ORDINAL AND CARDINAL UTILITY:

AN INTEGRATION OF THE TWO DIMENSIONS OF THE WELFARE CONCEPT

by

Bernard M.S. van Praag*

December 1989

* Professor of Economics, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL 3000 DR Rotterdam, The Netherlands.
Member of the Scientific Council for Government Policy (WRR), Plein 1813 no.2, 2514 JN The Hague.

ORDINAL AND CARDINAL UTILITY: AN INTEGRATION OF THE TWO DIMENSIONS OF THE WELFARE CONCEPT

by

Bernard M.S.van Praag[1].

Abstract

In this paper we distinguish two "dimensions" of the utility concept. The first is the "behavioral" dimension, described by indifference curves in a commodity space. It may be estimated by observing consumer purchase behavior. The second dimension is the "welfare" dimension, i.e., the cardinal utility levels corresponding to indifference curves. The second dimension may be estimated by means of the income evaluation approach.
In this paper we deal with methodological issues and show by means of empirical evidence the validity of the income evaluation approach. In the same time we propose some major modifications of the method. Secondly we show how the two dimensions may be combined. This is illustrated with respect to the AIDS- and the Translog- model. In this way we find how price and income variations influence measured individual welfare.

Contents

1 I thank Robert Jan Flik for his invaluable support with the empirical part of this paper.

1. INTRODUCTION

One of the key concepts in economics is *utility* or *welfare*[2]. The first thorough introductions of the concept were that by Gossen (1854), Jevons (1871), Edgeworth (1881). They assumed that a commodity bundle **x** in the commodity space $(R^+)^n$ contained an intrinsic utility value U(**x**). The consumer problem could then be described as looking for the bundle with the highest utility value that could be bought at prices **p** and income y.

Such a model was able to describe and to predict purchase behavior. This was the *behavioral* aspect. But the model was also to be used for *normative* purposes, where we compare utility differences between bundles $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ for a specific individual. This is called *intra*personal comparison. The utility of income levels y_1, y_2, y_3 may be calculated by means of the *indirect utility* function $V(y,\mathbf{p})$ which is defined as the maximum utility to be derived from income y at given prices **p**.

This led to the progressive income taxation rules, suggested by Cohen Stuart (1889) a.o. Actually the latter use implies also that utility differences are comparable between individuals. This is called *inter*personal comparability.

Then it would also be possible to define social welfare functions $W(U_1,...,U_n)$ where social welfare is a function of individual utilities. The most obvious application of that concept is to compare distributions of social wealth and to devise policies which will lead to a better distribution.

Pareto (1909) gave a fierce blow to the utility concept by showing that demand behavior was completely determined by the contour lines, defined by the equation

$$U(\mathbf{x}) = \text{constant}$$

carved out on $(R^+)^n$; they are called the (utility) indifference curves.

The result is that demand behavior does not define the utility function uniquely, but rather that there is a whole equivalence class of utility functions which will yield the same demand behavior. Those utility functions $\bar{U}(\mathbf{x})$ have the property that $\bar{U}=\varphi(U)$ where $\varphi(.)$ is any monotonously increasing function.

---

2 _ We will use the words utility and welfare indiscriminately.

This eroded utility concept is called the *ordinal* concept. The original one of e.g. Edgeworth is called the *cardinal* utility concept. Samuelson (1947) remarked that Edgeworth thought utility to be "as real as his morning jam". Pareto (1909) did not state that cardinal utility was a nonsensical concept, (see also Kirman (1987)) but only that it was not necessary to know the utility function to explain demand behavior, as knowledge of the contour lines of the utility surface on the commodity space are all that we need. Nevertheless, this was a very helpful finding for our science as it proved very difficult to measure utility in practice.

Robbins (1932) made a fierce attack on cardinal utility and stated that it was an unmeasurable concept altogether. Hicks and Allen (1934) and later on Houthakker (1950) gave rigorous explanations of demand behavior without applying the utility concept at all. Deaton and Muellbauer (1982) make similar observations in their authoritative survey. So the utility concept degenerated into just a handsome tool to describe choice behavior.

Still there was an undercurrent of "true believers" in the utility of cardinal utility, which included famous names like Tinbergen who uses and defends the log(income) as a utility function and Frisch (1934). (See also Sen (1979)).

Harsanyi (1987) writes on interpersonal utility comparisons in the New Palgrave :

"It seems to me that economists and philosophers influenced by logical positivism have greatly exaggerated the difficulties we face in making interpersonal utility comparisons with respect to the utilities and the disutilities that people derive from ordinary commodities and, more generally, from the ordinary pleasures and commodities of human life."

See also Shubik (1982).

Indeed the whole literature on income inequality and poverty would be reduced to a sterile exercise if we do not accept the implicit cardinal utility measurement and interpersonal comparability on which these concepts are based[3].

In general the questions which are posed by reality (e.g. in physics, medicine, sociology or economics) are answered by the development of science.

---

3 See also Atkinson (1970), Jorgenson et al.(1984, 1990) and Apps and Savage (1989) who are implicitly using a cardinal interpersonally comparable welfare concept. However their cardinal utility functions have no empirically based welfare dimension, but they are based on implicit value judgments only.

If reality poses questions on which science has no answers, only few people would doubt that the correct way for scientific researchers will be to dig into these questions and to develop an operational theory as an answer.

However, in the case of utility history seems to have taken a different course. Reality abounds of normative questions with respect to distribution issues, income inequality, poverty, equivalence scales, etc. which are clearly answered in an intuitive way by policy makers in such a way that at least groups of the population feel that they agree with the policies proposed and the underlying value judgments. Such a utility base is also nearly indispensable for the evaluation of equilibria and the application of game-theoretic models (see also Shubik (1984)) . Nevertheless, it seems as if many economists have declared those questions non-issues which cannot be solved scientifically except for looking for Pareto -optimality. It is rather remarkable that mainstream economics for half a century since Robbins has followed a way which is so different from what is going on in the development of most sciences. Mostly science is *following* reality instead of *ignoring* it.

In this paper we will show that the cardinal dimension of the utility concept may and should be identified by the use of other data than those which can be derived from the observation of demand behavior.

In Section 2 we consider the precise relation between the cardinal and ordinal utility concepts on the commodity space and outline the way in which cardinal utility can be measured. It turns out that we need a combination of *two* measurement instruments, which we consider in Section 3. The first measurement tool is the observation of consumer behavior from household budgets as described in surveys; the second tool of measurement is the Income Evaluation Question (IEQ), as developed by a.o. Van Praag (1968, 1971), Van Praag, Kapteyn (1973) and Van Praag, Van der Sar (1988). In Section 4 we consider the possibilities of translating qualifying verbal labels into figures, where we leave the subject- matter which has to be evaluated. In Section 5 we describe the IEQ-method, and we consider its validity and results. In Section 6 we report on some empirically estimated relationships with respect to income evaluation. In Section 7 we present an integrated cardinal utility function on $(R^+)^n$ by merging the two types of information and we apply the method on two well- known demand models, viz., the Almost Ideal Demand System and the Translog Demand System. Section 8 concludes with a summary and discussion.

Up to now the IEQ- methodology has been used by practitioners but it has not been embedded in mainstream consumer theory for two reasons. The first was the use of specific assumptions ("equal quantiles" and "log- normality") in the measurement method, for which there was no independent empirical evidence. A second point was that it was unclear how price variations would influence the results. Both objections will be met in this paper.

2. THE BEHAVIORAL AND THE NORMATIVE ASPECTS OF UTILITY.

The utility hill in $(R^+)^{n+1}$ described by the function U(x) on $(R^+)^n$ may be sketched like in Figure 1. (see Shubik,(1984),p.43). The contour lines, corresponding to the utility levels U_1, U_2 are projected into the (x_1, x_2)-plane yielding the corresponding indifference curves, which we assume to be convex with respect to the origin. It is obvious that knowledge of the indifference curve implies knowledge of the demand curves, while knowledge of the demand behavior implies that the net of indifference curves may be constructed. Demand behavior says everything about "the *horizontal* dimension of utility". However, demand behavior as such says nothing about "the *vertical* dimension".

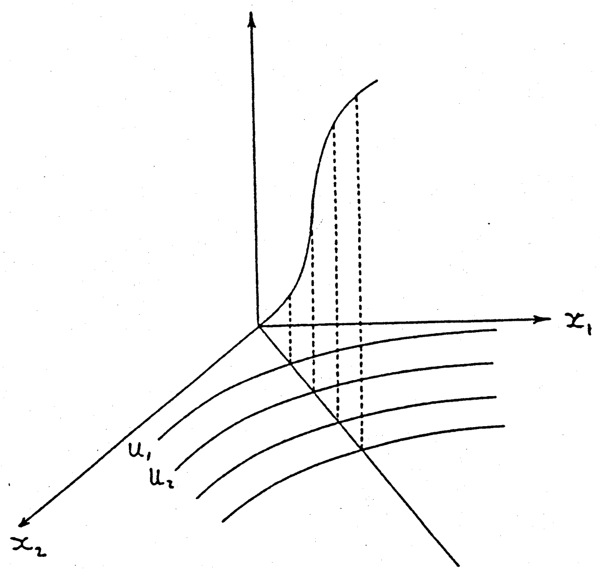

Figure 1. The utility hill.

The horizontal dimension defines an equivalence class of utility functions. Only if we put additional constraints to the functional specification such that in the equivalence class only one member satisfies the additional constraints, we are able to identify one member of the equivalence class as *the* utility function. We do not know of credible constraints.

A much better approach is to develop a measurement method to measure the vertical dimension. If we would succeed to identify indifference curves with utility labels U_1, U_2, etc., which can be identified as expressions of a level of well-being attached to those indifference curves we would have succeeded in the construction of a *cardinal* utility or welfare function.

Actually to succeed we only need to know the welfare evaluation U of *one* point at each indifference curve, as for by definition the welfare evaluation of all other points on that indifference curve is then known as well. It follows that we will have identified the cardinal welfare function, if we know the function $U(\mathbf{x}(t))$ where $\mathbf{x}(t)$ is a continuous and increasing vector function in $(R)^n$ with respect to t. The simplest way is clearly to take prices as given and to consider $U(x_1(y), x_2(y))=\tilde{U}(y)$ where $\mathbf{x}(y)$ stands for the demand vector when income y rises and the price vector $\mathbf{p}$ is constant.

Let us assume that we know the horizontal dimension of the utility function, either directly from the indifference curves or indirectly by knowing the household cost function $c(u;\mathbf{p})$, where $c(u;\mathbf{p})$ is the minimum cost (at prices $\mathbf{p}$) to reach the indifference curve labeled by the ordinal utility level u.

If $\tilde{U}(y)$ is the *true* cardinal utility level corresponding to $(y;\mathbf{p})$ for a fixed price vector, the relation between the ordinal utility label u and the true utility level $\tilde{U}$ is

$$\tilde{U}(y) = \tilde{U}(c(u;\mathbf{p}))$$

This relation describes the monotonically increasing function $\tilde{U}=\varphi(u)$, which links the ordinal utility label u to the (cardinal) height $\tilde{U}$ of the utility hill. The same procedure is possible with a variable price-structure, as long as we have one $\tilde{U}$-value for each indifference curve u.

Using this transformation we may now shift to the cardinal utility definition by introducing the cardinal version of the household cost function

$$y=c(\varphi(u);\mathbf{p}) = \tilde{c}(\tilde{U};\mathbf{p})$$

In this latter case

$$\frac{dy}{d\tilde{U}} = \frac{d\tilde{c}}{d\tilde{U}} = \frac{1}{\lambda}$$

where λ is the well-known marginal (cardinal) utility of income. At the moment the estimation of the ordinal household cost function c(u;p) is well-established in the literature. In the next two sections we shall concentrate on the empirical estimation of the function U(y), the vertical dimension.

3. MEASUREMENT OF THE CARDINAL DIMENSION.

The fundamental problem is evidently how to measure the cardinal[4] welfare evaluation U corresponding to an income level y. The problem is very similar to the measurement problem met in other disciplines. The starting point is a pre- scientific notion like temperature, time, loudness, humidity, hunger, health. We all know from empirical experience whether it is "warm" or "cold", what is "long ago" and what is "short ago", what sound is loud and what not, whether it is "dry" or "wet", whether we are "hungry" or not, whether we are feeling "healthy" or not. Nevertheless, it took a long time for mankind to devise a method of measurement for temperature that was acceptable for most people. The same holds for time where a reliable clock and calendar are relatively recent inventions. When such a measurement method and its unit of measurement is generally accepted, we face a process of convergence between the empirical notion and the measurable concept. It can still be observed when children learn to use the clock or the thermometer. Gradually " early in the day " is replaced by eight o'clock and "cold" by -5 C^o. The exact numerical descriptions get an emotional meaning as such, as they steadily refer to a situation which is considered to be the same. However, there are situations and people for whom eight o'clock is "late". In winter sport resorts -5 C^o may be rather hot. A temperature of -5 C^o may also be very cold when there blows a chilly wind.

The adoption of a specific frame of reference implies that the rather simple mostly one-dimensional measurement result is sometimes felt to be an inadequate description of the real empirical notion or rather the perception of it. It is an empirical question whether the measured counterpart of the

---

[4] We drop the tilde, when confusion is improbable.

empirical notion, as defined by the measurement method, is reflecting the empirical notion sufficiently. This depends of course on the degree of fit between the pre-scientific perceptions and the empirically measured observation results. The same caveats hold obviously and even to a larger extent when we try to measure feelings like "hunger" or a situation like "health" or the perception of brightness of light. The reason is clearly that those concepts are much more subject-related and much more subjective: that is, the same objective situation is differently perceived and evaluated by different individuals. Probably the evaluation made by one individual of his situation even varies over time according to his mood, etc..Nevertheless, such types of feelings may also be made more objective in the sense that there are basic objective explanatory factors.

The same holds for the evaluation of income by individuals, which will be the subject of this paper. How people evaluate their income is basically a feeling. If we like to know how an individual evaluates his own income or other levels of income we have to ask for it. More precisely, can we extract information from the individual what he considers to be a "good" income, a "bad" income, etc.?

We started in 1969 by formulating the so-called Income Evaluation Question (IEQ) which runs as follows:

> "Please try to indicate what you consider to be an appropriate amount for each of the following cases? Under my (our) conditions I would call a net household income per week/month/year of
>
> | | | |
> |---|---|---|
> | about | (250) | *very bad* |
> | about | (350) | *bad* |
> | about | (450) | *insufficient* |
> | about | (700) | *sufficient* |
> | about | (1200) | *good* |
> | about | (1600) | *very good* |
>
> Please enter an answer on each line and under line the period you refer to."

The answers between parentheses represent an answer by an American respondent

referring to monthly net household income. The idea of this question is that a verbal label sequence i=1,...,6 is supplied to the respondent and that the respondent reacts to this stimulus by giving income levels c_1 ,..., c_6. The respondent may select the payment period with which he is most familiar.
The use of this type of questioning is rather unfamiliar to mainstream economists. According to Sen's (1982, p.9.) statement on empirical economic methodology there is a strong predilection among economists for *observable behavior*. He pursues:

> *"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 solid information, where as intro-spection is not open to observation .... Even as behaviorism this is particularly limited since verbal behavior (or writing behavior, including response to questionnaires) should not lie outside the scope of the behaviorist approach."*

Indeed, the IEQ may be considered as an example of the observation of *verbal* behavior in the sense of Sen.
We do not stick to the precise wording of those labels in the IEQ. They have to be distinguishable from each other and they have to suggest a specific order. Several different wordings have been used both in Europe and in the U.S.A. Although most questionnaires have used a six-level question, in the earliest research (1971,1973) we utilized versions with eight or nine levels. Also those questions were well-answered, although the non-response rate was higher. In mail-back questionnaires the valid response is about 50%. In oral surveys it is over 90%.
For quantitative analysis of the I.E.Q. we need to translate those verbal levels into numerical figures. We devote the next section to this problem.

4. TRANSLATION OF EVALUATIVE VERBAL LABELS IN A CONTEXT-FREE SETTING

The translation of verbal evaluations into numerical figures is a matter of routine in many rating procedures. For instance, in Dutch universities and schools results of exams are rated on a (0,10)-scale where the significance of a 6 is "sufficient", 7 is "amply sufficient" and so on. The same numerical rating is used for rather esoteric things like the quality of music performances, ice-dancing, etc. As a matter of fact, frequently people become so used to numerical translations of evaluations that one forgets about the original verbal evaluations and takes the numerical figures themselves as evaluation; for instance "I got an eight for my exam". In England and the U.S.A. the usual letter evaluation ("A-level") plays a similar role.
Clearly the correspondence between verbal labels and figures depends on the *number* of different levels. The first question is then whether such a translation of verbal labels into numerical values is unambiguous. That is, *is the meaning of such a sequence of verbal labels about the same for all respondents*? The best way to study this question is to study it in a *context-free setting*, that is ,without specification of what matter is evaluated.
In order to get an answer on this question we carried out the following survey experiment (see also Van Doorn and Van Praag (1988)). In a survey where the respondent could type in his responses on a desk computer we asked two questions:

1. assign *numbers* between one and 1000 to 5 verbal labels, where one stands for the very worst and 1000 for the very best.

   very bad *bad* *not bad*, *not good* *good* *very good.*

2. assign *line segments* to the five labels, where a line segment of one unit stands for the worst and a line segment of 40 units for the best.

Notice that these two questions do not make any reference to an evaluation context. It does not refer to a specific subject-matter to be evaluated, e.g., income. The sample consists of 364 net observations.
The responses to the first question are denote by $v_{i,n}$ (i = 1, ...,5;

n= 1, ...,364) and the responses to the second question are denoted by $w_{i,n}$. Actually, the label sequence is intended to describe on a discrete scale a phenomenon which is evaluated "in the mind" on a continuous scale. The continuous value scale is the real interval [0,1] which is divided into adjacent intervals I_1 ,..., I_5, as sketched in Figure 2.

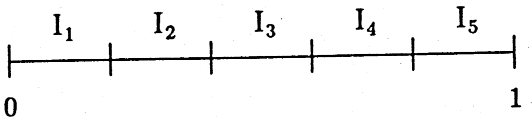

Figure 2. A discrete division of the evaluation scale.

The verbal scale is efficiently used if the underlying intervals have *equal length*[5]. That this is so may be seen as follows. Let us assume that a random phenomenon has to be evaluated by X and that the *a priori* distribution of values X offered is homogeneous on [0,1]. We write $p_i = P[X \in I_i]$. Then the entropy $\Sigma p_i \ln p_i$ is a measure which describes the discriminatory power of the discrete scale. That entropy is maximal if $p_1=p_2=...=p_5=1/5$. Hence, our prediction is that ideally v_i and w_i are situated at the midpoint of I_i, or that

$$v_i = w_i = (2i-1)/10 \qquad i=1,...,5$$

We call this the "equal-interval assumption" (E.I.A.).

In Table 1 we present the mean observations for both measurement methods with their corresponding sample standard deviations. In the last column we give our theoretical prediction (2i -1)/10..

5 Similar quantitative arguments may be found in Van Praag (1971) and a generalization of it in Kapteyn (1977).

Table 1a. Translation into numbers:

	Empirical mean	st.dev.	Th.Pred.
very bad	$\bar{v}_1 = 0.0892$	0.0927	0.10
bad	$\bar{v}_2 = 0.2013$	0.1234	0.30
not bad, not good	$\bar{v}_3 = 0.4719$	0.1117	0.50
good	$\bar{v}_4 = 0.6682$	0.1169	0.70
very good	$\bar{v}_5 = 0.8655$	0.0941	0,90

Table 1b. Translation into line-segments:

	Empirical	St.dev.	Th.Pred.
very bad	$\bar{w}_1 = 0.0734$	0.0556	0.10
bad	$\bar{w}_2 = 0.1799$	0.0934	0.30
not bad, not good	$\bar{w}_3 = 0.4008$	0.1056	0.50
good	$\bar{w}_4 = 0.5980$	0.1158	0.70
very good	$\bar{w}_5 = 0.8230$	0.1195	0.90

These results show that our predictions are not exactly verified. However, if we take account of the difficulty of the questions and the fact that there is an error term involved, we find the results quite satisfactory. In only one case ($\bar{w}_2$) the estimate lies outside a one σ-interval about the predicted value.

It appears that there is a downward bias of all w's and v's which is perhaps a measurement effect. Every respondent takes care not to take too much room in the beginning to keep space for the latter scale values.

The two measurement results of what is claimed to be two different descriptions of the same basic rating phenomenon are also highly consistent. We regressed v on w and we found the following regression

$$v_{i,n} = \underset{(0.0049)}{0.0557} + \underset{(0.0097)}{0.9737}\, w_{it} \qquad R^2=0.848$$

for $364 \times 5 = 1820$ observations, where we did not take into account that the level disturbances per individual will be strongly correlated.

On the basis of this exercise we conclude that

a. *a verbal label sequence is rather similarly understood by different respondents, irrespective of the context or the individual respondent,*

b. *a verbal label sequence may be translated either on a numerical scale or on a line scale; in both cases the translations are rather uniform over individuals,*

c. *translations via various translation mechanisms (lines and figures) are consistent with each other. That is we measure the same thing, irrespective of whether we use line segments or numbers.*

d. *the verbal evaluations are translated on a bounded scale in accordance with the Equal Interval Assumption.*

The interesting point is that these results have been found in a context-free setting, viz., the respondents did not know what concept they were evaluating.

5. THE EVALUATION OF INCOME BY THE I.E.Q.

Let us now return to the I.E.Q. introduced in the previous section. We are looking for the evaluation of c_i and we call that evaluation the *utility* or *welfare* derived from c_i. This is an empirical definition. If there is a "natural" numerical translation of a verbal label sequence, following the previous section, then it stands to reason that the evaluations of the k income levels $c_1,...,c_k$ are given by $v_1,...,v_k$ or by $w_1,...,w_k$ or by $(2i-1)/2k$ $(i=1,...,k)$. This is now justified as we showed in the previous section that verbal label sequences may be translated unambiguously to an interval scale.
It follows that we have to look for a functional specification $U_n(c)$, the welfare function of individual n or the Individual Welfare Function (WFI) such that

$$U_n(c_i) = (2i-1)/2k \qquad (i=1,...,k)$$

where U_n is a distribution function after appropriate scaling. We call this assumption of equal quantiles on the income axis the Equal Quantile Assumption

(EQA). Considering the response patterns of various respondents it is seen that a rich man may quote \$50,000 as a good net household income (c_5) while a poor man would call \$20,000 as his c_5-level. Hence, we have to recognize that different individuals n will evaluate income levels on different scales. Therefore we indexed the function U by n.
More precisely we parameterize the function as

$$U_n(c_i) = U(c_i;\theta_n)$$

where θ_n stands for an individual parameter vector. The next problem we are facing is *the specification of the function* U(.).

The evaluation of income and many other concepts is basically done by situating the income into the perceived distribution of incomes. It is comparative evaluation[6]. As all income distributions have an approximately log- normal shape, especially if we think on the smoothing regularizing perception of it by individuals, it stands to reason that we assume that

$$U_n(c) = \Lambda(c;\mu_n, \sigma_n) = N(\ln c;\ \mu_n, \sigma_n)$$

where Λ and N stand for the log-normal and normal distribution functions respectively. If the EQA is true, we will have per individual roughly symmetrical answers about μ_n and it follows that μ_n may be estimated by

$$m_n = \frac{1}{k} \sum_{i=1}^{k} \ln c_{i,n}$$

In a similar way we estimate σ_n^2 by

$$s_n^2 = \frac{1}{(k-1)} \sum_{i=1}^{k} (\ln c_{i,n} - m_n)^2$$

6 It is well-known that people when faced with an unusual phenomenon do not know how to evaluate it, as they have "nothing to compare it with". It is also well-known that student's performances are frequently evaluated on an explicitly comparative basis as one being in the top 5% or being in the lower half, etc. This link was first laid by Kapteyn, Wansbeek, Buyze (1980); see also Van Praag (1981, 1988).

(see also Van der Sar, Van Praag, Dubnoff (1988) and Van Praag (1988)).
Now we "standardize" the answers $c_{i,n}$ by defining

$$u_{i,n} = \frac{\ln c_{i,n} - m_n}{s_n}$$

Then the EQA combined with log- normality of U(.) implies that there should hold

$$N(u_{i,n};0,1) = (2i-1)/2k \qquad i=1,...,k$$

In our experimental questionnaire where we asked for the translation of the verbal *five*-label sequence into numbers v or the segments w we posed the IEQ as a *five*-level question as well. Then there should hold that the verbal labels in the IEQ referring to income should give the same stimulus as when these labels are used in a subject-free context. Hence, we expect

$$N((\ln c_{in} - \mu_n)/\sigma_n;\ 0,1) = v_{in} \qquad t=1,...,364 \quad i=1,...,5$$

and

$$N((\ln c_{in} - \mu_n)/\sigma_n;\ 0,1) = w_{in}$$

Here we do *not* assume that the EQA holds.
Estimating μ_n and σ_n as above from the 5-level question we should approximately get

$$\ln c_{i,n} \approx \mu_n + \sigma_n\, N^{-1}(v_{in})$$

$$\ln c_{i,n} \approx \mu_n + \sigma_n\, N^{-1}(w_{in})$$

OLS-regression yields

$$\ln c_{in} = \underset{(0.1605)}{-0.0398} + \underset{(0.0213)}{1.0136\ \mu_n} + \underset{(0.0090)}{0.6292\ \sigma_n N^{-1}(v_{in})} \qquad R^2=0.796$$

$$\ln c_{in} = \underset{(0.1392)}{-0.3521} + \underset{(0.0185)}{1.0634\ \mu_n} + \underset{(0.0097)}{0.8197\ \sigma_n\ N^{-1}(w_{in})} \qquad R^2=0.847$$

N=364

These results suggest that the hypothesis of log- normality holds approximately , even if we do not assume EQA. That it does not hold exactly seems to be due to random errors and inexperience of the respondents in answering the questions with respect to v and w.

Finally we consider the OLS-equation where we assume that the EQA holds as well. We find

$$\ln c_{in} = \underset{(0.0831)}{0.00005} + \underset{(0.0110)}{0.9999\ \mu_n} + \underset{(0.0065)}{0.9834\ \sigma_n N^{-1}\left(\frac{2i-1}{10}\right)} \qquad R^2= 0.945$$

The last result shows that the assumptions of log normality and equal quantiles are very reasonable indeed.

On the basis of this combined evidence it appears justified to assume that

$$U(c) = N((\ln c - \mu)/\sigma\ ;\ 0,1)$$

where μ and σ are estimated per individual from the IEQ by the formulae

$$m = \frac{1}{k}\sum_{i=1}^{k} \ln c_i, \qquad s^2 = \frac{1}{k-1}\sum_{i=1}^{k} (\ln c_i - m)^2$$

Additional evidence on a six- level version of the IEQ confirms the previous evidence. We consider the hypothesis for a sample of 500 American respondents, created by Dubnoff for the Boston region. We present in Table 2 the mean values of $u_{1,n},\dots,u_{6,n}$, the corresponding sample standard deviations $\sigma(u_i)$, the resulting average welfare values $N(u_i;0,1)$ and the predicted values according to the EQA.

Table 2. Average u-levels, sample s.d., resulting welfare levels and prediction

	$\bar{u}_i$	$\sigma(u_i)$	(u_i)	prediction
1.	-1.291	0.236	0.09	0.083
2.	-0.778	0.190	0.22	0.250
3.	-0.260	0.241	0.40	0.417
4.	0.259	0.239	0.60	0.583
5.	0.760	0.190	0.78	0.750
6.	1.311	0.229	0.90	0.917

We see that the prediction is pretty well validated by Table 2..

At this point we mention some important differences between this result and the now usual[7] method described in Van Praag (1971). In the 1971-approach we *a priori* postulated a log-normal function on the basis of theoretical arguments in Van Praag (1968). Also we postulated the EQA without empirical evidence for it.

Then we *estimated* μ_n and σ_n per person by applying OLS on the equation

$$\ln c_{in} = \mu_n + \sigma_n \hat{u}_i$$

per individual n, where u_i was based on the EQA-postulate.

In the present approach we reverse the line of reasoning. We estimate μ_n and σ_n^2 as the mean and the variance of the logarithmic IEQ-responses.

Having defined the u_{in} we see that the u_{in}'s may be interpreted as equal-quantile responses if U(.) is a $\Lambda(.;\mu,\sigma^2)$-distribution function. Hence, we conclude that the log- normality-assumption is acceptable.

It may be shown quite easily that the new μ-definition and the old one according to OLS are identical[8]. For σ^2 this is not true except in the case of exact log- normality. The present estimator and the original one appear to be

7 The IEQ is now experimentally used by EUROSTAT, most official statistical offices in the European Community and by Statistics Canada. It is especially used for the definition of the poverty line (see Goedhart et.al. (1977), Hagenaars (1986), Ghiatis (1989))

8 This was first brought to my attention by Reuben Gronau.

almost equal. Actually the correlation coefficient of the two estimators for σ applied to all individuals in the sample equals 0.99. This is additional strong evidence for the validity of the log normal distribution.

At this point we notice that there are many other functions suggested and estimated by Van Herwaarden, Kapteyn (1981). They concluded that the log-normal *and* the logarithmic function fitted best. As the logarithm is unbounded this does not conform with the usual everyday practice of a bounded evaluation interval.

6. EMPIRICALLY ESTIMATED RELATIONSHIPS

As already hinted at μ_n and σ_n^2 vary over individuals. The variance σ_n^2 (sometimes called the welfare sensitivity parameter) is very hard to explain by objective individual factors. It seems established that σ_n^2 is heavily dependent on the income inequality in a country as measured by the log-variance. This is in accordance with the idea that the function U(y) is a reflection of the objective income distribution function. Within a country we take the variable as constant, say, σ. The individual variable μ_n is explainable to a large extent as has been shown in many of the references. We refer to those papers for details. The basic relationship, first established by Van Praag (1971) and Van Praag, Kapteyn (1973) appears to be

$$\mu_n = \beta_0 + \beta_1 \ln fs_n + \beta_2 \ln y_{c,n}$$

were fs stands for family size (head count) and y_c for current net household income. The typical estimates are

$$\mu_n = \beta_0 + 0.10 \ln fs_n + 0.60 \ln y_{c,n} \qquad R^2 \approx 0.6$$

The intuitively plausible result says that μ_n depends on the own *current* income level. It follows that the indirect utility function drifts with the level of current own income $y_{c,n}$. We may write

$$U_n(c) = N((\ln c - \mu_n) / \sigma;\ 0,1)$$

$$= N((\ln c - \beta_0 - \beta_1 \ln fs_n - \beta_2 \ln y_{c,n})/\sigma \; ; \; 0,1)$$

We have called this effect *preference drift* and the parameter β_2 the *preference drift rate*.

Obviously this drift will not be immediately realized as a result of an income change. Let us assume that log-income shifts from $\ln y_c^{(0)}$ to $\ln y_c^{(1)}$, the difference being Δy_c. Then we may define the *ex ante* $\mu_n^{(0)}$ and the *ex post* $\mu_n^{(1)}=\mu_n^{(0)}+\beta_2\Delta y_c$ accordingly. It follows that people with different incomes will have different (cardinal) utility functions. This variation will increase if β_2 increases. If $\beta_2=1$ it would imply that everybody would evaluate his own income by

$$N((-\beta_0-\beta_1 \ln fs)/\sigma)$$

irrespective of his income level. Fortunately almost always β_2 is estimated far below one.

The ex ante function with fixed μ_o is called the *virtual* WFI. The function $N(\ln y_c - \mu(y_c);0,\sigma)$ is called the *true* WFI. It describes the welfare evaluation of an individual for his own current income.

Similarly it is clear that family size has a (cost) effect on the evaluation of income. Family equivalence scales are derived by setting

$$\ln y_c - \beta_0 - \beta_1 \ln fs - \beta_2 \ln y_c = \text{constant}$$

which yields a constant welfare equation

$$\ln y_c = \frac{\beta_0+\beta_1 \ln fs}{1-\beta_2}$$

We may also interpret it as the *true* household cost function.
We find the family size elasticity

$$\frac{d\ln y_c}{d\ln fs} = \frac{\beta_1}{1-\beta_2}.$$

We refer to Van Praag, Van der Sar (1988) for similar findings in an ordinal setting.

7. THE INTEGRATION INTO ORDINAL CONSUMER BEHAVIOR MODELS.

In this Section we shall link our results with the traditional neo-classical models. We shall show that the vertical dimension can be added in a fairly simple way to ordinal models. As examples we shall consider the Almost Ideal Demand System, developed by Deaton and Muellbauer (1980) and the Translog Demand System, originally proposed by Christensen, Jorgenson and Lau(1975) in its version by Jorgenson and Slesnick(1983).
Both models are based on the simple cost function

$$\ln c = u\, b(p) + a(p)$$

where u stands for an ordinal utility label. In this case we may write the ordinal utility-index as an explicit (indirect utility) function of c and p, viz.

$$u = \frac{\ln c - a(p)}{b(p)}$$

Let us assume we know the cardinal utility function $U_n(c;p_o)$ for a specific p_0 vector and there holds for a specific individual with (μ_0,σ_0)

$$U_n(c;p_0) = N\left(\frac{\ln c-\mu_o}{\sigma_0};\ 0,1\right).$$

It follows that we may write $\mu_0 = a(p_0)$, $\sigma_0=b(p_0)$ in which case

$$u = \frac{\ln c-\mu_0}{\sigma_0}$$

and

$$U = \varphi(u) = N(u;\ 0,1).$$

Now we extend this to arbitrary price vectors p. We define $B(p) = b(p)/b(p_0)$ and $A(p) = a(p)/a(p_0)$.

Then we may write for general prices

$$u = \frac{\ln c - \mu_0 A(p)}{\sigma_0 B(p)}$$

and

$$U(c;p) = N\left(\frac{\ln c - \mu_0 A(p)}{\sigma_0 B(p)} \; ; \; 0,1\right)$$

with $A(p_0) = B(p_0) = 1$.

The corresponding cost function is

$$\ln c = u\sigma_0 B(p) + \mu_0 A(p)$$

where $u = N^{-1}(U;0,1)$.

The price-differentiated welfare parameters are then

$$\mu(p) = \mu_0 \cdot A(p)$$

and

$$\sigma(p) = \sigma_0 \cdot B(p)$$

If the demand system is described by the Almost Ideal Demand System (Deaton and Muellbauer (1980)) we have

$$a(p) = a_0 + \Sigma\alpha_k \ln p_k + \frac{1}{2}(\ln p)' \Gamma \ln p$$

$$b(p) = \beta_0 \, \Pi \, p_k^{\beta_k}$$

If the demand system is of the Translog-type we find the same (Jorgenson, Lau and Stoker (1980, 1981, 1982)) Gorman formula, where

$$b(p) = 1 / (1-\iota B_{pp}\ln p)$$

$$a(p) = a_0 + \Sigma\alpha_k \ln p + \frac{1}{2}(\ln p)'\Gamma \ln p.$$

Family size-effects are now included by setting

$$\mu_0^{(1)} = \mu_0^{(0)} + \beta_1 \Delta \ln fs.$$

Notice that this introduces a term in $\Delta\ln fs.\beta_1 A(p)$. This implies that price effects may depend on family size via an interaction term. We may write

$$\ln c = u\ \sigma_0\ B(p) + (\mu_0^{(0)} + \beta_1 \Delta \ln fs) A(p).$$

In a similar way we get dependence on own current income y_c by

$$\ln c = u\ \sigma\ B(p) + (\mu_0^{(0)} + \beta_1 \Delta fs + \beta_2\ \Delta \ln y_c)\ A(p)$$

From this formula we may find a *true* household cost function setting $\ln c = \ln y_c$. It follows that

$$\ln y_c = \frac{1}{1-\beta_2 A(p)} \left[u\ \sigma_0 B(p) + (\mu_0^{(0)} + \beta_1 \Delta \ln fs)\ A(p)\right]$$

The family size elasticity becomes

$$\frac{d\ln y_c}{d\ln fs} = \frac{1}{1-\beta_2 A(p)}\ \beta_1 A(p)$$

Consider now the indirect utility of log-income. It equals

$$U(c) = N\left(\frac{\ln c - \mu}{\sigma}\right)$$

and hence

$$U'(c) = \frac{dU}{dc} = \frac{1}{c}\ \frac{1}{\sqrt{2\pi}}\ \exp\left(-\frac{1}{2}\left(\frac{\ln c-\mu}{\sigma}\right)^2\right)$$

The relative risk aversion becomes (see also Van Praag (1971))

$$\frac{d\ln U'}{d\ln c} = -1- \frac{(\ln c-\mu)}{\sigma^2}$$

8. DISCUSSION AND CONCLUSION

In this paper we argue that a cardinal welfare function is a measurable concept. It may be split up into an ordinal and a cardinal part. The ordinal part is estimated in the traditional way by studying consumer demand behavior. The cardinal part, an indirect welfare function of income (WFI) , is estimated from the IEQ, a specific type of an attitude question battery. It is argued on empirical evidence that this function is approximately a log - normal distribution function with welfare parameters μ and σ. The two dimensions are knitted together into the cardinal welfare function. It is shown how the welfare parameters take their place in the Almost Ideal Demand System or the Translog Demand System.

The main new results in this paper are the following ones:

a. the log-normal form of the individual welfare function of income is derived from empirical evidence instead of postulated on the basis of a theoretical argument like in earlier studies,
b. the assumption of equal quantiles is verified on empirical evidence,
c. the WFI-concept is shown to be interpretable as an indirect utility or welfare function,
d. it is found how μ and σ change when prices change.
e. it is shown how any ordinal demand system can be combined with any cardinal welfare function.

In this paper we did not attempt to evaluate social welfare or welfare inequality. It can be shown that most inequality indices may be based on the log normal welfare function (see e.g. Van Praag (1975)). Also we do not make use of the empirically estimated welfare functions to evaluate market equilibria or to model games, based on utility functions. Clearly here lies a wide field of applications.

REFERENCES

Apps, P. and E. Savage (1989), "Labour Supply, Welfare Rankings and the Measurement of Inequality", Journal of Public Economics, **29**, pp.335-64.

Atkinson, A.B. (1970), "On the Measurement of Inequality", Journal of Economic Theory, **2**, pp.244-263.

Cohen Stuart, A.J. (1889), Bijdrage tot de Theorie der Progressieve Inkomensbelastingen, Nijhoff, The Hague..

Deaton, A. and J. Muellbauer (1980), "An Almost Ideal Demand system", American Economic Review, **70**, pp.312-26.

Deaton, A. and J. Muellbauer (1982), Economics and Consumer Behavior, Cambridge, Cambridge University Press.

Edgeworth, F.Y. (1881), Mathematical Psychics: an Essay on the Application of Mathematics to the Moral Sciences, London: C. Kegan Paul.

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

Ghiatis A., "Low Income Groups Obtained by Enhanced Processing of the household Budget surveys in the EC-Summary Figures for Italy and the Netherlands", presented at the EUROSTAT-seminar on Poverty Statistics in the European community, Noordwijk, 1989.

Goedhart, Th., V. Halberstadt, A. Kapteyn and B.M.S. van Praag (1977),"The Poverty Line: Concept and Measurement", The Journal of Human Resources, vol.12, pp.503-520.

Gossen, H.H. (1854), Entwicklung der Gesetze des Menschlichen Verkehres und der daraus fliebenden Regeln für menschliches Handeln, Braunschweig.

Hagenaars, A.J.M. (1986), The Perception of Poverty, Amsterdam: North Holland Publishing Cy.

Hicks, J.R. and R.G.D. Allen (1934),"A Reconsideration of the Theory of Value", Economica, vol.1, pp. 52-75, 196-219.

Houthakker, H.S. (1950), "Revealed Preference and the Utility Function", Economica, vol.17, pp.159-174.

Jevons, W.S. (1871) The Theory of Political Economy, MacMillan, London.

Jorgenson, D.W., L.J. Lau and T.M. Stoker (1980), "Welfare Comparison Under Exact Aggregation", American Economic Review, 70, pp.268-272.

Jorgenson, D.W., L.J. Lau and T.M. Stoker (1981), "Aggregate Consumer Behavior and Individual Welfare", in D. Currie, R. Nobay and D. Peel (eds.), Macro economic Analysis, London, Croom-Helm, pp.35-61.

Jorgenson, D.W. and D.T. Slesnick (1984), "Aggregate Consumer behaviour and the Measurement of Inequality", Review of Economic Studies, L1, pp.369-392.

Jorgenson, D.W. and D.T. Slesnick (1990), "Inequality and the Standard of living", Journal of Econometrics 43, pp. 103-121.

Kapteyn, A. (1977), A Theory of Preference Formation, unpublished Ph.D. thesis, Leyden University.

Kapteyn, A., T.J. Wansbeek and J. Buyze (1980), "The Dynamics of Preference Formation", Journal of Economic Behavior and Organization, 1, pp.123-157.

Kirman, A.P. (1987), "Pareto as an economist", the New Palgrave, a Dictionary of Economics, J. Eatwell et al. (eds).

Moriani, C. (1989), "Italian Experience on Poverty Harmonized Measures", presented at the EUROSTAT-seminar on Poverty Statistics in the European community, Noordwijk, 1989.

Pareto, V. (1909), Manuel d'Economie Politique, Paris: V. Giard & E. Brière.

Robbins, L. (1932), An Essay on the Nature and Significance of Economic Science, 1st. ed. London: MacMillan.

Samuelson, P.A. (1947), Foundations of Economic Analysis, Cambridge, Mass.: Harvard University Press.

Sen, A.K. (1979), "Interpersonal Comparisons of Welfare", in: M. Boskin (ed.), Economics and human welfare: Essays in honor of Tibor Scitovsky, New York, Academic Press.

Sen, A.K. (1982), Choice, Welfare and Measurement, Cambridge, Mass.: M.I.T. Press.

Shubik, M. (1982), Game Theory in the Social Sciences, Cambridge, Mass.: M.I.T. Press.

Shubik, M. (1984), A Game-Theoretic Approach to Political Economy, Cambridge, Mass.: M.I.T. Press.

Van der Sar, N.L., B.M.S. van Praag and S. Dubnoff (1988), "Evaluation Questions and Income Utility", in: B.Munier (ed.) Risk, Decision and Rationality, Reidel Publishing Company, Dordrecht, pp.77-96.

Van Herwaarden, F.G. and A. Kapteyn (1981) "Empirical Comparison of the Shape of Welfare Functions", European Economic Review, vol.15, pp.261-286.

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

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

Van Praag, B.M.S. and A. Kapteyn (1973), "Further Evidence on the Individual Welfare Function of Income: An Empirical Investigation in The Netherlands", European Economic Review, vol.4., pp.33-62.

Van Praag, B.M.S. (1978), "THe Perception of Income Inequality" in W. Krelle and A.F. Shorrocks (eds.),Personal Income Distribution, North-Holland Publ.Cy., Amsterdam.

Van Praag, B.M.S. (1981), "Reflections on the Theory of Individual Welfare Functions", Report 81.14, Center for Research in Public Economics, Leyden University, Proceedings of the American Statistical Association.

Van Praag, B.M.S. and N.L. van der Sar (1988), "Household Cost Functions and Equivalence Scales", in: Journal of Human Resources, vol.23, pp.193-210.

Van Praag, B.M.S. (1988), "The Relativity of the Welfare Concept", forthcoming in M. Nussbaum and A. Sen (eds): Quality of life, Clarendon press, Oxford.

Van Praag, B.M.S. and L. van Doorn (1988), "On the Measurement of Income Satisfaction", in: W.E. Saris and J.N. Gallhofer (eds.), McMillan Press Ltd., Houndmills, Sociometric Research, vol.1, pp.230-246, McMillan, London.