, it follows that $P\left(\left(\frac{U(y_j)}{U(y_i)}\right)^b \le q\right)$ is decreasing as a function of $v(y_j)/v(y_i)$. Hence, the result of the corollary follows.

Q.E.D.

References

Anderson, S.P., Palma, A.D., Thisse, J. F., 1992. Discrete Choice Theory of Product Differentiation. MIT Press, Cambridge, Massachusetts.

Barnard, G.A., 1965. The Use of the Likelihood Function. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability I**.** University of California Press, Berkeley, pp. 27-40.

Buyze, J.,1982. The Estimation of Welfare Levels of a Cardinal Utility Function. European Economic Review 17, 325-332.

Dagsvik, J.K., 1994. Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes. Econometrica 62,1179-1205.

Dagsvik, J.K., 1995. How Large is the Class of Generalized Extreme Value Models? Journal of Mathematical Psychology 39, 90-98.

Dagsvik, J.K., 2002. Discrete Choice in Continuous Time: Implications of an Intertemporal Version of the IIA Property. Econometrica 70, 817-831.

Dagsvik, J.K., and Røine, S., 2005. Testing Dimensional Invariance in a Stochastic Model of the Utility of Income. Statistics Norway (forthcoming).

Debreu, G., 1958. Stochastic Choice and Cardinal Utility. Econometrica 26, 440-444.

Ellermeier, W., Faulhammer, G., 2000. Empirical Evaluation of Axioms Fundamental to Stevens' Ratio-scaling Approach. I. Loudness Production. Perception & Psychophysics 62, 1505-1511.

Ellingsen, T., 1994. Cardinal Utility: A History of Hedonimetry. In: Allais, M., Hagen, O., (Eds.), Cardinalism. Kluwer Academic Publishers, London.

Falmagne, J.-C., 1985. Elements of Psychophysical Theory. Oxford University Press, New York.

Fechner, G.T., 1860/1966. Elements of Psychophysics. In: Howes, D.H., Boring, E.C., (Eds.), Adler, H.E. , (Translated from German. Originally published in 1860.) Holt, Rinehart and Winston, New York.

Fechner, G. T., 1882/1965. Some thoughts on the Psychophysical Representation of Memories. Psychological Research 49, 209-212.

Fishburn, P.C., 1998. Stochasitc Utility. In: Barbera, S., Hammond, P.J., Seidl, C.,(Eds.), Handbook of Utility Theory, Vol. I, Kluwer Academic Publishers, Boston.

Fisher, I., 1892. Mathematical Investigations in the Theory of Value and Prices. Transactions of the Connecticut Academy of Arts and Sciences 9, New Haven.

Fisher, I., 1918. Is "Utility" the Most Suitable Term for the Concept It Is used to Denote? American Economic Review 8, 335-337.

Fisher, I., 1927. A Statistical Method for Measuring Marginal Utility and Testing the Justice of a Progressive Income Tax. In: Hollander, J. H., (Eds.), Economic Essays in Honor of John Bates Clark, The Macmillan Company, New York.

Frisch, R., 1926. Sur un Problème d'Économie Pure, Norsk Matematisk Forenings Skrifter 16, 1-40, English translation in Chipman, J., Hurwicz., L., Richter, M. K., Sonnenschein, H. F., (Eds.), 1971, Preferences, Utility and Demand. Harcourt Brace Jovanovich Inc. London.

Frisch, R., 1932. New Methods of Measuring Marginal Utility. Mohr, Tübingen.

Frisch, R., 1959. A Complete Scheme for Computing All Direct and Cross demand Elasticities in a Model with Many Sectors. Econometrica 27, 177-196.

Gescheider, G. A., 1997. Psychophysics: The Fundamentals. Lawrence Erlbaum Associates, Publishers, London.

Gescheider, G.A., Bolanowski, S.J., Jr., 1991a. Introduction to Conference on Ratio Scaling of Psychological Magnitudes. In: Bolanowski, Jr. S.J., Gescheider, G.A., (Eds.), Ratio Scaling of Psychological Magnitude: In Honor of the Memory of S. S. Stevens. Lawrence Erlbaum Associates, Inc., Hillsdale, N.J.

Gescheider, G.A., Bolanowski, S.J. Jr., 1991b. Final Comments on Ratio Scaling of Psychological Magnitudes. In: Bolanowski, Jr. S.J., Gescheider, G.A., (Eds.), Ratio Scaling of Psychological Magnitude: In Honor of the Memory of S. S. Stevens. Lawrence Erlbaum Associates, Inc., Hillsdale, N.J.

Gorman, W.M., 1953. Community Preference Fields. Econometrica 21, 63-80.

Heidelberger, M., 2004. Nature from Within: Gustav Theodor Fechner and his Psychophysical Worldview. University of Pittsburgh Press, Pittsburgh, PA.

Johansen, L., 1960. A Multi-Sectoral Study of Economic Growth. North-Holland, Amsterdam.

Kapteyn, A., Wansbeck, T., 1985. The Individual Welfare Function. A Review. Journal of Economic Psychology 6**,** 333-363.

Luce, R.D., 1959a. Individual Choice Behavior. Wiley, New York.

Luce, R.D., 1959b. On the Possible Psychophysical Laws. The Psychological Review 66, 81-95.

Luce, R. D., Suppes, P., 1965. Preference, Utility and Subjective Probability. In: Luce, R.D., Bush, R.R, Galanter, E., (Eds.), Handbook of Mathematical Psychology, Vol. 3. Wiley, New York.

Mach, E., 1886/1996. The Analysis of Sensations. The Origins of Modern Philosophy of Science 1830-1914. Routledge/Thoemmes, London.

McFadden, D., 1973. Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka, P., (Eds.), Frontiers of Econometrics. MIT Press, Cambridge, pp. 105-142.

McFadden, D., 1978. Modelling the Choice of Residential Location. In: Karlquist, A., Lundquist, L., Snickars,F., Weibull, J.J. (Eds.), Spatial Interaction Theory and Residential Location. North Holland, Amsterdam, pp. 75-96.

McFadden, D., 1981. Econometric Models of Probabilistic Choice. In: Manski, C. F., McFadden, D.,(Eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, MA, pp. 198-272.

McFadden, D., 2001. Economic Choices, American Economic Review 91, 351-378.

Narens, L., 1996. A Theory of Magnitude Estimation. Journal of Mathematical Psychology 40, 109-129.

Patel, P, Subrahmanyam, M., 1978. Utility Theory and Participation in Unfair Lotteries. Journal of Economic Theory 19, 555-57.

Quandt, R.E., 1956. A Probabilistic Theory of Consumer Behavior. Quarterly Journal of Economics 70, 507-536.

Resnick, S.I., 1987. Extreme Values, Regular Variation and Point Processes. Springer-Verlag, New York.

Samorodnitsky, G., Taqqu, M.S., 1994. Stable Non-Gaussian Random Processes. Chapman & Hall, New York.

Seidl, C., 1994. How sensible is the Leyden Individual Welfare Function of Income? European Economic Review 38, 1633-1659.

Sen, A., 1982. Choice, Welfare and Measurement. MIT Press, Cambridge, MA.

Shepard, R.N., 1981. Psychological Relations and Psychophysical Scales: On the Status of "Direct" Psychophysical Measurement. Journal of Mathematical Psychology 24**,** 21-57.

Simon, H.A., 1986. Rationality in Psychology and Economics. Journal of Business 594, 209-224.

Smith, R.L., 1994. Nonregular Regression. Biometrika 81, 173-183.

Smith, V., 1990. Experimental Economics: Behavioral Lessons for Microeconomic Theory and Policy, Nancy L. Schwartz Memorial Lecture at J. L. Kellogg Graduate School of Management, Northwestern University.

Stevens, S.S., 1946. On the Theory of Scales of Measurement. Science 161, 677-680.

Stevens, S.S., 1951. Mathematics, Measurement, and Psychophysics. In S.S. Stevens (Eds.), Handbook of Experimental Psychology, Wiley, New York.

Stevens, S.S., 1957. On the Psychophysical Law. Psychological Review 64, 153-181.

Stevens, S.S., 1975. Psychophysics: Introduction to Its Perceptual Neural, and Social Prospects. Wiley, New York.

Strauss, D., 1979. Some Results on Random Utility Models. Journal of Mathematical Psychology 17, 32-35.

Suppes, P., Luce, R. D., Krantz, D., Tversky, A., 1989. Foundations of Measurement: Geometrical, Threshold, and Probabilistic Representations. Academic Press, London.

Thurstone, L.L., 1927. A Law of Comparative Judgment. Psychological Review 34, 272-286.

Torgerson, W.S., 1961. Distances and Ratios in Psychological Scaling. Acta Psychologica 19, 201-205.

Van Herwaarden, F.G., Kapteyn, A., van Praag, B.M.S., 1977. Twelve Thousand Individual Welfare Functions. European Economic Review 9, 283-300.

Van Herwaarden, F.G. Kapteyn, A., 1981. Empirical Comparisons of the Shape of Welfare Functions. European Economic Review 15, 261-286.

Van Praag, B.M.S., 1968. Individual Welfare Functions and Consumer Behavior: A Theory of Rational Irrationality. North-Holland, Amsterdam.

Van Praag, B.M.S., 1971. The Welfare Function of Income in Belgium: An Empirical Investigation. European Economic Review 2, 337-369.

Van Praag, B.M.S., 1991. Ordinal and Cardinal Utility: An Integration of the Two Dimension of the Welfare Concept, Journal of Econometrics 50, 60-89.

Van Praag, B.M.S., 1994. The Relativity of the Welfare Concept. In Nussbaumm, M., and Sen, A., (Eds.), The quality of Life. Clarendon Press, Oxford.

Van Praag, B.M.S., Kapteyn, A., 1994. How Sensible is the Leyden Individual Welfare Function of Income? A Reply. European Economic Review 83, 1817-1825.

Yellott, J.I., Jr., 1977. The Relationship between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment and the Double Exponential Distribution. Journal of Mathematical Psychology 15, 109-146.

Zanakis, S.H., Kyparisis, J., 1986. A Review of Maximum Likelihood Estimation Methods for the Three Parameter Weibull Distribution. Journal of Statistical Computation and Simulation 25, 53-73.

Tables:

Table 1. Income Evaluation Questions. All numbers are in NOK as of September 1995

Utility level u_j	Disposable Income	Mean	Std.	Minimum	Maximum
Bad	$Y(u_1)$	160428	76268	30000	500000
Insufficient	$Y(u_2)$	203776	91260	50000	600000
Sufficient	$Y(u_3)$	247452	110469	70000	700000
More than Sufficient	$Y(u_4)$	296180	131622	100000	800000
Good	$Y(u_5)$	349212	155918	120000	1000000
Very good	$Y(u_6)$	435932	214094	140000	1500000
# observations		250			

Table 2. Summary statistics of basic characteristics of the sample

	Mean	Std.	Minimum	Maximum
Family size	2.40	1.31	1	6
# children less than 6	0.31	0.62	0	3
# children between 7 and 15	0.28	0.59	0	2
# children above 16	0.16	0.48	0	3
Education Years	16.39	3.09	8	24
Income (100 000 NOK)	3.67	2.03	0.3	19
Number of Individual with positive debt	217			
Debt (those with positive debt, 100 000 NOK)	4.88	3.66	0.05	21
# of observations	250			

Table 3. Maximum likelihood estimates of the parameters of the utility of income

		General model		Model under EQA	
Variables/parameters		Estimates	Std.	Estimates	Std.
Box-Cox exponent τ		0.048	0.020	0.037	0.019
Dispersion parameter σ		0.166	0.004	0.166	0.004
Characteristic exponent α		0.328	0.015	0.332	0.015
Transformed utility levels $a_j = d_j + t \log I$	Constant level 1, d_1	0.364	0.307		
	Constant level 2, d_2	1.729	0.310		
	Constant level 3, d_3	3.048	0.317		
	Constant level 4, d_4	4.372	0.326		
	Constant level 5, d_5	5.723	0.339		
	Constant level 6, d_6	7.545	0.362	7.365	0.300
	Log income, t	1.517	0.194	1.486	0.190
	Utility ratio ρ			1.683	0.430
Loglikelihood		-1141.7		-1143.4	

Table 4. Summary statistics, Norwegian Households with Positive Disposable Household Income, 1995

	Mean	Std	Minimum	Maximum
Family size	2.23	1.30	1	12
# children less than 6	0.22	0.55	0	4
# children between 7 and 15	0.25	0.60	0	5
# children above 16	0.05	0.23	0	3
Income (100 000 NOK)	2.33	1.95	0.37	114
Debt (100 000 NOK)	3.16	6.12	0	306
# of observations	1 902 367			

Figures:

Figure 1. Plot of K(y) against ln(y)

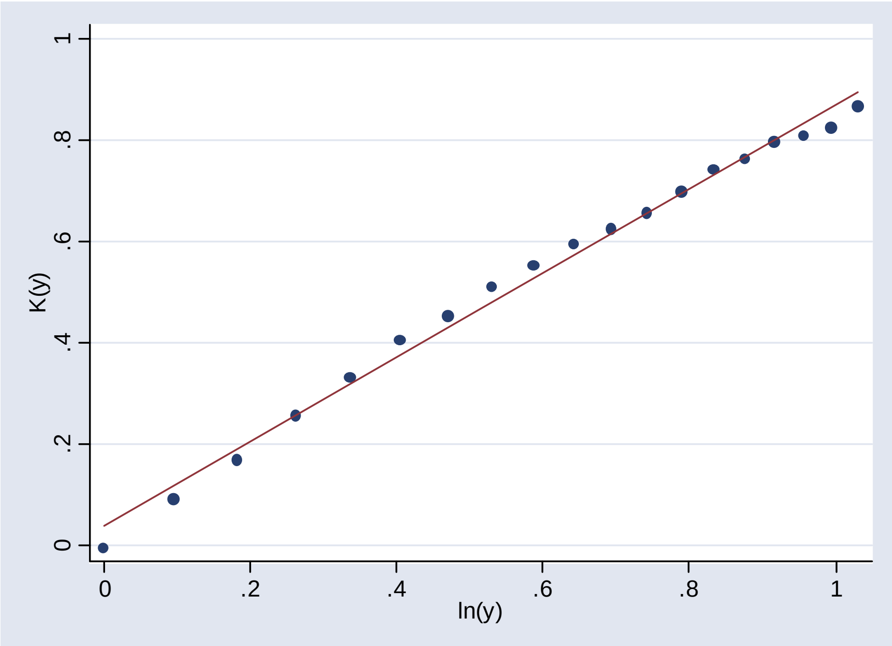

Figure 2. Distribution of the Money Flexibility