

Dagsvik, John K.; Strøm, Steinar; Jia, Zhiyang

**Working Paper**

## A stochastic model for the utility of income

Memorandum, No. 2003,32

**Provided in Cooperation with:**

Department of Economics, University of Oslo

*Suggested Citation:* Dagsvik, John K.; Strøm, Steinar; Jia, Zhiyang (2003) : A stochastic model for the utility of income, Memorandum, No. 2003,32, University of Oslo, Department of Economics, Oslo

This Version is available at:

<https://hdl.handle.net/10419/63208>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# MEMORANDUM

No 32/2003

**A Stochastic Model for the Utility of Income**

The seal of the University of Oslo is a circular emblem. It features a central figure of a woman in classical attire, holding a lyre. The text "UNIVERSITAS OSLOENSIS" is inscribed around the top inner edge of the circle, and "MDCCLXI" is at the bottom. Two small dots separate the top and bottom text on the left and right sides.

**John K. Dagsvik, Steinar Strøm and  
Zhiyang Jia**

ISSN: 0801-1117

---

**Department of Economics  
University of Oslo**

This series is published by the  
**University of Oslo**  
**Department of Economics**

P. O.Box 1095 Blindern  
N-0317 OSLO Norway  
Telephone: + 47 22855127  
Fax: + 47 22855035  
Internet: <http://www.oekonomi.uio.no/>  
e-mail: [econdep@econ.uio.no](mailto:econdep@econ.uio.no)

In co-operation with  
**The Frisch Centre for Economic  
Research**

Gaustadalleén 21  
N-0371 OSLO Norway  
Telephone: +47 22 95 88 20  
Fax: +47 22 95 88 25  
Internet: <http://www.frisch.uio.no/>  
e-mail: [frisch@frisch.uio.no](mailto:frisch@frisch.uio.no)

List of the last 10 Memoranda:

|       |  |
|-------|--|
| No 31 | Karine Nyborg, Richard B. Howarth, and Kjell Arne Brekke<br>Green consumers and public policy: On socially contingent moral motivation. 23 pp.   |
| No 30 | Halvor mehlum<br>A Finer Point in Forensic Identification. 13 pp.  |
| No 29 | Svenn-Erik Mamelund<br>Effects of the Spanish Influenza Pandemic of 1918-19 on Later Life Mortality of Norwegian Cohorts Born About 1900. 30 pp. |
| No 28 | Fedor Iskhakov<br>Quasi-dynamic forward-looking model for joint household retirement decision under AFP scheme. 59 pp.                           |
| No 27 | Henrik Wiig<br>The Productivity of Social Capital. 22 pp.  |
| No 26 | Tao Zhang<br>Identifying treatment effects of active labour market programmes for Norwegian adults. 45 pp.                                       |
| No 25 | Tao Zhang<br>A Monte Carlo study on non-parametric estimation of duration models with unobserved heterogeneity. 89 pp.                           |
| No 24 | Karine Nyborg and Kjetil Telle<br>The Role of Warnings in Regulation: Keeping Control with Less Punishment. 31 pp.                               |
| No 23 | Maria Kalvaraskaia<br>Savings behaviour when households have access to occupational pensions. 50 pp.   |
| No 22 | Rolf Golombek and Michael Hoel<br>Climate Policy under Technology Spillovers. 37 pp.   |

A complete list of this memo-series is available in a PDF® format at:  
<http://www.oekonomi.uio.no/memo/>

October 3, 2003.

# **A Stochastic Model for the Utility of Income**

by

John K. Dagsvik, Statistics Norway,

Steinar Strøm, University of Oslo

and

Zhiyang Jia, University of Oslo

## **Abstract**

In this paper we propose a particular approach to measuring utility of income. To this end 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 ideas and principles that are used in modern psychophysical research and theories of probabilistic choice.

The empirical part of the paper is based on “Stated Preference” data (SPD). In the present context this means that individuals participating in a laboratory type of experiments are asked to rank order a set of hypothetical alternatives presented.

*Keywords:* Utility of income, Random utility, Invariance principles

Corresponding author: Steinar Strøm, Department of Economics, University of Oslo, P.O.Box 1095, Blindern, 0317 Oslo, Norway. Email: [steinar.strom@econ.uio.no](mailto:steinar.strom@econ.uio.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 total income (see Ellingsen ,1994, for a review of the use of cardinal utility in the economic literature). 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 vary with income are rare<sup>1</sup>.

In this paper we develop a stochastic model for the utility of income, which subsequently is estimated on micro data. Our empirical analysis is based on interview data of the “Stated Preference” (SP) type. We consider this type of 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 experiment runs. 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, which might be estimated separately for each individual.<sup>2</sup>

Irrespective of the type of data that is used, being it market data or SP data, problems related to functional form and the distribution of unobservables remain. The tradition in economics has been to employ ad hoc assumption with regard to functional form and the distribution of unobservables, alternatively to rely on non-parametric approaches. In this paper we propose an alternative strategy, namely an axiomatic approach to justify the choice of functional form of utility functions and the distribution of unobservables. 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. We have adopted and modified ideas and principles from this literature. In fact we demonstrate how the application of invariance principles similarly to the ones employed in psychophysics, combined with several versions of the “Independence from Irrelevant Alternatives” property, lead to explicit characterizations of functional form and the distribution of random terms of the utility function. We consider these invariance

---

<sup>1</sup> Of course, there are a large number of studies where assumptions about the utility function are tested “indirectly” through empirical analysis of the corresponding derived demands.

<sup>2</sup> A good example of the use of SPD is a study of individuals’ preferences over parking attributes (fee and walking distance) conducted by Fisher and Nagin (1981). With SPD more information is obtained on behavioral responses, which may prove helpful in the detection of the structure and distribution of preferences.

principles to be intuitive and plausible as a theoretical rationale for restricting the class of admissible specifications, as we shall discuss below.

The paper is organized as follows. In Section 2 we give a brief review of the literature on the measurement of utility. Section 3 presents and discusses the axioms and their implications for the stochastic model for utility of income. In the Section 4 we discuss how to introduce unobserved heterogeneity in the model, and in Section 5 we discuss the empirical specification of the model. In Section 6 we describe the data we use and empirical results. Section 7 discusses how the marginal utility of income shall be understood within our framework.

## 2. A brief review of the literature

Here we shall briefly discuss some selected works that analyze theoretical and empirical issues related to the measurement of utility. 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 Irving Fisher (1892, 1918, 1927). However, to the best of our knowledge, the first one to estimate the marginal utility of money was Ragnar Frisch (1926)<sup>3</sup>.

A key part in Frisch (1926) was the introduction of certain behavioral-or choice axioms. The choice axioms Frisch referred to were of two types. The first was a preference ordering axiom concerning completeness, transitivity and regularity. The second axiom gave restrictions on the ordering of changes from one position in the commodity space to another and was named intensity of preferences. The two axioms were called inter-local choice axioms and implied that individuals were able to compare and rank changes in the commodity space, given the points in the commodity space. The inter-local choice axioms imply the existence of a utility function measurable on an interval scale, or a cardinal utility function. In Frisch (1947) he expresses 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

---

<sup>3</sup> “ I approached the problem of utility measurement in 1923 during a stay in Paris. There were three objects I had in view :  
(I) To point out the choice axioms that are implied when we think of utility as a quantity, and to define utility in a rigorous way by starting from a set of such axioms;  
(II) To develop a method of measuring utility statistically;  
(III) To apply the method to actual data.

The results of my study along these lines are contained in a paper “Sur un Problème d’Économic Pure”, published in the Series Norsk Matematisk Forenings Skrifter, Serie I, Nr 16, 1926. In this paper, the axiomatics are worked out so far as the static utility concept is concerned. The method of measurement developed is the method of isoquants, which is also outlined in Section 4 below. The statistical data to which the method was applied were sales and price statistics collected by the “Union des Coopérateurs Parisien”. From these data I constructed what I believe can be considered the marginal utility curve of money for the “average” member of the group of people forming the customers of the union. To my knowledge, this is the first marginal utility curve of money ever published”. (Quoted from Frisch, 1932, p. 2-3).

were asked to rank utility differences. Instead he assumed an additive 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 (Frisch 1926, 1959). In his earliest work referred to above, he assumed that there existed at least one good with the property that its marginal utility of consumption was independent of the consumption of other goods. The additivity 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), van Herwaarden, Kapteyn and van Praag (1977), van Herwaarden and Kapteyn (1981), Kapteyn and Wansbeck (1985). A recent 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 use are typically collected through Income Evaluation Questions (IEQ). This means that respondents were asked to answer the questions given in Table 1 below.

**Table 1. Income Evaluation Questions of the Leyden school**

|                   |                    |   |
|-------------------|--------------------|---|
| Excellent         | If it were above   | \$...y <sub>7</sub> .....                               |
| Good              | If it were between | \$...y <sub>6</sub> ..... and \$...y <sub>7</sub> ..... |
| Amply sufficient  | If it were between | \$...y <sub>5</sub> ..... and \$...y <sub>6</sub> ..... |
| Sufficient        | If it were between | \$...y <sub>4</sub> ..... and \$...y <sub>5</sub> ..... |
| Barely sufficient | If it were between | \$...y <sub>3</sub> ..... and \$...y <sub>4</sub> ..... |
| Insufficient      | If it were between | \$...y <sub>2</sub> ..... and \$...y <sub>3</sub> ..... |
| Bad               | If it were between | \$...y <sub>1</sub> ..... and \$...y <sub>2</sub> ..... |
| Very bad          | If it were below   | \$...y <sub>1</sub> .....                               |

The  $\{y_j\}$  represent the respective income boundaries reported by the respondents. For simplicity, we have suppressed the indexation of the individuals. It is assumed that the respondents evaluate income on the basis of the utility that they derived from income. Thus, the answers may be used to recover an underlying utility function. It is not implied in the IEQ scheme that the respondent's 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 respondent minimizes the average inaccuracy of his/her answer, given that he/she has to represent a continuous utility function in income in terms of discrete categories. More precisely, the procedure is the following: Let  $U(y)$  denote the utility of income. The respondent's evaluation of an income in the income interval  $[y_j, y_{j+1}]$  is on average given by  $U^*(y_j) = (U(y_{j+1}) + U(y_j))/2$ . It is

assumed that the average inaccuracy of the respondent's answers can be represented by the following loss function:

$$\sum_{j=0}^{n-1} \int_{y_j}^{y_{j+1}} [U(y) - U^*(y)]^2 dU(y).$$

Thus, the contribution to inaccuracy from each interval is weighted by the utility mass in that interval, with  $y_0 = 0$  and  $y_{n+1} = \infty$ . From the loss function we note that this entails that *utility differences* are being compared and thus a cardinal utility function is implied. Given that  $U(y_0) = 0$  and  $U(y_{n+1}) = 1$ , the utility function that follows from minimizing the loss function is given by  $U(y_{i+1}) - U(y_i) = 1/n$ , for  $i = 0, 1, \dots, n-1$ . Thus, the optimal solution, also called the Equal Quantile Assumption (EQA), says that the respondents maximize informational content by letting the perceived difference in utility between two adjacent labels be equal. EQA was rejected in a test performed by Buyze (1982), but it was concluded that EQA still 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  $U(y)$  is assumed to have the same functional form as a lognormal 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 lognormal utility function. However, the authors preferred the lognormal 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 utility function, its boundedness from above and from below, and the Cobb-Douglas form of the iso-price functions, 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 to the utility income. According to Seidl both laws have strong theoretical support and convincing empirical confirmation, while the ad hoc assumptions in these theories are kept at a minimum.

While economists traditionally express considerable uneasiness when confronted with the issue of how to measure utility, 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 the mathematical laws for the relation between physical stimuli (money) and sensory response (utility) seem to have started with Fechner (1860), and Stevens (1946, 1951). Stevens (1946) describes the work of the special committee of the British Association for Advancement of Science, instructed to investigate the possibility of quantitative estimates of sensory events. One of the questions being addressed was: Is it possible to measure human sensation? Stevens (1957, 1975), among others, 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 Bolanowski and Gescheider (1991a,b). There are several different types of survey questionnaires applied by Stevens and his followers to obtain SPD. 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?” In another version of the method the subject is asked to “produce” a number that matches the intensity of the stimulus. 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 or light 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 particular, to an orthodox economist it may seem to be complete madness to ask questions like this. The reason for this attitude is the convention, established purely by habit, that agents are only able to make ordinal rank orderings. A good illustration of the skepticism 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 are consistent with the power law and that 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. Stevens (1975), p. 18, presents the following views on why the power function appears plausible in the perceptual domain:

“Since, in order to survive, we must be able to move about effectively, perception must to a certain degree achieve stable and veridical representations. It must tell us how matters stand out there. But the universe is in constant flux. We move about and other things also move. Day turns into night. Sound sources approach and recede. How can perceptual stability be achieved in the face of the ongoing flux?

We can perhaps formulate a better question by asking what aspect of the universe most needs stability. For example, is it the differences or the proportions and ratios that need to remain constant in perception? Apparently it is the proportions - the ratios. When we walk towards a house, the relative proportions of the house appear to remain constant: the triangular gable looks triangular from almost any distance. A photograph portrays the same picture whether we view it under a bright or a dim light: the ratio between the light and the shaded parts of the photograph seem approximately the same even though the illumination varies. The perceived relations among the sounds of speech remain the same whether the speech is soft or highly amplified. In other words, the perceptual domain operates as though it had its own ratio requirement—not a mathematically rigid requirement, as in physics, but a practical and approximative requirement.”

The argument above leads to the power function because the power function is the only one that has the property that relative change in output is proportional to relative change in input. Note also that the psychophysical experiments described above imply that utility is *cardinal*, a point that we shall discuss further in Section 7.

With Steven's 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 three 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 (1975). A good reference source to these kinds of theories, can be found in the book by Falmagne (1985).

Among economists, an early contribution in the tradition of Fechnerian psychophysics is due to Debreu (1958). Without relying on a random utility formulation he postulated a stochastic choice model with binary choice probabilities that were assumed to satisfy certain conditions (equivalent to Axiom 1 in Section 3 below). Given these conditions he 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<sup>4</sup>.

---

<sup>4</sup> 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. Relationship between the cardinal utility function and the binary choice probabilities has been established by Falmagne (1985).

### 3. The model

Similarly to Lancaster (1971) we identify consumer goods with quantities, prices and non-pecuniary attributes. We shall use the term “characteristics” to mean prices and non-pecuniary attributes. Thus, a “quantity” and a vector of characteristics represent a commodity. Let  $\Omega$  denote the set of characteristics, which we assume is uncountable. Thus,  $\Omega$  is a continuum, for example a subset of a Euclidian space. Let  $\mathfrak{B}$  denote the family of all *finite* subsets with elements from  $\Omega$ . We shall assume that preferences over characteristics can be represented by a random utility function,  $\tilde{U}(z, y)$ ,  $z \in \Omega$ , where  $y$  denotes income,  $y \geq \gamma$ , and  $z$  is a vector of characteristics.<sup>5</sup> Here,  $\gamma$  is interpreted as a subsistence level that may be specific to the agent. Let  $P(z_i, z_j)$  be the probability that the agent prefers  $z_i$  over  $z_j$ ,  $(z_i, z_j) \in \Omega$ , and let  $R_+ = (0, \infty)$ .

#### 3.1. The individual utility function as a stochastic process

In this section we shall discuss assumptions concerning the stochastic properties of the utility function. The following axiom is due to Debreu (1958).

##### Axiom 1

Let  $z_k, z_r, z_i, z_j \in \Omega$ . The binary choice probabilities satisfy

- (i) *The Balance condition;  $P(z_i, z_j) + P(z_j, z_i) = 1$ .*
- (ii) *The Quadruple condition;  $P(z_i, z_j) \geq P(z_k, z_r)$  if and only if  $P(z_i, z_k) \geq P(z_j, z_r)$ . Moreover, if either antecedent inequality is strict so is the conclusion.*
- (iii) *Solvability; For any  $s \in (0, 1)$  satisfying  $P(z_i, z_j) \geq s \geq P(z_i, z_k)$ , there exists a  $z \in \Omega$  such that  $P(z_i, z) = s$ .*

The Balance condition in Axiom 1 means that either  $z_i$  is preferred to  $z_j$  or  $z_j$  is preferred to  $z_i$ . Thus, indifference is ruled out. The Solvability condition (iii) means that the “amount of variety” in  $\Omega$  is so rich that to any given attribute vector  $z_i \in \Omega$  and any population fraction  $s$ , one can always find a  $z \in \Omega$  such that the fraction who prefers  $z_i$  over  $z$  equals  $s$ .

To clarify the intuition of Axiom 1, note that in an additive random utility model,

$U(z_j) = v(z_j) + \varepsilon_j$ , the binary choice probabilities can be written as

---

<sup>5</sup> It is implicit in this notation that the alternatives are completely represented through their respective vector of characteristics.

$$(3.1) \quad P(z_j, z_k) = P(U_j > U_k) = H(v(z_j) - v(z_k))$$

where  $H$  is the c.d.f. of  $\varepsilon_k - \varepsilon_j$ , that is assumed to be independent of  $j$  and  $k$ . Clearly the choice probabilities given by (3.1) satisfy the Quadruple condition (ii). Furthermore, if the distribution function  $H$  is continuous, the Balance condition (i) is also satisfied for this model. Finally, if for any  $x$  there exists a  $z \in Z$  such that  $v(z) = x$ , then clearly the Solvability condition (iii) holds provided  $H$  is continuous and strictly increasing.

The next result is essentially due to Debreu (1958).

### **Theorem 1**

*Axiom 1 implies that there exists a continuous and strictly increasing c.d.f.  $G$  with  $G(x) + G(-x) = 1$  and a mapping  $\kappa$  from  $\Omega$  onto some interval such that the binary choice probabilities can be represented as*

$$(3.2) \quad P(z_j, z_k) = G(\kappa(z_j) - \kappa(z_k))$$

*for  $(z_j, z_k) \in \Omega^2$ . The mapping  $\kappa$  is unique up to an increasing linear transformation. The function  $G$  is unique in the sense that if  $G_1$  and  $G_0$  are two representations, then  $G_0(x) = G_1(bx)$ , where  $b > 0$  is a constant.*

The representation given in Theorem 1 is known as the *Fechnerian representation*. It is a formalization of an idea originally due to Fechner (1860), and it has been studied by many authors such as Debreu (1958), Falmagne (1985), Suppes et al. (1989), p. 391.

The proof of Theorem 1 is by Dagsvik (2003a) but for the reader's convenience it is also provided in Appendix A.

### **Corollary 1**

*Theorem 1 implies that the agent's utility function can be expressed as*

$$(3.3) \quad \tilde{U}(z, y) = \tilde{v}(z, y) \tilde{\varepsilon}(z, y), \quad z \in \Omega,$$

*where  $\tilde{v}(\cdot)$  is a positive deterministic mapping from  $\Omega \times [\gamma, \infty)$  to  $R_+$ , for some suitable constant  $\gamma > 0$ , and  $\tilde{\varepsilon}(z, y)$  is a random function satisfying*

$$P\left(\prod_{j=1}^n \left(\tilde{\varepsilon}(z_j, y) \leq x_j\right) \middle| \left\{v(z_k, y), k = 1, 2, \dots, n, y > 0\right\}\right) = P\left(\prod_{j=1}^n \left(\tilde{\varepsilon}(z_j, y) \leq x_j\right)\right)$$

for any set  $\{z_1, z_2, \dots, z_n\} \in \mathfrak{B}$ . The representation is unique in the sense that if

$U^*(z, y) = v^*(z, y) \varepsilon^*(z, y)$  is another representation, then there is a constant  $b > 0$  such that

$$v^*(z, y) = \tilde{v}(z, y)^b$$

and  $\varepsilon^*(z_j, y) / \varepsilon^*(z_k, y)$  has the same distribution as  $(\tilde{\varepsilon}(z_j, y) / \tilde{\varepsilon}(z_k, y))^b$ , for  $z_j, z_k \in \Omega, z_j \neq z_k$ .

The proof of Corollary 1 is given in Dagsvik (2003a) (Corollary 1).

The random terms  $\{\tilde{\varepsilon}(z, y)\}$  are supposed to capture the effect of unobserved variables that affect the agent's preferences<sup>6</sup>. Note that the property that the c.d.f. of  $\tilde{\varepsilon}(z, y)$ , for a given  $y$  is independent of the deterministic terms does not necessarily imply that the law of the stochastic process  $\{\tilde{\varepsilon}(z, y), y \geq \gamma\}$  is independent of the (deterministic) process  $\{\tilde{v}(z, y), y \geq \gamma\}$ .

In Corollary 1 we establish that the deterministic part of the utility function,  $\tilde{v}(z, y)$ , is unique up to an increasing power transformation. Thus, the deterministic part of the utility function is cardinal. However, this does not imply that the utility function, including the random term, that is  $\tilde{U}(z, y) = \tilde{v}(z, y) \tilde{\varepsilon}(z, y)$ , is cardinal.

For a given income  $y$  and a given choice set  $B = \{z_1, z_2, \dots, z_m\} \in \mathfrak{B}$ , let  $J(y)$  denote the index of the preferred attribute in  $B$ , i.e.,

$$J_B(y) = j \Leftrightarrow \tilde{U}(z_j, y) = \max_{z \in B} \tilde{U}(z, y).$$

## **Axiom 2 (DIM)**

For  $B \in \mathfrak{B}$

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

---

<sup>6</sup> Thurstone (1927) argued that these variables are uncertain, even to the agent himself in the sense that he is unable to fully assess the precise value of the alternatives (to him). Thus, if the agent faces repetitions of a choice experiment he may choose differently at each occasion.

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

For  $x = y$ , Axiom 2 is analogous to 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. Our motivation for DIM is as follows: Since the values of the alternatives are fully captured by the corresponding utilities, the information about which alternative that yields utility maximum should be irrelevant for the assessment of its value.

Moreover, Axiom 2 states that the choice that can be made under equal- or less income than  $y$  should be irrelevant for the evaluation of the most preferred alternative at income  $y$ . This is so because the alternatives available at income  $x$  also are available at income  $y$  when  $x \leq y$ . (Dagsvik 2002, has discussed an analogous axiom employed in the context of discrete choice over time.)

Let  $B$  be the agent's choice set which we assume belongs to  $\Omega$ , and let

$$U(y) = \max_{z \in B} \tilde{U}(z, y).$$

The process  $\{U(y), y \geq \gamma\}$  is the utility-of-income process. We shall now derive a very important characterization of this process, but first we need additional assumptions.

### **Axiom 3**

*The utility process is non-decreasing 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 (may be very small) that the agent's utility will remain constant even if income increases.

### **Condition 1**

*For any  $x \in R_+$  and  $y \geq \gamma$  there exists a  $z \in \Omega$  such that  $\tilde{v}(z, y) = x$ .*

Condition 1 states that for any (given)  $y \geq \gamma$ ,  $\tilde{v}(\cdot, y)$  maps  $\Omega$  onto  $R_+$ .

**Condition 2**

Let  $z_1, z_2, \dots, z_m$ , be distinct alternatives in  $\Omega$ . The  $m$ -dimensional utility process

$\left\{ \left( \tilde{U}(z_1, y), \tilde{U}(z_2, y), \dots, \tilde{U}(z_m, y) \right), y \geq \gamma \right\}$  (with  $y$  as parameter) is a multivariate max-stable process with type I extreme value marginals.

Recall that a max-stable process has finite dimensional distributions that are multivariate extreme value distributions. The type I distributions, defined on  $\mathbb{R}_+$ , have one-dimensional marginal distributions equal to  $\exp(-1/x)$  for  $x > 0$  (in standardized form). 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. The reason why is that the “multiperiod” random utility model (with income  $y$  as parameter) can be approximated arbitrarily closely by random utility models generated from max-stable processes.

**Condition 3**

The utility process  $\{U(y), y \geq \gamma\}$  is separable and continuous in probability.

**Theorem 2**

Assume that Axioms 1 to 3 and Conditions 1 to 3 hold. Then the process  $\{U(y), y \geq \gamma\}$  is an extremal process defined by

$$(3.4) \quad U(y_2) = \max(U(y_1), V(y_1, y_2))$$

for  $y_2 > y_1 \geq \gamma$ , with  $U(\gamma) = 0$ , where  $V(y_1, y_2)$  is a random variable with c.d.f.

$$(3.5) \quad P(V(y_1, y_2) \leq u) = \exp\left(-\frac{(v(y_2) - v(y_1))}{u}\right)$$

for  $u > 0$ , where  $v(y) = \sum_{z \in B} \tilde{v}(z, y)$  is increasing in  $y$  and  $v(y) = 0$  for  $y < \gamma$ . The random variable

$V(y_1, y_2)$  is independent of  $U(y_1)$ .

The proof of Theorem 2 is given in Appendix A. Note that if  $v(y) = 0$  for  $y < \gamma$ , then also  $U(y) = 0$ .

It follows from (3.4) and (3.5) that

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

which means that

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

where  $\varepsilon(y)$  is type I standard extreme value distributed, i.e.,

$$(3.6) \quad P(\varepsilon(y) \leq x | v(y)) = \exp\left(-\frac{1}{x}\right), \quad x > 0.$$

At first glance the result of Theorem 2 does not appear to be intuitive. However, the following discussion shows that the extremal process property is in fact quite intuitive. To this end consider the following discrete/continuous choice setting. Let  $z$  denote a  $n$ -dimensional vector of goods quantities and let  $j=1,2,\dots$  index "basic" discrete alternatives. Let  $U^*(z, j)$  denote the utility of the joint choice  $(z, j)$ . Suppose that

$$(3.7) \quad U^*(z, j) = m(z)b(j)\varepsilon(z, j),$$

where  $m(\cdot)$  and  $b(\cdot)$  are positive deterministic functions and  $\varepsilon(z, j)$ ,  $j=1,2,\dots$ , are i.i.d. with c.d.f. given by (3.6). For given  $j$ ,  $\{\varepsilon(z, j), z \in \mathbb{R}_+^n\}$  is a multiparameter max-stable process. The max-stable assumption means that the finite dimensional distributions of the process are multivariate extreme value distributed. The budget constraint is given by

$$(3.8) \quad pz + c_j \leq y,$$

where  $p$  is a vector of goods prices and  $c_j$  is the cost of alternative  $j$ . As a result, the utility of  $z$  can be expressed as

$$(3.9) \quad \tilde{U}(z, y) = \max_{c_j \leq y - pz} U^*(z, j).$$

The utility function  $\tilde{U}(z, y)$  is a conditional indirect utility, given  $z$ . For simplicity, we suppress the price vector in the notation here. Thus, the utility function  $\tilde{U}(z, j)$  is here interpreted as an "aggregate" utility function in the sense that some suitable basic goods have been "maximized out". In fact, one can in practice always think of utility functions in this way. For simplicity, we have here assumed that the basic goods are discrete, but the utility construction above holds with minor modifications also in the continuous case.



It follows immediately that  $\{\tilde{U}(z, y), y \geq \gamma\}$  for given  $z$ , is an extremal process with  $y$  as parameter. To realize this, note that for  $y_1 < y_2$ , we can write

$$\tilde{U}(z, y_2) = \max_{c_j \leq y_2 - pz} U^*(z, j) = \max \left( \max_{c_j \leq y_1 - pz} U^*(z, j), \max_{y_1 - pz < c_j \leq y_2 - pz} U^*(z, j) \right) = \max(\tilde{U}(z, y_1), V(z, y_1, y_2))$$

where

$$V(z, y_1, y_2) = \max_{y_1 - pz < c_j \leq y_2 - pz} U^*(z, j).$$

It follows from the distributional assumption of  $\{\varepsilon(z, j)\}$ , that  $V(z, y_1, y_2)$  is extreme value distributed, and that  $\tilde{U}(z, y_1)$  is independent of  $V(z, y_1, y_2)$ . Hence, we have demonstrated that  $\{\tilde{U}(z, y), y \geq \gamma\}$  is an extremal process.

Let  $z_1$  and  $z_2$  be two arbitrary consumption vectors and consider

$$\max(\tilde{U}(z_1, y), \tilde{U}(z_2, y)).$$

By Theorem 2 this process must also be extremal provided DIM holds. We shall next examine under which conditions  $\max(\tilde{U}(z_1, y), \tilde{U}(z_2, y))$  is also an extremal process. For this to be true one must have that  $\tilde{U}(z_k, y_1)$  and  $V(z_r, y_1, y_2)$  are independent for all  $r, k = 1, 2$ . Since it may happen that  $z_k p < z_r p$ , in which case the sets  $(-\infty, y_1 - z_k p)$  and  $(y_1 - z_r p, y_2 - z_r p)$  will overlap,  $\tilde{U}(z_k, y_1)$  and  $V(z_r, y_1, y_2)$  will be dependent unless  $\varepsilon(z_k, j)$  and  $\varepsilon(z_r, j)$  are independent. A completely similar argument applies in the general case for the indirect utility with respect to any set of consumption bundles. An immediate implication of this result is that, at the individual level, the probability of choosing a specific consumption bundle from a discrete set of feasible consumption bundles will satisfy IIA. Note that when unobserved heterogeneity is introduced, cf. Section 4 below, then IIA does not necessarily hold.

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

In this section we postulate an axiom which enables us to derive strong restrictions on the functional form of  $v(\cdot)$ . We assume first that the function  $v(y)$  is defined for  $y \geq \gamma$  (subsistence level) and is continuous and strictly increasing.

**Axiom 4**

Suppose that  $y_1, y_2, y_1^*, y_2^*$  are equal to or greater than  $\gamma$  and such that

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

Then for all  $\lambda > 0$

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

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 subsistence level are multiplied by an arbitrary positive constant  $\lambda$ . The intuition is as follows: Associate the different income levels  $y_1, y_2, y_1^*, y_2^*$  to consumption profiles (when prices are given). To the consumers relative income 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  $U(y_2) > U(y_1)$ , for  $y_2 > y_1$ , is *not* equal to one because there is a positive probability that  $U(y_2) = U(y_1)$ . The intuition for the latter 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.). We realize that if satiation can happen, then evidently Axiom 4 cannot in general be true.

**Theorem 3**

Assume that  $v(\cdot)$  is continuous. Then Axiom 4 implies that  $v(y)$  has the structure

$$(3.10) \quad v(y) = \exp \left( \delta \left( \frac{(y - \gamma)^\tau - 1}{\tau} \right) \right)$$

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

The proof of Theorem 3 is given in the Appendix A.

**Axiom 5**

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

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

Axiom 5 is evidently stronger than Axiom 4, and it means that income beyond subsistence level is perceived in a strict relative sense, that is, the number 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  $\lambda$ . Note that this property is *not* implied by Axiom 4.

**Theorem 4**

Assume that  $v(\cdot)$  is continuous. Then Axiom 5 implies that  $v(y)$  has the structure

$$(3.11) \quad v(y) = (y - \gamma)^\delta$$

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

A proof of Theorem 4 is given in the Appendix A.

We note that (3.11) is obtained as a special case of (3.10) with  $\tau = 0$ .

**4. Heterogeneity in preferences**

In the empirical specification described later we will specify some *observed* co-variates that affect the individual's evaluation of income. These observed co-variates capture some of the heterogeneity in the population, but obviously not all. Hence, 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 multiplicative component that varies across the population according to some probability distribution (random effect). Thus, the utility function, modified to include this random effect, becomes

$$(4.1) \quad U(y) = w v(y) \varepsilon(y),$$

where  $w > 0$  is the random effect. Note that the way we include  $w$  is analogous to allow for an additive constant term in an additive separable utility representation (which is seen by taking logarithm in (4.1)).

Now, let the random income process  $\{Y(u), u > 0\}$  be given by

$$(4.2) \quad U(Y(u)) = u.$$

Due to the properties of the utility function  $U(y)$ , which in fact is a stochastic process with parameter  $y$ , we have the following results.

**Proposition 1**

Assume that (3.4) and (3.5) hold. For  $0 < u_1 \leq u_2 \leq \dots \leq u_m$ , and  $\gamma \leq y_1 \leq y_2 \leq \dots \leq y_m$ , we have

$$(4.3) \quad G(y_1, y_2, \dots, y_m) \equiv P(Y(u_1) > y_1, Y(u_2) > y_2, \dots, Y(u_m) > y_m) = E \exp(-wH(y_1, y_2, \dots, y_m))$$

where

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

For  $y_1 < y_2 < \dots < y_m$ , the corresponding joint density of  $(Y(u_1), Y(u_2), \dots, Y(u_m))$  equals

$$(4.4) \quad g(y_1, y_2, \dots, y_m) = v'(y_m)u_m^{-l} \prod_{j=1}^{m-1} (u_j^{-l} - u_{j+1}^{-l}) v'(y_j) E(w^m \exp(-wH(y_1, y_2, \dots, y_m))).$$

A proof of Proposition 1 is given in the Appendix.

**Corollary 2**

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

$$(4.5) \quad Y(u_1) = \min \left( Y(u_2), v^{-1} \left( \frac{u_1 u_2 w \eta(u_1, u_2)}{u_2 - u_1} \right) \right)$$

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  $Y(u_2)$ .

The proof of Corollary 2 follows readily since any finite dimensional marginal distribution function of the process  $\{Y(u), u > 0\}$ , given by (4.5), has a form that is analogous to the corresponding ones for the utility process given by (3.4).

Next, we shall propose a theoretical justification for the distribution of  $w$ . For the sake of notational precision 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

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

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  $A$  of population  $S$ , 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  $S \setminus A$  is irrelevant for assessing who is the highest ranked individual in  $S$ . We recognize Axiom 6 as a particular version of the Independence from Irrelevant Alternatives Property, cf. Luce (1959a). We note that Axiom 6 requires that individual utilities can be compared and ranked. If utility is ordinal, then Axiom 6 only permits a common monotonic transformation of the utility function across individuals. If utility is cardinal, then a common affine transformation is implied.

**Theorem 5**

Assume that Axioms 1, 2, 3 and 6 and Conditions 1 to 3 hold, and that  $w_i$  and  $\varepsilon_i(y_i)$  are independent across the individuals,  $i \in S$ . Then the distribution of  $w_i$  is strictly  $\alpha$ -stable and totally skew to the right with  $\alpha < 1$ .

The proof of Theorem 5 is given in the Appendix.

Recall that the family of  $\alpha$ -stable distributions, often denoted by  $\{S_\alpha(c, \beta, \mu)\}$ , is characterized by four parameters, namely  $(\alpha, c, \beta, \mu)$ , where  $0 < \alpha \leq 2$  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  $\mu$  is a location parameter. When  $\alpha = 2$ , one obtains as a special case the normal distribution. It is strictly  $\alpha$ -stable when  $\mu = 0$  and totally skew to the right when  $\beta = 1$ . When  $\alpha < 1$ , neither the variance nor the mean of the stable random variable exist. (See Samorodnitsky and Taqqu, 1994).

As mentioned in Section 3, the choice among characteristics will, under the assumptions of Theorem 2, satisfy IIA. This is still true if (4.1) holds since the "random effect"  $w$  vanishes in utility comparisons. IIA does not seem overly restrictive in our case since the characteristics have not been given an explicit empirical content. It is however possible to motivate more general representations of

unobserved heterogeneity, as we shall discuss below. This extension consists in extending the utility representation to

$$(4.6) \quad \tilde{U}(z, y) = \tilde{w}(z)\tilde{v}(z, y)\tilde{\varepsilon}(z, y),$$

where  $\{\tilde{w}(z), z \in \Omega\}$  is a strictly stable process that is totally skew to the right with  $\alpha < 1$ . It can be demonstrated that (4.6) implies that the choice of characteristics model will have a Generalized Extreme Value (GEV) structure (Dagsvik, 2003b). 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  $\{w(z), z \in \Omega\}$  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, we have in this paper chosen to base our empirical model on the special case (4.1) rather than on (4.6). However, in Appendix B we have given the details of the derivation of the likelihood function when (4.6) holds.

## 5. Empirical specification

Motivated by the result of Theorem 4 we assume that

$$(5.1) \quad v(y) = \exp\left(\frac{\left(\frac{y-\gamma}{s}\right)^\tau - 1}{\tau\sigma}\right),$$

where  $\tau \leq 1$ ,  $\sigma > 0$  are parameters. By comparing (5.1) and (3.10) we note that  $\delta = 1/\sigma$ . The term  $s$  is assumed to be a function of variables that controls for variation in size and household composition. In fact  $s$  can be interpreted as a subjective equivalence scale. We assume that

$$(5.2) \quad \log s = b_1 \log N + b_2 C_{<6} + b_3 C_{7-15} + b_4 C_{>16} + b_5 Y + b_6 D + b_7 E,$$

where  $N$  is the size of the household,  $C_{<6}$  is the number of children of age less than or equal to 6,  $C_{7-15}$  is the number of children with age between 7 and 15,  $C_{>16}$  is the number of children with age equal to or greater than 16,  $Y$  is household income after tax,  $D$  is household debt and  $E$  is the length of schooling (in years) of the respondent.

Consistent with the result of Theorem 5, let

$$w \sim S_\alpha(c, 1, 0).$$

It turns out to be convenient to let  $c = k$ , where

$$(5.3) \quad k^\alpha = \frac{1}{\cos\left(\frac{\alpha\pi}{2}\right)}.$$

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 following lemma will be useful for establishing the likelihood function.

**Lemma 2**

Let  $w$  be  $\alpha$ -stable,  $S_\alpha\left(\left(\cos\left(\frac{\alpha\pi}{2}\right)\right)^{-1/\alpha}, 1, 0\right)$  with  $\alpha < 1$ . Then for  $\lambda \geq 0$

$$(5.4) \quad E \exp(-\lambda w) = \exp(-\lambda^\alpha),$$

$$(5.5) \quad \begin{aligned} \psi(\lambda; \alpha) := E(w^5 \exp(-\lambda w)) = & \left[ (\alpha \lambda^{\alpha-1})^5 + 10(1-\alpha)\alpha \lambda^{\alpha-2} (\alpha \lambda^{\alpha-1})^3 \right. \\ & + 10(1-\alpha)(2-\alpha)\alpha \lambda^{\alpha-3} (\alpha \lambda^{\alpha-1})^2 + 15(1-\alpha)^2 \alpha^2 \lambda^{2\alpha-4} (\alpha \lambda^{\alpha-1}) \\ & + 5\alpha(1-\alpha)(2-\alpha)(3-\alpha) \lambda^{\alpha-4} \alpha \lambda^{\alpha-1} + 10(1-\alpha)^2 \alpha^2 (2-\alpha) \lambda^{2\alpha-5} \\ & \left. + \alpha(1-\alpha)(2-\alpha)(3-\alpha)(4-\alpha) \lambda^{\alpha-5} \right] \exp(-\lambda^\alpha). \end{aligned}$$

Moreover, it is convenient to reparametrize the model by introducing  $\{a_j\}$  defined by

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

for  $j \leq 5$  and  $a_6 = \sigma \log u_6$ , where  $u_j$  are unknown threshold levels associated with the ordered structure of the questionnaire we are using.

By inspection of the likelihood function it is easily realized that it is increasing in  $\gamma$  so that the maximum likelihood estimate equals  $\hat{\gamma} = Y(u_1)$ . We shall use the remaining observations  $Y(u_2), Y(u_3), \dots, Y(u_6)$ , to estimate the remaining parameters. In the following we treat  $\tilde{\gamma}$  as a given parameter, which means that we ignore that  $\tilde{\gamma}$  is an estimate that involves error. From Proposition 1, (5.1) and Lemma 2 we get

$$\begin{aligned}
(5.6) \quad G(y_2, y_3, \dots, y_6) &\equiv P(Y(u_2) \geq y_2, Y(u_3) \geq y_3, \dots, Y(u_6) \geq y_6) \\
&= \exp \left[ - \left\{ \sum_{j=2}^6 \exp \left( \frac{\left( \frac{y_j - \hat{\gamma}}{s} \right)^\tau - 1 - \tau a_j}{\tau \sigma} \right) \right\}^\alpha \right],
\end{aligned}$$

and

$$\begin{aligned}
(5.7) \quad g(y_2, y_3, \dots, y_6) \\
= G(y_2, y_3, \dots, y_6) \prod_{j=2}^6 \left( \exp \left( \frac{\left( \frac{y_j - \hat{\gamma}}{s} \right)^\tau - 1 - \tau a_j}{\tau \sigma} \right) \frac{(y_j - \hat{\gamma})^{\tau-1}}{\sigma} \right) \cdot \psi \left( \sum_{j=2}^6 \exp \left( \frac{\left( \frac{y_j - \hat{\gamma}}{s} \right)^\tau - 1 - a_j \tau}{\tau \sigma} \right); \alpha \right)
\end{aligned}$$

## 6. 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 above. Actually, as shown in Table 2 a revised scheme was used with only 6 income intervals.  $Y(u_i)$ ,  $i = 1, 2, \dots, 5$ , means the upper bound of intervals, while  $Y(u_6)$  means that an income at this level is considered to be “very good”.  $Y$  stands for disposable household income and  $u_i$  indicates the utility level related to the income bound  $i$ .

**Table 2. Income Evaluation Questions. All numbers are in NOK as of September 1995**

|                |          | Mean    | Standard deviation | Minimum | Maximum   |
|----------------|----------|---------|--------------------|---------|-----------|
| Very bad       | $Y(u_1)$ | 159 950 | 76 640             | 30 000  | 500 000   |
| Bad            | $Y(u_2)$ | 202 290 | 91 200             | 50 000  | 600 000   |
| Insufficient   | $Y(u_3)$ | 246 110 | 110 650            | 70 000  | 700 000   |
| Sufficient     | $Y(u_4)$ | 294 340 | 131 300            | 100 000 | 800 000   |
| Good           | $Y(u_5)$ | 346 880 | 155 650            | 120 000 | 1 000 000 |
| Very good      | $Y(u_6)$ | 433 490 | 213 870            | 140 000 | 1 500 000 |
| # observations |          | 254     |                    |         |           |



The response rate was slightly above 50 per cent, of which 254 proved able to fill in answers on all the income intervals in the income evaluation scheme, and where none of the income bounds were equal. 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 3 below gives the summary statistics for the observed covariates.

**Table 3. Summary statistics for the observed covariates**

|                             | Mean  | Standard deviation | Minimum | Maximum |
|-----------------------------|-------|--------------------|---------|---------|
| Family size                 | 2.39  | 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                   | 16.42 | 3.07               | 8       | 24      |
| Income (100 000 NOK)        | 3.59  | 2.08               | 0       | 19      |
| Debt (100 000 NOK)          | 4.23  | 3.77               | 0       | 21      |
| # of observations           | 254   |                    |         |         |

Next we report maximum likelihood estimates based on the loglikelihood function

$$\ell = \sum \ln g_i(y_{i2}, y_{i3}, \dots, y_{i6})$$

where  $g_i$  is the joint density of the observation for individual  $i$ , as given in (5.7).

**Table 4. Maximum likelihood estimates of the parameters of the utility of income**

| Variables/parameters             |                             | Estimates | Standard deviation | t-value |
|----------------------------------|-----------------------------|-----------|--------------------|---------|
| Box-Cox exponent $\tau$          |                             | 0.0049    | 0.017              | 0.3     |
| Dispersion parameter $\sigma$    |                             | 0.1862    | 0.005              | 36.7    |
| Characteristic exponent $\alpha$ |                             | 0.3120    | 0.015              | 21.1    |
| Equivalent scale $s$             | Family size                 | 0.2459    | 0.133              | 1.8     |
|                                  | # children less than 6      | -0.1140   | 0.081              | -1.4    |
|                                  | # children between 7 and 15 | -0.0293   | 0.079              | -0.4    |
|                                  | # children above 16         | -0.2282   | 0.084              | -2.7    |
|                                  | Income (100 000 NOK)        | 0.0983    | 0.024              | 4.0     |
|                                  | Education                   | 0.0103    | 0.011              | 0.9     |
|                                  | Debt (100 000 NOK)          | -0.0100   | 0.013              | -0.8    |
| Transformed utility levels       | $a_2$                       | -1.0647   | 0.231              | -4.6    |
|                                  | $a_3$                       | -0.4150   | 0.228              | -1.8    |
|                                  | $a_4$                       | 0.0166    | 0.228              | 0.1     |
|                                  | $a_5$                       | 0.3520    | 0.228              | 1.5     |
|                                  | $a_6$                       | 0.7745    | 0.229              | 3.4     |
|                                  | Loglikelihood               |           | -346.6             |         |

First, we notice that the estimate of the characteristic exponent  $\alpha$  is less than one which means that neither the mean or variance exist of the random variable that captures unobserved heterogeneity in preferences. Second, we observe that the Box-Cox exponent  $\tau$  is not significantly different from zero. Thus, the deterministic part of the utility function is not significantly different from the power function given in (3.11). The income variable affecting  $s$  may be interpreted as capturing preference drift. The impact on  $s$  is significant and positive, which imply that the utility of income, given the income, and evaluated by the deterministic part of the utility function, is shifted downwards with the preference drift variable. In further use of this utility function the income appearing in the  $s$ -parameter may be associated with lagged income. The only other variable that is significant is the number of children above 16 years of age. The estimated impact on  $s$  is negative, which means that the higher the number of these children are, the stronger the upward shift is in the utility of income. Finally we observe that  $a_j$  is increasing in  $j$ , which is as desired.

Table 5 gives the estimates of regressing the observed minimum levels  $\{\hat{\gamma}_i\}$  against selected covariates.

**Table 5. Estimates of the effects of co-variates on the minimum level,  $\gamma$** 

| Variable               | Estimates | t-values |
|------------------------|-----------|----------|
| Constant               | 0.4414    | 1.84     |
| Family size            | 0.2097    | 2.29     |
| Child less than 6      | -0.0723   | -0.58    |
| Child between 7 and 15 | -0.1424   | -1.16    |
| Child above 16         | -0.1237   | -1.02    |
| Income                 | 0.1292    | 5.42     |
| Education              | 0.0056    | 0.44     |
| Debt                   | 0.0427    | 3.32     |

As expected the higher the family size and the higher the debt is, the higher is the minimum level. We observe that the higher household income is, the higher is also the required minimum income level.

## 7. The cardinal utility of income

### 7.1. Determination of the cardinal transformation

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 empirical evidence from the psychophysical literature not only seems to indicate that utility is perceived as cardinal, but also the specific functional form of this utility can be determined. Recall from Section 2 that the different type of psychophysical experiments imply that utility is cardinal. In particular, Stevens and his followers' claim that the average utility of income (average across individuals) is approximately proportional to the square root of income (possibly minus subsistence level), see Stevens (1975). This claim is established on the basis of a large number of "laboratory" experiments, of the type described in Section 2. Now let us examine if the results obtained in this paper can be made consistent with this claim from the psychology camp. From the empirical results to be discussed in Section 6 it follows that the utility function has the form

$$(7.1) \quad U(y) = (y - \gamma)^{1/\sigma} \varepsilon(y)w .$$

This yields

$$(7.2) \quad EU(y)^b = (y - \gamma)^{b/\sigma} M ,$$

where  $b$  is a suitable constant and

$$M = E(w \varepsilon(y))^b .$$

It can easily be demonstrated that  $M < \infty$  provided  $b < 1$ . We thus realize that with  $b = 0.5\sigma$ , our results are consistent with those of Stevens. Since we have estimated  $\sigma$  to be approximately equal to 0.186, we may claim that the cardinal utility of income is an extremal process in  $y$  given by

$$\{U(y)^b, y > 0\} \text{ with } b = 0.5\sigma = 0.093 .$$

## 7.2. The marginal utility of income

Consider next the notion of marginal utility with the present framework. Evidently  $U(y)^b$  is not differentiable in  $y$  (due to (3.4)). We shall therefore propose to study utility difference or ratios. Since our cardinal utility takes values on  $R_+$  and is unique up to a scale transformation, it may be natural to consider utility ratios.

We have the following result:

### Proposition 2

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

$$(7.3) \quad P\left(\left(\frac{U(y_2)}{U(y_1)}\right)^b \leq q\right) = \begin{cases} 0 & \text{if } q < 1 \\ \frac{q^{1/b}v(y_1)}{q^{1/b}v(y_1) + v(y_2) - v(y_1)} & \text{if } q \geq 1. \end{cases}$$

Moreover, we have that

$$(7.4) \quad E\left(\frac{U(y_2)}{U(y_1)}\right)^b = 1 + \frac{(v(y_2) - v(y_1))^\infty}{v(y_1)} \int_1^\infty \frac{bv(y_1)x^{b-1}dx}{v(y_1)x + v(y_2) - v(y_1)} .$$

The proof of Proposition 2 is given in the Appendix A.

From (7.4) it follows that

$$(7.5) \quad \begin{aligned} \lim_{y_2 \rightarrow y_1} \frac{E\left(\frac{U(y_2)}{U(y_1)}\right)^b - 1}{y_2 - y_1} &= \lim_{y_2 \rightarrow y_1} E\left(\frac{U(y_2)^b - U(y_1)^b}{(y_2 - y_1)U(y_1)^b}\right) \\ &= \lim_{y_2 \rightarrow y_1} \left(\frac{v(y_2) - v(y_1)}{y_2 - y_1}\right) \int_1^\infty x^{b-2} dx = \frac{v'(y_1)}{v(y_1)} \cdot \frac{b}{1-b} . \end{aligned}$$

Consider next the analogous concept based on the mean utility,  $EU(y)^b$ . It follows that

$$(7.6) \quad \lim_{y_2 \rightarrow y_1} \frac{EU(y_2)^b - EU(y_1)^b}{(y_2 - y_1)EU(y_1)^b} = \frac{\partial \log EU(y)^b}{\partial y} = \lim_{y_2 \rightarrow y_1} \frac{v(y_2) - v(y_1)}{(y_2 - y_1)v(y_1)} = \frac{v'(y_1)}{v(y_1)}.$$

If we compare (7.5) with (7.6) we realize that they are equivalent (equal apart from a multiplicative constant). This means that, whether one studies mean relative changes in utility, or relative changes of the mean utility, turns out to be equivalent at the margin. Due to the fact that our estimation results yield that  $v(y)$  is a power function we get that the marginal utility of income is equivalent to

$$(7.7) \quad \frac{\partial \log EU(y)^b}{\partial y} = \frac{1}{y - \gamma}$$

for  $y > \gamma$ .

We note that the elasticity of the marginal utility of income with respect to income is given by

$$(7.8) \quad -\frac{1}{1 - \frac{\gamma}{y}}$$

and it varies from  $-1(y = \infty)$  to  $-\infty(y = \gamma)$ .

## 8. 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 derived a characterization of the utility of income, viewed as a stochastic process in income. The basic assumptions are specific behavioral and invariance postulates, which we believe have intuitive appeal. These assumptions yield an explicit characterization of the probability law of the utility of income process. Specifically, it turns out that the implied utility function can be represented by an *extremal process*.

Subsequently, we have applied Stated Preference data to estimate the unknown parameters of the distribution of the utility of income process. 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). Finally, we have discussed how one can apply empirical results obtained by psychologists to determine a cardinal utility function, from which follows the marginal utility of income.

## Appendix A

### Proof of Theorem 1:

Debreu (1958) has proved that Axiom 1 implies that there exists a cardinal representation  $\kappa(z), z \in \Omega$ , such that

$$(A.1) \quad P(z_i, z_j) < P(z_k, z_r) \Leftrightarrow \kappa(z_i) - \kappa(z_j) < \kappa(z_k) - \kappa(z_r).$$

It is immediately verified that the Balance- and the Quadruple conditions are necessary for such a representation to exist. Debreu proved that if in addition the Solvability condition holds this imply (A.1).

Now let  $z_0 \in \Omega$  be fixed. Then it follows from (A.1) that if  $z_j$  and  $z_k$  satisfy

$P(z_j, z_0) = P(z_k, z_0)$  one must have that  $\kappa(z_j) = \kappa(z_k)$ . But this means that we can write

$$P(z_j, z_k) = \tilde{P}(\kappa(z_j), \kappa(z_k))$$

for some suitable function  $\tilde{P}$ . Now let  $z_j, z_0, z_k, z_r \in \Omega$  satisfy

$$P(z_j, z_0) = P(z_k, z_r) = \tilde{P}(\kappa(z_k), \kappa(z_r)) = \tilde{P}(\kappa(z_j), \kappa(z_0))$$

where  $z_0$  is fixed. By (A.1) this can hold only if

$$\kappa(z_j) = \kappa(z_0) + \kappa(z_k) - \kappa(z_r)$$

which yields

$$P(z_k, z_r) = \tilde{P}(\kappa(z_j), \kappa(z_0)) = \tilde{P}(\kappa(z_0) + \kappa(z_k) - \kappa(z_r), \kappa(z_0)) = G(\kappa(z_k) - \kappa(z_r))$$

where

$$G(x) = \tilde{P}(\kappa(z_0) + x, \kappa(z_0)).$$

Evidently  $G(x)$  is increasing and take values in  $[0,1]$ . Without loss of generality it can be chosen to be a c.d.f.

Q.E.D.

**Proof of Theorem 2:**

Dagsvik (2002) proves that the assumptions of the Theorem imply that

$$(A.2) \quad U(y_2)h(y_2) = \max(U(y_1)h(y_1), V(y_1, y_2))$$

where  $h$  is a positive deterministic function. This implies that the support of  $U(y_1)$  and  $U(y_2)$  satisfies

$$\frac{U(y_2)}{U(y_1)} \geq \frac{h(y_1)}{h(y_2)}.$$

Evidently, Axiom 3 implies that  $h(y_1) \geq h(y_2)$ . This is so because otherwise the support of  $(U(y_1), U(y_2))$  would allow that  $U(y_2) < U(y_1)$ . But  $h(y_1)$  cannot be strictly greater than  $h(y_2)$  because then the probability that  $U(y_2) = U(y_1)$  equals zero, and this contradicts Axiom 3. Thus we must have that  $h(y_1) = h(y_2)$ . Without loss of generality we can therefore choose  $h(y) = 1$ .

Q.E.D.

**Proof of Theorem 3:**

From Theorem 2 it follows for  $y_1, y_2 > \gamma$ , that

$$(A.3) \quad P(U(y_2) > U(y_1)) = P(V(y_1, y_2) > U(y_1)).$$

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

$$(A.4) \quad P(V(y_1, y_2) < U(y_1)) = \frac{v(y_1)}{v(y_1) + v(y_2) - v(y_1)} = \frac{v(y_1)}{v(y_2)}$$

for  $y_2 \geq y_1 \geq \gamma$ . Note that the right hand side of (A.4) is strictly increasing in  $v(y_1)/v(y_2)$ . When we combine (A.3) and (A.4) 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<sup>7</sup>

(A.5)

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

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

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

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

Let

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

where  $a \in (\gamma, y_1)$  is fixed. From (A.5) we get

$$(A.6) \quad \frac{v(a)}{v(y_2)} = F(x + z)$$

and

$$(A.7) \quad \frac{v(a)}{v(y_1)} = F(z).$$

When (A.6) and (A.7) are combined with (A.5) we get

---

<sup>7</sup> Note that Theorem 14.19, p.338, in Falmagne (1985) can be expressed more compactly as

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

where  $\frac{x^\theta - 1}{\theta}$  is defined as  $\lim_{\theta \rightarrow 0} (x^\theta - 1)/\theta = \log x$ .



$$(A.8) \quad F(x+z) = F(x)F(z).$$

Eq. (A.8) is a Cauchy functional equation which only continuous solution is the exponential function. Consequently, for  $y \geq \gamma$ ,

$$(A.9) \quad \log v(y) = \frac{\beta((y-\gamma)^\tau - 1)}{\tau}.$$

Q.E.D.

**Proof of Theorem 4:**

From (A.3) and (A.4) we get that

$$(A.10) \quad P(U(y_2) > U(y_1)) = 1 - \frac{v(y_1)}{v(y_2)}.$$

Hence, Axiom 5 implies that

$$(A.11) \quad \frac{v(\lambda(y_2 - \gamma) + \gamma)}{v(\lambda(y_1 - \gamma) + \gamma)} = \frac{v(y_2)}{v(y_1)}$$

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

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

for  $x \geq 0$ . With  $y_1 = \gamma + 1$ ,  $y_2 = x + \gamma$ , we get from (A.11) that

$$(A.12) \quad g(\lambda x) = g(x)g(\lambda).$$

But (A.12) is a functional equation of the Cauchy type which only continuous solution is the power function

$$(A.13) \quad g(x) = x^\delta$$

for some constant  $\delta$ . Since

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

the result of Theorem 4 follows (apart from an arbitrary multiplicative and positive constant).

Q.E.D.

**Proof of Proposition 1:**

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

$$\begin{aligned}
 G(y_1, y_2) &= P(U(Y(u_1)) \geq U(y_1), U(Y(u_2)) \geq U(y_2)) \\
 (A.14) \quad &= P(u_1 \geq U(y_1), u_2 \geq U(y_2)) = P(U(y_1) \leq u_1, \max(U(y_1), V(y_1, y_2)) \leq u_2) \\
 &= P(U(y_1) \leq \min(u_1, u_2), V(y_1, y_2) \leq u_2) = P(U(y_1) \leq \min(u_1, u_2))P(V(y_1, y_2) \leq u_2).
 \end{aligned}$$

Eq. (3.4) and (3.5) imply that

$$(A.15) \quad P(U(y_2) \leq u) = P(U(y_1) \leq u)P(V(y_1, y_2) \leq u).$$

Since

$$(A.16) \quad P(U(y) \leq u) = \exp(-v(y)u^{-1}), u > 0,$$

we obtain from (A.16) that for  $y_2 \geq y_1$

$$(A.17) \quad P(V(y_1, y_2) \leq x) = \exp(-(v(y_2) - v(y_1))x^{-1}), x > 0.$$

When (A.14) to (A.17) are combined we obtain that

$$(A.18) \quad G(y_1, y_2) = \exp(-v(y_1) \max(u_1^{-1}, u_2^{-1}) - (v(y_2) - v(y_1))u_2^{-1}).$$

In particular, when  $u_2 > u_1$ , (A.18) reduces to

$$(A.19) \quad G(y_1, y_2) = \exp(-v(y_1)(u_1^{-1} - u_2^{-1}) - v(y_2)u_2^{-1}).$$

The multivariate case is completely analogous. It follows readily that when  $u_1 \leq u_2 \leq \dots \leq u_m$ ,

$y_1 \leq y_2 \leq \dots \leq y_m$ ,

$$(A.20) \quad G(y_1, y_2, \dots, y_m) = \exp\left(-v(y_m)u_m^{-1} - \sum_{j=1}^{m-1} v(y_j)(u_j^{-1} - u_{j+1}^{-1})\right).$$

For  $y_1 < y_2 < \dots < y_m$ , the corresponding joint density function equals

$$(A.21) \quad g(y_1, y_2, \dots, y_m) = G(y_1, y_2, \dots, y_m) v'(y_m) u_m^{-1} \prod_{j=1}^{m-1} (u_j^{-1} - u_{j+1}^{-1}) v'(y_j).$$

The general case with random  $w$  now follows readily from (A.20) and (A.21).

Hence, the proof is complete.

Q.E.D.

**Proof of Theorem 5:**

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

$$(A.22) \quad E \left( \frac{w_i v(y_i)}{\sum_{r \in B} w_r v(y_r)} \right) = P \left( v(y_i) \xi_i = \max_{r \in B} (v(y_r) \xi_r) \right).$$

Since the error terms  $\{\xi_i\}$  are independent, we know from Yellott (1977) that Axiom 4 (IIA) can only be satisfied if the errors  $\{\xi_i\}$  are type I extreme value distributed, i.e.,

$$(A.23) \quad P(\xi_i \leq x) = \exp(-x^{-\alpha})$$

for  $x > 0$ , where  $\alpha$  is a positive constant. But (A.29) implies that

$$(A.24) \quad P(\xi_i \leq x) = P(w_i \varepsilon_i(y_i) \leq x) = E P \left( \varepsilon_i(y_i) \leq \frac{x}{w_i} \mid w_i \right) = E \exp \left( -\frac{w_i}{x} \right).$$

The last equality in (A.24) follows because by assumption  $P(\varepsilon(y) \leq x) = \exp(-1/x)$ . When we combine (A.23) and (A.24) we obtain that for any  $\lambda > 0$

$$(A.25) \quad E \exp(-\lambda w_i) = \exp(-\lambda^\alpha).$$

The left hand side of (A.25) is the Laplace transform of the distribution of  $w_i$ . From Samorodnitsky and Taquq (1994) Proposition 1.2.12, p. 15, it follows that when (A.25) holds  $w_i$  must be a strictly  $\alpha$ -stable random variable that is totally skew to the right and with  $\alpha < 1$ .

Q.E.D.

**Proof of Proposition 2:**

We have from (3.4) that  $U(y_2) = \max(U(y_1), V(y_1, y_2))$ , where  $y_2 \geq y_1 \geq \gamma$ . From Corollary 1 it follows that our utility transformation is unique up to a power transformation. Let  $b$  be a constant,  $0 < b < 1$ , and let  $q \geq 1$ . Hence,

$$(A.26) \quad P\left(\left(\frac{U(y_2)}{U(y_1)}\right)^b \leq q\right) = P(U(y_2) \leq q^{1/b}U(y_1)) = P(\max(U(y_1), V(y_1, y_2)) \leq q^{1/b}U(y_1)).$$

Note that  $V(y_1, y_2)$  and  $q^{1/b}U(y_1)$  are independent type I extreme value distributed with parameters  $v(y_2) - v(y_1)$  and  $q^{1/b}v(y_1)$ , respectively. Hence, from standard results in discrete choice theory we get that

$$(A.27) \quad P(V(y_1, y_2) \leq q^{1/b}U(y_1)) = \frac{q^{1/b}}{q^{1/b} + K(y_1, y_2)}$$

where

$$(A.28) \quad K(y_1, y_2) = \frac{v(y_2) - v(y_1)}{v(y_1)}$$

and where  $K(y_1, y_2) = 0$ , when  $y_1 = y_2$ .

Thus, we get

$$(A.29) \quad P\left(\left(\frac{U(y_2)}{U(y_1)}\right)^b \leq q\right) = \begin{cases} 0 & \text{if } q < 1, \\ \frac{q^{1/b}}{q^{1/b} + K(y_1, y_2)} & \text{if } q \geq 1. \end{cases}$$

This proves the first part of the Proposition.

Recall that one can express the expectation of a positive random variable  $X$  (say) as

$$EX = \int_0^{\infty} P(X > y) dy.$$

The expected value of the power transformation of the utility ratio equals

$$(A.30) \quad E\left[\left(\frac{U(y_2)}{U(y_1)}\right)^b\right] = \int_0^{\infty} P\left(\left(\frac{U(y_2)}{U(y_1)}\right)^b > q\right) dq = \int_0^1 dq + \int_1^{\infty} \frac{K(y_1, y_2)}{q^{1/b} + K(y_1, y_2)} dq.$$

By using the substitution  $x = q^{1/b}$ , we get

$$(A.31) \quad E\left[\left(\frac{U(y_2)}{U(y_1)}\right)^b\right] = 1 + K(y_1, y_2) \int_1^{\infty} \frac{bx^{b-1}}{x + K(y_1, y_2)} dx.$$

which proves the second part of the Proposition.

Q.E.D.

## Appendix B

Assume that (4.6) holds where  $\{\tilde{w}(z), z \in \Omega\}$  is a strictly stable process which is totally skew to the right with  $\alpha < 1$ . Under rather mild regularity conditions one can thus express  $\tilde{w}(z)$  as

$$(B.1) \quad \tilde{w}(z) = \sum_k h_k(z) \eta_k$$

where  $\{h_k(z)\}$  are suitable non-negative deterministic weights and  $\eta_1, \eta_2, \dots$  are i.i.d. strictly stable random variables that are totally skew to the right with same  $\alpha$  as the process  $\{\tilde{w}(z), z \in \Omega\}$ , cf.

Proposition 2.3.7, p.70, in Samorodnitsky and Taqqu (1994).

As a consequence we get from our Proposition 1 that

$$(B.2) \quad G(y_1, y_2, \dots, y_m) = E \exp\left(-\sum_k \eta_k S_k(y_1, y_2, \dots, y_m)\right),$$

where

$$(B.3) \quad S_k(y_1, y_2, \dots, y_m) = u_m^{-1} v_k^*(y_m) + \sum_{j=1}^{m-1} v_k^*(y_j) (u_j^{-1} - u_{j+1}^{-1})$$

and

$$(B.4) \quad v_k^*(y_j) = \sum_{z \in B} \tilde{v}(z, y_j) h_k(z).$$

Hence, by Lemma 2

$$(B.6) \quad G(y_1, \dots, y_m) = E \exp\left(-\sum_k \eta_k S_k(y_1, \dots, y_m)\right) = \exp\left(-\sum_k S_k(y_1, \dots, y_m)^\alpha\right).$$

Next suppose one the following extension of (5.1) holds;

$$(B.7) \quad v_k^*(y) = \psi_k \exp\left(\frac{\left(\frac{y-\gamma}{\tau\sigma_k}\right)^\tau - 1}{\tau\sigma_k}\right).$$

Then

$$(B.8) \quad S_k(y_1, \dots, y_m) = u_m^{-1} \psi_k \exp\left(\frac{\left(\frac{y_m - \gamma}{\tau\sigma_k}\right)^\tau - 1}{\tau\sigma_k}\right) + \sum_{j=1}^{m-1} (u_j^{-1} - u_{j+1}^{-1}) \psi_k \exp\left(\frac{\left(\frac{y_j - \gamma}{\tau\sigma_k}\right)^\tau - 1}{\tau\sigma_k}\right).$$

## References

- 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): A Behavioral Characterization of Intertemporal Random Utility Models in Continuous Time. Mimeo, Statistics Norway.
- Dagsvik, J.K. (2003a): Nonparametric Identification of Random Utility Models. Mimeo, Statistics Norway.
- Dagsvik, J.K. (2003b): Mixtures and Exact Aggregation in Random Utility Models. Work in progress, Statistics Norway.
- Debreu, G. (1958): Stochastic Choice and Cardinal Utility. *Econometrica*, **26**, 440-444.
- Ellingsen, (1994): Cardinal Utility: A History of Hedonimetry. In M. Allais and O. Hagen (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 D.H. Howes & E.C. Boring (eds.), H.E. Adler, trans. (Originally published in 1860.) Holt, Rinehart and Winston, New York.
- 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 J. H. Hollander (ed.) *Economic Essays in Honor of John Bates Clark*, The Macmillan Company, New York.
- Fischer, G.W. and D. Nagin (1981): Random versus Fixed Coefficient Quantal Choice Models. In Manski, C.F. and D. McFadden (eds.), *Structural Analysis of Discrete Choice Data*, MIT Press, Cambridge, Mass, USA.
- Frisch, R. (1926): Sur un Problème d’Économie Pure, *Norsk Matematisk Forenings Skrifter* **16**, 1-40, English translation in J. Chipman, L. Hurwicz., M. K. Richter and H. F. Sonnenschein (eds.) (1971), *Preferences, Utility and Demand*. Harcourt Brace Jovanovich Inc. London.
- Frisch, R. (1932): *New Methods of Measuring Marginal Utility*. Mohr, Tübingen.
- Frisch, R. (1947): Notes in Economic Theory, 4th edition. (In Norwegian). University Press, Oslo.
- 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. and S.J. Bolanowski, Jr. (1991a): Introduction to Conference on Ratio Scaling of Psychological Magnitudes. In S.J. Bolanowski, Jr. and G.A. Gescheider (eds.), *Ratio Scaling of Psychological Magnitude*, Lawrence Erlbaum Associates, Inc., Hillsdale, N.J., USA.
- Gescheider, G.A. and S.J. Bolanowski, Jr. (1991b): Final Comments on Ratio Scaling of Psychological Magnitudes. In S.J. Bolanowski, Jr. and G.A. Gescheider (eds.): *Ratio Scaling of Psychological Magnitude*, Lawrence Erlbaum Associates, Inc., Hillsdale, N.J., USA.
- Kapteyn, A. and T. Wansbeek (1985): The Individual Welfare Function. A Review. *Journal of Economic Psychology*, **6**, 333-363
- Lancaster, K.J. (1971): *Consumer Demand: A New Approach*. Columbia University Press, New York.
- 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.
- Patel, P and M. Subrahmanyam (1978): Utility Theory and Participation in Unfair Lotteries. *Journal of Economic Theory*, **19**, 555-57
- Samorodnitsky, G. and M.S. Taqqu (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.
- 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 (ed.), *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., R. D. Luce, D. Krantz and A. Tversky (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.
- van Herwaarden, F.G., A. Kapteyn and B.M.S. van Praag (1977): Twelve Thousand Individual Welfare Functions. *European Economic Review*, **9**, 283-300.
- van Herwaarden, F.G. and A. Kapteyn (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 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. and A. Kapteyn (1994): How Sensible is the Leyden Individual Welfare Function of Income? A Reply. *European Economic Review*, **38**, 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.