Praag, Bernard M.S. Van

Working Paper
The Connexion between Old and New Approaches to Financial Satisfaction

Tinbergen Institute Discussion Paper, No. 04-053/3

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Praag, Bernard M.S. Van (2004) : The Connexion between Old and New Approaches to Financial Satisfaction, Tinbergen Institute Discussion Paper, No. 04-053/3, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
http://hdl.handle.net/10419/86635

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.
The Connexion between Old and New Approaches to Financial Satisfaction

Bernard M.S. Van Praag

Universiteit van Amsterdam, Economics Dept./SCHOLAR, and Tinbergen Institute.
Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at http://www.tinbergen.nl.
THE CONNEXION BETWEEN OLD AND NEW APPROACHES TO FINANCIAL SATISFACTION

by

Bernard M.S. van Praag*

Forthcoming in:


Amsterdam, 5 may 2004

* The author is university professor at the University of Amsterdam. This paper is partly based on Chapter 2 of Happiness Quantified: A Satisfaction Calculus Approach (2004) by B.M.S. Van Praag and A. Ferrer-i-Carbonell.
ABSTRACT

THE CONNEXION BETWEEN OLD AND NEW APPROACHES TO FINANCIAL SATISFACTION

by

Bernard M.S. van Praag *

In this paper we compare the new satisfaction evaluation approach, developed in the nineties by Oswald, Clark, Blanchflower and others with the older income evaluation (IEQ) approach, developed by Van Praag and Kapteyn in the seventies of the previous century. We find that both approaches yield strikingly similar results with respect to financial satisfaction. The IEQ- approach yields additional insights, but it is not well applicable to other life domains than finance. It is argued that the usual Probit specification implies a specific cardinalization and, consequently, is less ordinal than usually thought. It is shown that the Probit-approach may be replaced by three other equivalent specifications that have some computational and intuitive advantages.

Keywords:financial satisfaction, income evaluation, Probit – models, cardinal utility.

* The author is university professor at the University of Amsterdam. This paper is partly based on Chapter 2 of Happiness Quantified : A Satisfaction Calculus Approach (2004) by B.M.S. Van Praag and A. Ferrer-i-Carbonell.
THE CONNEXION BETWEEN OLD AND NEW APPROACHES TO FINANCIAL SATISFACTION

by

Bernard M.S.van Praag*

1. Introduction

The subject 'happiness' of this timely book is one of the most pressing ones for the behavioural sciences in general and for economics in particular. As I am an economist by upbringing, I will restrict myself mostly to the economic viewpoint, although it is clearly unavoidable that there will be points of tangency with the other behavioural sciences, especially with psychology.

Economists agree on the fact that individuals strive for the greatest happiness. Let us assume two situations $x_1$ and $x_2$ and let us assume that the two situations generate happiness values $W_1$ and $W_2$, where $W_2 > W_1$, then the individual will choose $x_2$, if that situation is in his choice set. It follows that the function $W(x)$ describes a very basic aspect of human behaviour. Choice between scarce things is the core subject of economics.

It is therefore normal that economists developed this choice model, where human choice behaviour is described as maximizing a function $W$ on a relevant choice set.

One of the first economists who proposed it was Edgeworth in his book *Mathematical Psychics* (1881). He thought of $W$ as a cardinal concept. If $W(x_1)$=1 and $W(x_2)$=2, then the individual derives twice as much utility\(^1\) (or happiness) from $x_2$ as from $x_1$.

Pareto(1904) was the first to raise doubts about the practical possibility to observe and estimate the function $W$. Moreover, he showed that in the case of static consumer behaviour we do not need to know the function itself but only its contour lines, the so-called indifference curves, which are described by the equations:

---

* The author is university professor at the University of Amsterdam. This paper is partly based on Chapter 2 of Happiness Quantified: A Satisfaction Calculus Approach (2004) by B.M.S. Van Praag and A. Ferrer-i-Carbonell.
\[ W(x) = C \]  

(1)

where \( C \) stands for a constant. The same net of indifference curves is described by

\[ \varphi(W(x)) = \varphi(C) \]  

(2)

where \( \varphi(.) \) stands for an arbitrary monotonously increasing function.

The net of indifference curves defines an equivalence class of functions \( \varphi(W(x)) \), that have the same contour lines.

In the meantime there arose a growing aversion among economists towards 'psychologizing'. The term 'happiness' was abandoned for 'ophelimity' and later on for 'well–being', 'welfare' or for the still less emotionally loaded term 'utility' or 'satisfaction'.

The practical difficulty to estimate the function \( W \) was annoying for economists and it led to the gradual coming–up of the axiom that utility was immeasurable. Notice that an axiom is not proven but proposed and accepted. The high priest of this dogma was Lionel Robbins (1932) and it was supported by Hicks (1939), Samuelson (1947) and Houthakker (1950). We may say that after 1950 an economist was not taken seriously, if he did not profess his belief in this dogma.

As a consequence, in consumer theory the use of a cardinal utility function was either completely abandoned or it was used as just a handy instrument in order to describe the indifference curves. The latter function was called the ordinal utility function. To one net of indifference curves corresponds a whole equivalence class of ordinal utility functions, satisfying equation (2).

However, there were some problems left. The first point arose with decisions under uncertainty. If we accept the von Neumann- Morgenstern model, where we have a lottery \( L \) with outcomes \( x_1 \) and \( x_2 \) and corresponding probabilities \( p \) and \( (1-p) \), the individual is assumed to decide on expected utility

\[ E(W) = pW(x_1) + (1-p)W(x_2) \]  

(3)

\(^1\) We shall not differentiate between the terms utility, well-being, and happiness
Now it is clear that if we replace \( W \) by a non-linear monotonously increasing transform \( \varphi(W) \) and we have two lotteries \( L_1 \) and \( L_2 \), then it may be that \( L_1 \) is preferred to \( L_2 \) when we use \( W \) and that \( L_2 \) is preferred to \( L_1 \) when we use \( \varphi(W) \). It is obvious that in the setting of uncertainty we have to require that \( W \) is a cardinal utility concept. That is, \( \varphi(W) \) is defined up to a positive affine transformation, i.e., 
\[
\varphi(W) = \alpha W + \beta \quad \text{with} \quad \beta > 0.
\]
Actually, by observing lottery behaviour for various values of \( p \) we can estimate the values \( W \) up to a positive linear transform. We notice however that in the limiting situation where \( p \to 1 \) we are back in the situation of choice under certainty. Hence, if we assume that the certainty – situation is a special case of the more general uncertainty – situation we see that we cannot simultaneously maintain the validity of the unmeasurability axiom and VNM – theory.

Ragnar Frisch stated in 1959:

‘To me the idea that cardinal utility should be avoided in economics is completely sterile. It is derived from a very special and indeed narrow part of theory, viz., that of static equilibrium’.

In Van Praag (1968) I added to this quote (p.158): This author agrees completely with Frisch’s value judgment. The above – mentioned controversy seems to me the reason why there is found nowhere a really synthesizing analysis which brings under one denominator the theories of consumer behavior in a certain world and in an uncertain world. A similar story may be told with respect to decisions over time ( e.g. saving, investment).

It is indeed true that static consumer behaviour may be described by an ordinal utility concept. It follows that we cannot estimate and identify a cardinal utility function by observing static consumer behaviour, when prices vary. But this fact does not entail that cardinal utility would be unmeasurable \textit{per se} or a ridiculous concept. The statement only indicates that we should look for another measurement method.

There is still a rather basic observation to be made. When we have two alternative situations \( x_1 \), \( x_2 \) and \( x_3 \), in most cases individuals will not only be able to say that
they prefer \( x_2 \) to \( x_1 \) and \( x_3 \) to \( x_2 \), that is \( W(x_1) < W(x_2) < W(x_3) \), but they are also able to say whether the improvement going from \( x_1 \) to \( x_2 \) is more or less than the improvement associated with going from \( x_2 \) to \( x_3 \). Individuals are able to compare utility differences. But this is just what is necessary and sufficient for having a cardinal utility function (see Suppes and Winet (1954)).

One of the other methods by which we may investigate how individuals evaluate specific positions is by not observing their choice behavior, but much more simply by asking them how they evaluate those alternatives, either on a verbal scale ranging from 'very bad' to 'very good' or on a discrete (or ideally continuous) numerical finite scale, for instance from 0 to 10 or from 0 to 1.

In this framework we shall assume that satisfaction will always be measured on a finite interval scale, preferably [0,10]. That is, however the position will be described, either by a vector \( x \) or by a verbal description or by an image, vignet, etc., denoted by \( x \) as well, satisfaction will be described by a numerical function \( S(x) \), where worst positions are evaluated by 0 and best positions by 1.

This is the method, developed in the nineties, which employs so-called 'satisfaction questions'. Subjects are e.g. job satisfaction (Clark and Oswald, 1994), Health Satisfaction (Ferrer-i-Carbonell and Van Praag, 2002), Financial Satisfaction, or Satisfaction with 'life as a whole' (Van Praag et al., 2003).

In this paper we shall focus on Financial Satisfaction (FS). Let us assume that Financial Satisfaction is a function \( U(y; x) \) of household income \( y \) and other personal characteristics \( x \). Our question is then whether we can derive this function from the FS – question.

It turns out that there are four ways to derive a meaningful function \( U(.) \) from the FS – question. Their connexion, similarity and differences we will explain in Section 2.

As some readers will be aware of, in the seventies Van Praag (1971) and Van Praag and Kapteyn (1973) were considering the same problem. They formulated a different question module, the Income Evaluation Question (IEQ) and attempted to estimate a cardinal utility function of income \( U(y; x) \), which was called the Individual Welfare Function of Income (WFI).
In Section 3 we critically consider this second but earlier WFI -approach in more detail and we will see that the IEQ yields two utility function estimates, where one may be identified as a decision utility function and the other as an experienced utility function in the sense of Kahneman, Wakker and Sarin (1997). The functions, derived from the FS – approach may be identified as variants of the experienced utility function. It turns out that the experienced utility functions derived from either of the two roads are identical up to a positive linear transformation.

In Section 4 we present some empirical evidence.

In Section 5 we conclude that the IEQ stores more information than the FSQ at the expense of the fact that it requires more information from the respondent and is thus harder to answer. A second point is that it seems hard (but perhaps not impossible) to apply the IEQ -approach with respect to other domains than Financial Satisfaction. Moreover, we make some observations on the state of the art and the embedding of the happiness results in behavioural sciences in general and in economics in particular.

2. Four methods of analysis for the satisfaction question

The income satisfaction question we are using runs as follows:

```
How satisfied are you with your household income……………………………………..
(Please answer by using the following scale, in which 0 means totally unhappy and 10 means totally happy)
```

This question is posed in the German Socio–Economic Panel (GSOEP). A similar question is posed in the British Household Panel Survey (BHPS). The only difference between both modules is that the British survey questionnaire distinguishes between seven categories, while the German question has eleven response categories. The question is put in many other surveys as well. We notice that the response categories are explicitly described in terms of numerical grades, e.g., from 0 to 10. In other modules the response categories are verbal, ranging from 'very bad' to 'very good'. The verbal version is somewhat older and preferred by some as being better
understandable by respondents, although it is also thought that words may not carry
the same feeling for each respondent, diminishing the validity and the interpersonal
comparability of the question.

The numerical version gives less room for ambiguity. As most individuals are well-
acquainted with numerical evaluations since their school days, it may be surmised that
the evaluations may be interpreted as cardinal evaluations. It is obvious that, although
the satisfaction question requires a categorical answer, the underlying phenomenon is
continuous.

Let us assume that individuals $t$ in the population are ordered on the interval $[0,1]$ according to their satisfaction, such that in the interval $[0,t]$ is a fraction $t$ of the population. It is obvious that we may define the satisfaction $S(t)$ of individual $t$ as $S(t) = t$. In that case satisfaction is a purely relative phenomenon. For instance, if $t = 0.6$ it implies that 60% of the population is less satisfied than $t$ and hence $t$ 's satisfaction is 0.6. However, it is obvious, as we do not know the function $S$, that any other increasing function on $[0,1]$ may be just as credible. It is the objective of this paper to get more clarity on that. The response categories correspond with the intervals of a partition of the unit interval. We assume a model of the type

$$ S(t) = g[\alpha \ln(y) + \beta \ln(fs) + \gamma + \varepsilon] $$

(4)

where $g(.)$ is an increasing function, where $y$ stands for household income, $fs$ stands for family size and $\varepsilon$ is a random disturbance term with $E(\varepsilon) = 0$, which is uncorrelated with the structural part. In the sample $y$ and $fs$ are random variables as well. For convenience we define the variables $\ln(y)$ and $\ln(fs)$ as deviations from their means. This may evidently be rewritten as

$$ Z = \alpha \ln(y) + \beta \ln(fs) + \gamma + \varepsilon $$

(5)

where $Z = g^{-1}(S(t))$. As this paper does not concentrate on empirical results we take here only two explanatory variables. We notice that (5) for a constant $Z$- value describes an indifference curve in $(y;fs)$ – space. The label of the indifference curve is the value $Z$. 
A. Ordered Probit

Let us assume a random sample of size $N$, consisting of respondents $n$. The usual and first method to estimate this relation is by Ordered Probit, where it is assumed that the error term $\varepsilon$ is $\text{N}(0,1)$-distributed. The relation (5) implies that $Z \in (-\infty, \infty)$. The response categories for $Z$ correspond with intervals $(z_{i-1}, z_i]$. In terms of the original $t$ the response classes are $(t_{i-1}, t_i]$. For an individual who evaluates his financial satisfaction by $i_n$ it implies that $z_{i_n-1} < Z_n \leq z_{i_n}$. The log-likelihood of the sample is

$$L(z, \alpha, \beta) = \prod_{n=1}^{N} \left[ N(z_{i_n} - \alpha \ln(y_n) - \beta \ln(f_{n}) - \gamma) - N(z_{i_n-1} - \alpha \ln(y_n) - \beta \ln(f_{n}) - \gamma) \right]$$ (6)

where $N(.)$ stands for the standard – normal distribution function, is maximized with respect to $z_i, \alpha, \beta$. Generally, we are most interested in the parameters $\alpha, \beta$, which determine the trade-off ratio between $y$ and $f$. The parameters $z_i$ are called the nuisance parameters; they are mostly overlooked. Here, we are especially interested in those $z_i$, because they give insight in the cardinalization, which is implicitly applied by using the Probit model. It is easy to assess the $z_i$’s. We have for the conditional probability that individual $n$'s satisfaction will be found in the $i^{th}$ satisfaction interval

$$P(z_{i-1} < Z_n \leq z_i | y_n, f_{n}) = N(z_i - \alpha \ln(y_n) - \beta \ln(f_{n}) - \gamma) - N(z_{i-1} - \alpha \ln(y_n) - \beta \ln(f_{n}) - \gamma)$$ (7)

The marginal probability for an arbitrary individual $n$ to be found in the $i^{th}$ satisfaction interval is then (see Greene (1991)) the average of those individual conditional probabilities averaged over the individuals in the sample. We find

$$P(z_{i-1} < Z \leq z_i) = \text{plim}_N \frac{1}{N} \sum_{1}^{N} P(z_{i-1} < Z \leq z_i | y_n, f_{n})$$ (8)

$$= \text{plim}_N \frac{1}{N} \sum_{1}^{N} [N(z_i - \alpha \ln(y_n) - \beta \ln(f_{n}) - \gamma) - N(z_{i-1} - \alpha \ln(y_n) - \beta \ln(f_{n}) - \gamma)] = p_{i_i}$$

$(i = 1, \ldots, k-1)$
where $p_i$ stands for the sample fraction of respondents who answered the response category $i$.

From these equations we find the $z_i$'s. It follows that the latent variable $Z$ is not ordinal but cardinal. Although its value cannot be observed exactly, we know that a response $i$ corresponds to a value $Z \in (z_{i-1}, z_i]$. The remaining inexactness is only due to the discrete observation mode. If we would be able to make an infinitely-fine differentiation of response categories, the cardinalization of $Z$ would be exact. That is, *the Probit-model defines a cardinalization of $Z$.*

\[ B^2. \text{Probit OLS. (POLs)} \]

Although Ordered Probit is now included in all relevant software packages, it is still less easy and significantly less flexible than good old OLS. Equation (10) suggests that OP might be substituted by an OLS - procedure. We call that procedure Probit OLS (POLs).

In the POLS-approach we start from the other end so to speak. We assume that the labels $Z$ of the indifference curves within a population are distributed according to a continuous distribution function $G(z)$, that is, there is no indifference curve with a discrete mass of observations on it. Then $G(z)$ is the fraction of the population that is situated on or at a lower satisfaction level than the one associated with the indifference curve $z$. We repeat that in the ordinal approach there is no cardinal meaning attached to the values $z$. This means that we can replace the values $z$ by $\tilde{z} = \varphi(z; \zeta)$, where the function $\varphi$ is monotonically increasing to preserve the order

and where $\zeta$ is a set of $\varphi$-specific parameters. Notice that if $z = z(x; \beta)$, then $\tilde{z} = \varphi(z(x; \beta); \zeta)$. This implies that both representations $z$ and $\tilde{z}$ describe the same net of indifference curves. Consider now the frequency distribution of labels in the sample after application of the monotonic transformation. The distribution function of the distribution of $\tilde{z}$ is $H(\tilde{z}) = G(\varphi^{-1}(\tilde{z}))$. This shows that the distribution function of the label distribution depends on the specific labeling system. Inversely, it

---

\[ ^2 \text{This section is partly revised without changing the gist of the original text. This description of the POLS-method is partly derived from the description in Van Praag, Ferrer-i-Carbonell (2006).} \]
follows that the label distribution may be any continuous distribution on the real axis, depending on the appropriate choice of the re-labeling function $\varphi(.)$. It follows that there is a specific labeling system, for which the distribution of $\tilde{Z}$ will be standard normal, i.e., $H(\tilde{z}) = N(\tilde{z}; 0,1)$. We call this labeling system the normal labeling system. This amounts to another cardinalization than that implicit in Ordered Probit. We drop the tilde from now on.

Let us now assume that we observe satisfaction in terms of a few discrete response categories, for example ranging from 'very dissatisfied' to 'very satisfied'. The range of labels is partitioned in response categories that represent $k$ adjacent intervals $(z_{i-1}, z_i]$, such that a response $I=i$ ($i=1,...,k$) implies that the latent variable $Z \in (z_{i-1}, z_i]$. We define $z_0 = -\infty, z_k = \infty$. The categorical frequencies (i.e. the frequency of responses found in each $k$ category) are $p_1,..., p_k$. Now, if we start off from the normal labeling system the variable $Z$ is $N(0,1)$-distributed in the population. Moreover, we assume a model where $Z$ may be decomposed into a structural part, say $f(X)$ and a residual part $\varepsilon$, such that the two components are mutually independent. A rather deep theorem in probability theory, first proved by H. Cramèr in 1937 (see Feller, 1966, Ch. XV, 8, also Rao, 1973, p.525), states that if $Z$ is normally distributed and if it is the sum of two mutually independent random variables, say $f(X)$ and a residual part $\varepsilon$, then those two variables have to be normally distributed as well. It implies that the structural part $f(X)$ will be normally distributed as well. This does not imply that all $X$-variables separately have to be normal, for they are not assumed to be mutually independent. But it does imply that $f(X)$ cannot be restricted to a proper subset of the real axis $(-\infty, \infty)$ only. This excludes specifications where, for instance, $f(X) > 0$ by definition.

Given the normal distribution of the labels $Z$ over the population we may estimate the nuisance parameters $z_i$'s in a simple manner by solving the equations
\[ p_i = N(z_i) - N(z_{i-1}) \quad (i = 1, \ldots, k - 1) \quad (9) \]

(see for similar thoughts also Terza (1987), Stewart (1983), and Ronning and Kukuk (1996)).

These are \((k-1)\) equations in \((k-1)\) unknowns \(z_1, \ldots, z_{k-1}\). Notice that this result does not depend on the \(x\)-values, not brought into play yet, but only on the distribution of the response categories, that is, the unconditional distribution of \(Z\).

We can calculate the conditional expectation \(\bar{z}_i\) \(i,\) given \(Z \in (z_{i-1}, z_i]\) according to a well-known formula for the normal distribution as

\[ \bar{z}_i = E(Z \mid z_{i-1} < Z \leq z_i) = \frac{n(z_{i-1}) - n(z_i)}{N(z_i) - N(z_{i-1})} \quad (10) \]

Instead of taking a response category \(i_n\) as our observation to be explained, we take \(\bar{z}_i\) as the variable to be explained. We notice that \(\bar{z}_i\) can assume only \(k\) discrete values, where \(k\) is the number of response categories. We observe that this expression (10) does not depend on the individual characteristics \(y, f_s\). This is analogous to the usual regression situation where the 'left-hand' variable to be explained is directly observed without 'correcting' for additional information about the respondent, as revealed by explanatory variables.

It follows that we look at the regression model

\[ \bar{z}_i = \alpha \ln(y_n) + \beta \ln(f_{s_n}) + \gamma + \epsilon_n \quad (11) \]

We notice that the error term is a discrete random variable. However, if the number of observations is large, we may apply all large-sample results and deal with this OLS-equation as usual.

More precisely, we may write the model as
\[ Z = \bar{Z} + \eta \]  

(12)

The true latent observation is written as the sum of its conditional expectation plus a rounding-off error \( \eta \), caused by the fact that we can only observe the interval in which the true \( Z \) is situated. We may rewrite this regression equation as

\[ \bar{Z}_i = \alpha \ln(y) + \beta \ln(\bar{fs}) + \varepsilon + \eta \quad (i=1, \ldots, k) \]  

(13)

We summarize the POLS-method as follows:

1. Calculate the response fractions \( p_i \)
2. Calculate the nuisance parameters \( z_i \) using equation (9)
3. Calculate the values \( \bar{Z}_i \) according to (10)
4. Perform the regression (13)

We may raise the question whether we can also just as well take the untransformed response variable \( i (I = 0, \ldots, 10) \) itself as our dependent variable to be explained. This would yield the regression equation

\[ i = \alpha \ln(y) + \beta \ln(\bar{fs}) + \varepsilon \]  

(14)

This is a generalisation of the Linear Probability model (see Greene, 1991, p.813). Indeed we might do this but the results are statistically and intuitively not very attractive, unless we have only two response categories. There are two reasons for its unattractiveness. First, the range of the variable to be explained is finite instead of the real axis, which the model specification logically would require. Second, contrary to the practice in POLS, the values of the variable to be explained are equi-distanced by definition. In contrast, in POLS they are defined by the overall sample distribution. This explains as well, why the Linear Probability - model works for a phenomenon, which is two-valued, but not for multi-valued phenomena. In the two-valued case it is equivalent to POLS, except for an affine linear transformation.
If we drop our conventional prejudice towards cardinalism, we cannot deny that respondents who answer a satisfaction question by giving a numerical response are attempting to make a cardinal evaluation in terms of a finite interval scale. It stands to reason that responses are not very accurate, but the position that the respondent would have no intention to evaluate and, consequently, that his answers do not have any information value, may be safely discarded.

Now we look at a third method, which makes use of the cardinal information in the Financial Satisfaction Question as well. It is this cardinal information, which is neglected by Ordered Probit. If somebody is evaluating his satisfaction level by a 'seven', it does not imply that his satisfaction is exactly equal to 7. For instance, the exact evaluation might be 6.75 or 7.25, but due to the necessary discreteness of the responses we have to round it off at 7. However, it would be very improbable when the exact evaluation would be 7.75, for in that case we would round off to 8. More precisely, we assume that if somebody responds 7 his true evaluation will be in the interval (6.5, 7.5]. A similar reasoning holds for all other response values. For the extremes we use an obvious modification. The observed value 0 corresponds to the interval [0, 0.5] and the value 10 to (9.5, 10]. If we normalise the scale from [0,10] to the [0,1] - interval, the intervals will be [0,0.05],..., (0.95, 1).

Let us now assume that the satisfaction $S$ may be explained to a certain extent by a vector of explanatory variables $x$, including log-income. More precisely, we assume

$$ S = N(\beta'x + \beta_0;0,1) $$

(15)

where $N(.)$ stands for the (standard-) normal distribution function. We stress that (15) is a non-stochastic specification. For estimation purposes we add a random term and assume

$$ S = N(\beta'x + \beta_0 + \varepsilon ;0,1) $$

(16)
We see that satisfaction is determined by a structural part and a random disturbance \( \varepsilon \). We assume the random disturbance \( \varepsilon \) to be normally distributed with expectation equal to zero. Its variance \( \sigma^2 \) has to be estimated. As usual, we assume that \( \text{cov}(x, \varepsilon) = 0 \). Notice, that this model, and especially the specification of eqs. (15) and (16), is an assumption. If another model would fit the data better, we have to replace it. However, let us assume it holds.

In that case the chance on response 7 is

\[
P[0.65 < U \leq 0.75] = P[N^{-1}(0.65) < \beta'x + \beta_0 + \varepsilon \leq N^{-1}(0.75)]
\]

\[
= N(u_{0.75} - \beta'x - \beta_0) - N(u_{0.65} - \beta'x - \beta_0)
\]

(17)

Comparison with eq.(7) reveals that the likelihood is equal to the Probit-likelihood except that the unknown \( \mu_i \)'s are replaced by known normal quantiles \( u_i \). The \( \beta \)'s are estimated by maximizing the log-likelihood.

It follows that it is possible to estimate a cardinal satisfaction function from the same data by using the additional cardinal information. It is an empirical matter, which model is chosen.

This Cardinal Probit (CP) -approach is a special case of what is called in the literature sometimes the Group-wise or Interval Regression Method, where information on the regressand is only available group-wise. This is frequently the case in public statistics, eg., with respect to household income which is only known per income bracket.

**D. Cardinal OLS (COLS)**

The reader will not be surprised that the trick of eq.(10), which we used in order to define the POLS – method, can be used in the cardinal setting as well. We define

\[
\tilde{Z}_i = E(Z \mid u_{i-1} < Z \leq u_i) = \frac{n(u_{i-1}) - n(u_i)}{N(u_i) - N(u_{i-1})}
\]

(18)

and we formulate the regression equation
\[ \bar{Z}_i = \alpha \ln(y) + \beta \ln(fs) + \epsilon + \eta_i \quad (i=1, ..., k) \] (19)

In this section we listed four possible methods to estimate an explanatory model for satisfaction. The difference between methods A and C is that A does not employ the cardinal information in the satisfaction question, while C does employ that additional information. Methods B and D may be viewed as derivatives of A and C, respectively. The essential difference is between the acceptance or non-acceptance of the cardinal information. The two variations B and D are of much practical importance, as they make it possible to replace the non-linear Probit method by more easily applicable OLS.

The question is now how the different estimates are related. We leave this question for section 4 and go now to look after a much older competitor.

3. The income evaluation question

It is sometimes forgotten that the present wave of happiness research was preceded in the seventies by another attempt, which had certainly points in common with the present literature. This cluster of research is now frequently called the Leyden School after the Dutch university, where this research started. Van Praag, Kapteyn and Hagenaars were the main contributors. This line of research was started by Van Praag (1968, 1971) and it may be seen as a forerunner to present satisfaction question research.

In the spirit of the economic literature of that time it was assumed that satisfaction with income was synonymous with welfare or well-being. Although also economists (including this author) paid lip service to the idea that income was only dimension of life, this feature of reality was ignored in the practice of developing theory and applied research, where income was seen as the only determinant of welfare. Now we would say that the Leyden School was focusing on financial satisfaction. In this sense the subject of Leyden was narrower than that of present happiness research where various life domains, like job (Clark and Oswald, 1994), health (Ferrer-i-Carbonell and Van Praag, 2002) are studied as well. However, we should also realize that in those days so-called 'subjective' satisfaction questions were not put in surveys to
which economists had access. There were some 'soft' surveys organized by sociologists or psychologists, where such questions could be found, but those surveys did not contain reliable information about income and other 'economic' variables. Sociologists and psychologists were not interested in those mundane regions of life and left it to the 'dismal science' to bother about the effect of income.

The Leyden results are empirically based on the so-called Income Evaluation Question (IEQ). The IEQ has been posed in various countries. Here we are especially interested in comparing the outcomes with the previous results, derived from the Income Satisfaction Questions. Fortunately, the IEQ has been posed in the GSOEP-data set in the waves 1992 and 1997. This gives us the opportunity for a direct comparison between the results based on the Financial Satisfaction question with those derived from the IEQ. We utilise the 1997 wave.

Our first question here is whether the IEQ provides at least the same information as the Financial Satisfaction Question. Our second question is whether the results derived from the IEQ are comparable or nearly the same as the results, derived from the FS-question. Third, we are interested in the question whether the IEQ provides more information than the Financial Satisfaction Question.

The IEQ runs as follows:

**The Income Evaluation Question (IEQ) (mid-interval version).**

<table>
<thead>
<tr>
<th>Level Description</th>
<th>DM Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>a very low income if it would equal</td>
<td></td>
</tr>
<tr>
<td>a low income if it would equal</td>
<td></td>
</tr>
<tr>
<td>a still insufficient income if it</td>
<td></td>
</tr>
<tr>
<td>a just sufficient income if it</td>
<td></td>
</tr>
<tr>
<td>a good income if it would equal</td>
<td></td>
</tr>
<tr>
<td>a very good income if it would equal</td>
<td></td>
</tr>
</tbody>
</table>

There are several wordings of this question around. First, the number of levels has varied between five and nine. When it was first posed in a Belgian survey (Van Praag (1971)), nine verbally described levels have been used. In Russian surveys (see
Ferrer-i-Carbonell and Van Praag (2001) five levels have been used. Second, in the earliest versions (1971) the question was formulated as:

**The Income evaluation Question (interval - version).**

<table>
<thead>
<tr>
<th>An income below $???</th>
<th>as a very bad income</th>
</tr>
</thead>
<tbody>
<tr>
<td>An income between $??? and $???</td>
<td>as a bad income</td>
</tr>
<tr>
<td>An income between $??? and $???</td>
<td>as an insufficient income</td>
</tr>
<tr>
<td>An income between $??? and $???</td>
<td>as a sufficient income</td>
</tr>
<tr>
<td>An income between $??? and $???</td>
<td>as a good income</td>
</tr>
<tr>
<td>An income above $???</td>
<td>as a very good income</td>
</tr>
</tbody>
</table>

When introducing this type of question, which requires more from a respondent than the usual financial satisfaction question, survey agencies predicted that the response ratio would be very bad and that, if there would be any response, the respondents would not take this question seriously. It appeared in practice that those questions have a lower response than usual questions but not dramatically so. It may also be that the response is incomplete, but the question may still be used if at least three levels are filled in. Moreover, the amounts should be ordered in the sense that a good income requires a higher amount than a bad income. Finally, the response is considered to be unrealistic if a very bad income is much higher than the respondent’s current income or a very good income is much less than current income. Such cases represent a small percentage of the response and they are usually excluded from further analysis.

The essential difference between the FS- question and the IEQ is the inversion of stimulus and response. In the FS-question own current income, say $y_c$, is the stimulus and the individual's evaluation on a finite interval scale is the response. In the IEQ the stimuli are evaluations, expressed in terms of verbal labels like 'bad' and 'good'. The responses are income levels $y_{bad}$ and $y_{good}$. As different individuals have a different idea on what is a 'good' or a 'bad' income, it is obvious that we do not get one financial satisfaction function, but that each responding individual will have his own FS - function. Therefore, Van Praag (1971) used the term individual welfare function of income (WFI).
We now analyse the results of the IEQ. Let us denote the answers by of individual $n$ by $c_{i_1}, \ldots, c_6$. For analysis we have two possibilities. The first one is an ordinal analysis, where we consider the separate answers and look for regression equations

$$\ln(c_i) = \alpha_i \ln(y_i) + \beta_i \ln(f_s) + \gamma_i + \varepsilon_i \quad (i=1,\ldots,6) \quad (20)$$

The question is then what these coefficients are and whether these coefficients are equal over the six equations. We leave the empirical results for the next section.

Now we look at the cardinal concept of the Individual Welfare Function of Income (WFI). In Van Praag (1968) it was argued that individual welfare (read financial satisfaction in present days' terminology) was measurable as a cardinal concept between 0 and 1. In 1968 this was evidently an almost heretical idea, not in favour with mainstream economics (see e.g. Seidl (1994) for a fierce but belated critique).

The approximate relationship was argued to be a lognormal distribution function with parameters $\mu$ and $\sigma$. We notice that the specification (16) is also a log-normal specification if one of the dimensions of the vector $x$ is $\ln(y_i)$. In later years Van Praag (1971) and Van Praag and Kapteyn (1973) estimated the $\mu$ and $\sigma$ per individual on the basis of the response on the IEQ. They assumed for the 'mid-interval' version of the IEQ that the answers $c_{i_1}, \ldots, c_6$ correspond with satisfaction levels $1/12$, $(2i-1)/12$ and $11/12$ respectively; this was called the Equal Quantile Assumption (Van Praag (1991, 1994)) provided empirical evidence for this assumption. Moreover, it was assumed that satisfaction $U(c;\mu, \sigma) = \Lambda(c; \mu, \sigma)$ where $\Lambda(c; \mu, \sigma) = N(\ln(c); \mu, \sigma)$. The function $\Lambda(.)$ stands for the log-normal distribution function. Estimation of $\mu$ and $\sigma$ is possible per individual. We have six or, more generally, $k$ observations per individual and we assume that

$$N\left(\frac{\ln(c_i) - \mu}{\sigma}\right) = \frac{(2i-1)}{12} \quad (21)$$

We note that the $c$-value is comparable to $z$ in the Cardinal Probit situation of six observations per individual by COLS. The only difference is that the $c$'s are equated to interval *medians* instead of interval *means*. We estimate the parameters $\mu$ and $\sigma$ for each individual observation $n$ by

$$N\left(\frac{\ln(c_i) - \mu}{\sigma}\right) = \frac{(2i-1)}{12} \quad (21)$$
\[ \hat{\mu}_n = \frac{1}{6} \sum_{i=1}^{6} \ln(c_{in}) \quad \text{and} \quad \hat{\sigma}_n^2 = \frac{1}{5} \sum_{i=1}^{5} \ln(c_{in} - \hat{\mu}_n)^2 \] (22)

Then the estimated \( \hat{\mu}_n \) is explained over the sample of N observations by the equation

\[ \hat{\mu}_n = \alpha \ln(y_{c,n}) + \beta \ln(fs_n) + \gamma \] (23)

where \( y_{c,n} \) stands for the current income of individual \( n \). Later on we shall consider those regression results. Here we already notice that both coefficients are always estimated as significantly positive. The income effect \( \alpha \) equals roughly 0.6 and the family size coefficient \( \beta \) equals 0.10.

Up to now it has proved difficult to explain the \( \sigma \) - parameter, which was called by Van Praag (1968,1971) the welfare sensitivity, to an acceptable extent by means of individual explanatory variables. It seems to vary over individuals in a random manner. Like in many other studies, also here we assume \( \sigma \) to be constant over individuals in the same population. We set it equal to the average over individuals. Hagenaars (1986) found from international comparisons that the parameter \( \sigma \) appears to be related with the log – standard deviation of the population’s income distribution. Her result suggests that welfare sensitivity is higher in more unequal societies. In the present survey (GSOEP 97) we found an average value of \( \sigma = 0.453 \).

We can now find the evaluation of any income \( y \) by someone with individual parameters \((\mu(yc), \sigma)\). It equals

\[ N(\ln(y); \mu(yc), \sigma) = N\left(\frac{(\ln(y) - \mu(yc))}{\sigma}\right) = N\left(\frac{(\ln(y) - \alpha \ln(yc) - \beta \ln(fs) - \gamma)}{\sigma}\right) \] (24)

or using its ordinal equivalent on the \((-\infty, \infty)\)-axis

\[ \frac{(\ln(y) - \mu)}{\sigma} = \frac{(\ln(y) - \alpha \ln(yc) - \beta \ln(fs) - \gamma)}{\sigma} \] (25)
We notice that the IEQ effectively introduces two concepts of an Individual Welfare Function. The first concept is generated by keeping $\mu$ constant. It gives a schedule of how individuals evaluate varying (fictitious) income levels from the perspective of their own income, which is kept unchanged at the present level. We call it the virtual or short-term welfare function. It can be estimated for a specific individual by posing the IEQ to that individual.

The second concept is the welfare function according to which individuals with different incomes evaluate their own income in reality. It is an inter-individual concept. We call it the true or long-term welfare function. This function has to be derived by using a sample of different individuals. From (23) it follows that the true welfare function is

$$N\left( \frac{(\ln(y_c) - \mu(y_c))}{\sigma} \right) = N\left( \frac{(\ln(y_c) - \alpha \ln(y_c) - \beta \ln(fs) - \gamma)}{\sigma} \right)$$

$$= N(\ln(y_c); \frac{\beta \ln(fs) + \gamma}{1-\alpha}, \frac{\sigma}{1-\alpha})$$

(26)

Hence, it is also log-normal with parameters

$$\left( \frac{\beta \ln(fs) + \gamma}{1-\alpha}, \frac{\sigma}{1-\alpha} \right)$$
Fig. 1a. The virtual and true welfare functions.

In fig. 1.a we sketch both functions. We see that the true welfare function has a much weaker slope than the virtual. It implies that income changes are *ex ante* heavier perceived than when they are experienced in reality. Actually, the two concepts correspond with the two concepts of the *decision* and the *experienced* utility function, distinguished by Kahneman, Wakker and Sarin (1997). The virtual welfare function describes the way in which a specific individual evaluates different income levels, irrespective of whether it is his real income or a (remote) prospect. It is the perceived *ex ante* relationship between income and welfare on which the individual bases his decisions. The true welfare function describes how individuals, who *experience* those incomes themselves, evaluate incomes in reality.

The welfare function maps incomes on the evaluation range \([0,1]\). A second (and easier) way to consider the welfare function is to map the range to the real axis \((-\infty, +\infty)\) and to consider the function \(u(y)\)

\[
   u(y) = \frac{(1 - \alpha) \ln(y) - \beta \ln(f) - \gamma}{\sigma}
\]

The two representations are ordinally equivalent. We call the latter the linear transform. The linear transforms of the virtual and the true welfare functions are sketched in Fig.1.b. The short – term version corresponds to \(\alpha = 0\). It follows again that the short- term function is much steeper than the long -term function.
The difference between the short- and long term concepts is best explained by the following simple thought experiment. Let us assume that somebody with an initial income $y$ gets an income increase of $\Delta y$, yielding a new income $y + \Delta y$. Initially the increase will be evaluated by his short-term welfare function yielding an increase from point $A$ to point $B$. After a while income norms will adapt to the new situation and this will be reflected in the parameter $\mu$ that will increase by $\alpha \Delta y$. Hence, after a first euphoria there will be some disappointment, as the evaluation falls from point $B$ to point $C$.

This is the so-called preference drift effect, which was introduced and estimated by Van Praag (1971). It is only not there when $\alpha = 0$. If $\alpha = 1$, in the long term an income increase will not yield any increase in satisfaction.

We notice that the IEQ effectively introduces two concepts of an Individual Welfare Function. The virtual welfare function describes the way in which a specific individual evaluates different income levels. It is the perceived ex ante relationship between income and welfare on which the individual bases his decisions. The true welfare function describes how individuals, who experience those incomes themselves, evaluate incomes in reality.

The most interesting point is that different individuals have a different idea of what presents a ‘good’ income, a ‘sufficient’ income, etc.. It depends on their own net
household income and their household circumstances, in this case characterised by their household size. It shows that evaluations are relative. In case that \( \alpha = 0 \), the evaluations would be wholly absolute, that is, independent of current income. In case that \( \alpha = 1 \) the evaluation is completely relative. We will see from the table that we are somewhere in between, as \( \alpha \) turns out to be about 0.5 or 0.6. The phenomenon that evaluations are drifting along with rising income has been termed preference drift (see Van Praag (1971)). It is measured by \( \alpha \). This is similar to an effect, independently discovered by the psychologists Brickman and Campbell (1971). They called it the hedonic treadmill effect. These authors and also Easterlin (1974, 1995, 2001) tend to the hypothesis that adaptation would be complete, i.e. \( \alpha = 1 \). We were unable to establish this result empirically.

Obviously this is a puzzling effect. The evaluation of a specific income in combination with a specific household to support should, ideally and according to traditional economic models, be independent of the situation of the evaluator. However, we see that in practice it does depend on the income of the evaluator. It shows most clearly that the notion of 'a good income' is partly relative and psychologically determined. This holds as well for the situation of 'poverty' (see Goedhart et al. 1977).

4. Empirical evidence

Let us now go to the empirical analysis. We consider the GSOEP – data set and more precisely the 1997 wave\(^3\), where we restrict ourselves to the subset of West- German workers. The data set is so interesting because it contains the Financial Satisfaction Question and the IEQ simultaneously. We shall now compare the empirical outcomes of the approaches described before.

In Table 1 we tabulate the response fractions for the eleven categories.

<table>
<thead>
<tr>
<th>Satisfaction categories</th>
<th>Relative frequencies in %</th>
</tr>
</thead>
</table>

\(^3\) We use here a preliminary unauthorised release of the 1997-wave, which slightly differs from the final authorised version. In the final version there are found a few observations in category 0 as well. See also Plug et al. (1997).

\(^4\) See footnote 1.
In Table 2 we present the Ordered Probit - estimates of the unknown parameter values.
In Table 2 we present *three* estimates of the financial satisfaction equation to be estimated. The first version is the one, described by equation (5). We have two explanatory variables, viz. household income and the number of children *plus one* \(^5\).

In the second version we distinguish between children under 17 living at home and other persons living in the household. The latter class will be called ‘adults’. In the third specification we add a third variable ‘family structure’, which equals *zero* if the respondent lives alone, *one* if the household has two working adults, and *two* if there are two adults in the household of which only one has paid work. This ordering reflects the idea that household chores are pretty fixed and that they are easier to bear by a family with one person working outside than by a family where the two adults are working in a paid job. The single person or the incomplete family bears the heaviest burden in this respect.

The estimation results, when we use POLS, are given in Table 3.

---

\(^5\) We add one such that we do not get a non-defined logarithm of zero if the number of children equals zero.
Table 3. POLS results for the income satisfaction equations, GSOEP

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Household income)</td>
<td>0.678</td>
<td>19.829</td>
<td>0.681</td>
<td>19.581</td>
<td>0.681</td>
<td>19.584</td>
</tr>
<tr>
<td>Ln(Family Size)</td>
<td>-0.206</td>
<td>-6.624</td>
<td>-0.205</td>
<td>-6.122</td>
<td>-0.227</td>
<td>-6.597</td>
</tr>
<tr>
<td>Ln(Adults)</td>
<td>-0.118</td>
<td>-3.645</td>
<td>-0.152</td>
<td>-4.407</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(Children + 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.065</td>
<td>2.842</td>
</tr>
<tr>
<td>N</td>
<td>4964</td>
<td></td>
<td>4964</td>
<td></td>
<td>4964</td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.073</td>
<td></td>
<td>0.073</td>
<td></td>
<td>0.075</td>
<td></td>
</tr>
</tbody>
</table>

We see that the corresponding t-values are almost the same. The coefficients look multiples of each other except for the constant.
The cardinal CP- or interval- regression approach yields the following estimates and again we see that the ratios of coefficients are almost the same while the t-ratios hardly differ.

Table 4. The Financial Satisfaction Question estimated by Cardinal Probit.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Household income)</td>
<td>0.342</td>
<td>19.923</td>
<td>0.343</td>
<td>19.652</td>
<td>0.343</td>
<td>19.657</td>
</tr>
<tr>
<td>Ln(Family Size)</td>
<td>-0.102</td>
<td>-6.568</td>
<td>-0.101</td>
<td>-6.019</td>
<td>-0.112</td>
<td>-6.480</td>
</tr>
<tr>
<td>Ln(Adults)</td>
<td>-0.060</td>
<td>-3.697</td>
<td>-0.076</td>
<td>-4.430</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(Children + 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.032</td>
<td>2.765</td>
</tr>
<tr>
<td>Sigma</td>
<td>0.466</td>
<td>94.190</td>
<td>0.466</td>
<td>94.190</td>
<td>0.465</td>
<td>94.184</td>
</tr>
<tr>
<td>N</td>
<td>4964</td>
<td></td>
<td>4964</td>
<td></td>
<td>4964</td>
<td></td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-9500</td>
<td></td>
<td>-9500</td>
<td></td>
<td>-9496</td>
<td></td>
</tr>
<tr>
<td>Pseudo R2</td>
<td>0.0198</td>
<td></td>
<td>0.0199</td>
<td></td>
<td>0.0202</td>
<td></td>
</tr>
</tbody>
</table>
Finally we use the COLS – approach with yields table (2.11) and again we see that the ratios of coefficients and the t-ratios are almost the same.

Table 5. The Financial Satisfaction Question estimated by COLS

<table>
<thead>
<tr>
<th>Variable</th>
<th>effects</th>
<th>t-ratio</th>
<th>effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Household income)</td>
<td>0.354</td>
<td>19.388</td>
<td>0.355</td>
<td>19.123</td>
<td>0.355</td>
<td>19.126</td>
</tr>
<tr>
<td>Ln(Family Size)</td>
<td>-0.108</td>
<td>-6.515</td>
<td>-0.107</td>
<td>-5.966</td>
<td>-0.118</td>
<td>-6.427</td>
</tr>
<tr>
<td>Ln(Adults)</td>
<td>-0.108</td>
<td>-6.515</td>
<td>-0.107</td>
<td>-5.966</td>
<td>-0.118</td>
<td>-6.427</td>
</tr>
<tr>
<td>Ln(Children + 1)</td>
<td>-0.063</td>
<td>-3.669</td>
<td>-0.081</td>
<td>-4.033</td>
<td>0.034</td>
<td>2.766</td>
</tr>
<tr>
<td>Family Structure</td>
<td>0.034</td>
<td>2.766</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>4964</td>
<td>4964</td>
<td>4964</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.070</td>
<td>0.070</td>
<td>0.072</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The four methods used have the same objective, that is, the estimation of the equation

\[ Z = \alpha \ln(y) + \beta \ln(fs) + \gamma + \varepsilon \]  \hspace{1cm} (28)

where \( Z \) stands for a satisfaction index. The equation may be used for the derivation of family equivalence scales\(^6\). If \( \ln(fs) \) increases to \( \ln(fs) + \Delta fs \), the question arises by how much the individual has to be compensated in his income \( \ln(y) \). We find

\[ \Delta \ln(y) = -\frac{\beta}{\alpha} \Delta \ln(fs) \]  \hspace{1cm} (29)

It follows that the indifference curves between income and family size are described by

\[ y = y_0 (fs / fs_0)^{-\frac{\beta}{\alpha}} \]  \hspace{1cm} (30)

where \( y_0 \) and \( fs_0 \) stand for the reference income and reference family size, respectively. Now it is interesting to see whether the ratio \( \frac{\beta}{\alpha} \) is the same, irrespective of the four methods used. We give the different values in Table 6.

---

Table 6. Equivalence scale parameter calculated via different methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>OP</th>
<th>POLS</th>
<th>CP</th>
<th>COLS</th>
<th>IEQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta/\alpha$</td>
<td>0.32</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
<td>0.26</td>
</tr>
</tbody>
</table>

We see that the values of the ratio, estimated via four different methods, are virtually identical. Actually, this is less surprising than it looks like, if we realize that this equivalence scale describes an indifference curve, which is defined by the Financial Satisfaction - question. Everybody who evaluates his income by the same number is on the same indifference curve. The four methods yield different monotonic transforms of satisfaction, but their ordinal information is the same. The fifth column, derived from the IEQ, will be considered in a moment.

Let us now consider what is the relation between POLS and COLS - estimates. We denote as before the ordinal variable, belonging to a specific response category, by $\ln(\tilde{z})$ and the corresponding cardinal value by $\ln(\tilde{x})$. We assume $\ln(\tilde{x}) = f(\ln(\tilde{z}))$. As the categories are ordered, we may assume that $f(.)$ is a monotonically increasing function. Let us assume for a moment that both variables would be exactly measured on a continuous scale instead of on a discrete scale, then the marginal distributions of both variables would be normally distributed with parameters $(\mu, \sigma)$ and $(\bar{\mu}, \bar{\sigma})$, respectively.

We may express a fraction of respondents to a specific category either with respect to $\tilde{z}$ or with respect to $\tilde{x}$. We have

$$N(\ln(\tilde{z}); \bar{\mu}, \bar{\sigma}) \equiv N(f(\ln(\tilde{z})); \bar{\mu}, \bar{\sigma}) \tag{31}$$

which implies

$$\frac{\ln(\tilde{z}) - \bar{\mu}}{\bar{\sigma}} \equiv \frac{f(\ln(\tilde{z})) - \bar{\mu}}{\bar{\sigma}} \tag{32}$$

It follows that the function $f(.)$ is a linear affine transformation. We have
\[
\ln(\tilde{z}) = \frac{\sigma}{\bar{\sigma}} \ln(\tilde{z}) + D
\]

(34)

where \( D \) is a constant which can be easily calculated. Indeed if we apply this regression (on \( k \) observations) we find for the German data the regression result

\[
\ln(\tilde{Z}) = 0.5359 \ln(\tilde{Z}) + 0.1965.
\]

with an \( R^2 \) of 0.99.

It follows that, if \( \ln(\tilde{z}) \) is a linear combination of variables \( x \), then the same will hold for \( \ln(\tilde{z}) \), where the trade-off ratios will be the same.

It follows that OP, POLS, CP and COLS are for practical purposes equivalent for the computations of trade-off ratios. The C-versions employ the cardinal part of the information as well.

The implicit cardinalisation on which Probit and POLS are based will be called from now on the frequentist cardinalisation because it is based on the frequency distribution of satisfaction levels. The cardinalisation on which CP and COLS are based will be called the satisfaction cardinalisation from now on. We notice that one is a linear transformation of the other.

It is evident that we may also derive family equivalence scales from the IEQ. The estimates of equation (23) are presented in Table 7.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
<th>Effects</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>3.611</td>
<td>54.308</td>
<td>3.572</td>
<td>52.302</td>
<td>3.574</td>
<td>52.309</td>
</tr>
<tr>
<td>Ln(Household income)</td>
<td>0.527</td>
<td>61.964</td>
<td>0.533</td>
<td>60.667</td>
<td>0.533</td>
<td>60.644</td>
</tr>
<tr>
<td>Ln(Family Size)</td>
<td>0.121</td>
<td>14.819</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(Adults)</td>
<td>0.090</td>
<td>8.093</td>
<td>0.089</td>
<td>7.958</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(Children + 1)</td>
<td>0.096</td>
<td>11.976</td>
<td>0.083</td>
<td>5.555</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Structure</td>
<td></td>
<td></td>
<td>0.011</td>
<td>0.963</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.453</td>
<td>0.453</td>
<td>0.453</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( N )</td>
<td>3962</td>
<td>3962</td>
<td>3962</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.631</td>
<td>0.632</td>
<td>0.632</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is obvious that we may derive for the individual welfare function household equivalence scales by requiring that households with different family sizes \( f_{s0} \) and \( f_{s1} \)
enjoy an equal welfare level according to the true welfare function. This implies that the indifference curve is described by

\[ y = y_0 \left( \frac{fs}{fs_0} \right)^{0.121} \]  

(35)

where the ratio \( \beta/\alpha \) is replaced by \( \beta(1-\alpha) \). We notice that this power is 0.26. This value is evidently very well in line with the other values in Table 6. Hence our conclusion is that the ordinal information, which can be extracted from the true welfare function is the same as that which is provided by the FS- question.

Obviously, we may also try to explain the six separate responses on the IEQ, that is the household cost levels \( \ln(c_i) \). The resulting regression equations are given in table 8.

Table 8. Ordinal analysis of the six level equations of the IEQ.

<table>
<thead>
<tr>
<th></th>
<th>C_1</th>
<th></th>
<th>C_2</th>
<th></th>
<th>C_3</th>
<th></th>
<th>C_4</th>
<th></th>
<th>C_5</th>
<th></th>
<th>C_6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma_i )</td>
<td>3.499</td>
<td>33.653</td>
<td>3.422</td>
<td>42.488</td>
<td>3.447</td>
<td>46.647</td>
<td>3.558</td>
<td>51.176</td>
<td>3.788</td>
<td>51.033</td>
<td>3.961</td>
<td>41.326</td>
</tr>
<tr>
<td>( \alpha_i )</td>
<td>0.468</td>
<td>35.134</td>
<td>0.507</td>
<td>49.193</td>
<td>0.527</td>
<td>55.774</td>
<td>0.539</td>
<td>60.572</td>
<td>0.550</td>
<td>57.904</td>
<td>0.571</td>
<td>46.534</td>
</tr>
<tr>
<td>( \beta_i )</td>
<td>0.165</td>
<td>12.870</td>
<td>0.149</td>
<td>14.996</td>
<td>0.141</td>
<td>15.490</td>
<td>0.130</td>
<td>15.144</td>
<td>0.089</td>
<td>9.706</td>
<td>0.056</td>
<td>4.715</td>
</tr>
</tbody>
</table>

System Weighted R-Square: 0.2318

The errors are strongly correlated as we see from Table 9.

Table 9. The cross-model error correlation matrix.

<table>
<thead>
<tr>
<th></th>
<th>C_1</th>
<th></th>
<th>C_2</th>
<th></th>
<th>C_3</th>
<th></th>
<th>C_4</th>
<th></th>
<th>C_5</th>
<th></th>
<th>C_6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C_1</td>
<td>1</td>
<td>0.906</td>
<td>0.836</td>
<td>0.744</td>
<td>0.611</td>
<td>0.467</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_2</td>
<td>0.906</td>
<td>1</td>
<td>0.963</td>
<td>0.887</td>
<td>0.784</td>
<td>0.630</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_3</td>
<td>0.836</td>
<td>0.963</td>
<td>1</td>
<td>0.951</td>
<td>0.856</td>
<td>0.706</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_4</td>
<td>0.744</td>
<td>0.887</td>
<td>0.951</td>
<td>1</td>
<td>0.917</td>
<td>0.772</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_5</td>
<td>0.611</td>
<td>0.784</td>
<td>0.856</td>
<td>0.917</td>
<td>1</td>
<td>0.899</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C_6</td>
<td>0.467</td>
<td>0.630</td>
<td>0.706</td>
<td>0.772</td>
<td>0.899</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For a more extensive ordinal analysis see Van Praag and Van der Sar (1988), where similar results for other data sets were found. Our conclusion is that the coefficients for the separate levels are not equal, but that they follow exactly the same pattern as in Van Praag and Van der Sar. At a low level of satisfaction the dependency on own
income is considerable at 0.442, but it increases as the level of satisfaction increases up to 0.593 at the highest level of satisfaction.

The family size effect $\beta$ behaves just in the opposite way. It falls with rising levels of satisfaction (see also (Van Praag, Flik (1992)) for a comparison with other European data sets). We may stamp the effect of family size as reflecting real needs, while the dependency on own income stands for a psychological reference effect. Our findings may then be summarized as: when individuals become richer, their real needs become less pressing and their norms become more determined by reference effects.

We may calculate for each verbal level $i$ the income amount $y_i$, which is evaluated by $i$. For that level there holds

$$\ln(y_i) = \alpha_i \ln(y_i) + \beta_i \ln(fs) + \gamma_i \quad (i=1,...,6) \quad (36)$$

which yields

$$\ln(y_i) = \frac{\beta_i \ln(fs) + \gamma_i}{1 - \alpha_i} \quad (i=1,...,6) \quad (37)$$

We notice that the resulting family size elasticity is $\beta_i / (1 - \alpha_i)$. We notice that the elasticities and the corresponding household equivalence scales hardly vary between the satisfaction levels $i$.

Now we present the household equivalence scales for the six levels $i$ and those derived from the $\mu$ - equation side by side.

<table>
<thead>
<tr>
<th>Household size</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>$\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>81%</td>
<td>81%</td>
<td>81%</td>
<td>82%</td>
<td>87%</td>
<td>91%</td>
<td>84%</td>
</tr>
<tr>
<td>2</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>3</td>
<td>113%</td>
<td>113%</td>
<td>113%</td>
<td>112%</td>
<td>108%</td>
<td>105%</td>
<td>111%</td>
</tr>
<tr>
<td>4</td>
<td>124%</td>
<td>123%</td>
<td>123%</td>
<td>122%</td>
<td>115%</td>
<td>109%</td>
<td>119%</td>
</tr>
<tr>
<td>5</td>
<td>133%</td>
<td>132%</td>
<td>131%</td>
<td>129%</td>
<td>120%</td>
<td>113%</td>
<td>127%</td>
</tr>
<tr>
<td>6</td>
<td>141%</td>
<td>139%</td>
<td>139%</td>
<td>136%</td>
<td>124%</td>
<td>115%</td>
<td>133%</td>
</tr>
<tr>
<td>7</td>
<td>147%</td>
<td>146%</td>
<td>145%</td>
<td>142%</td>
<td>128%</td>
<td>118%</td>
<td>138%</td>
</tr>
<tr>
<td>8</td>
<td>154%</td>
<td>152%</td>
<td>151%</td>
<td>148%</td>
<td>131%</td>
<td>120%</td>
<td>143%</td>
</tr>
</tbody>
</table>

We see that the differences between all subjective scales are rather small.
Van Praag and Flik (1992) derived equivalence scales for various European countries by the same IEQ- method. They noticed that the scales in various countries are not the same, reflecting cultural differences and differences in social systems. See also Hagenaars (1986) and Goedhart et al. (1977).

Finally, it is interesting to compare the results derived from the IEQ- responses with our results, based on Financial Satisfaction -responses. The resulting trade- offs, derived from the true WFI, and the ratios found earlier are very similar. The additional result that we can derive from the IEQ and which we cannot find from financial satisfaction questions, because they only refer to current income, is the virtual WFI. As said before the true WFI corresponds to the experienced utility function and the virtual WFI to the decision utility function.

5. Conclusion

Let us now summarise the conclusions of this paper.

- We found that income satisfaction can be explained by objective factors. This yields trade - off coefficients between family size and income and trade-off coefficients between adults and children.
- We found that the Ordered Probit - method is based on an implicit frequentist utility assumption, which may be interpreted as a cardinalist approach as well.
- We saw that we may replace the O. Probit method by the method of Probit-Adjusted Ordinary Least Squares (POLS) and that the results do not vary except for a multiplication factor.
- We found that we can use the cardinal information in Financial Satisfaction Questions leading to a Cardinal Probit - and a Cardinal OLS - approach.
- The frequentist and the cardinalist approach imply two different cardinalisations of satisfaction, which are related by an affine linear transformation.
• The empirical estimates according to those four estimation methods are strongly related and yield (almost) the same trade-off ratios.
• The POLS- and COLS methods are computationally easier.
• An earlier way to study income satisfaction in a quantitative way has been developed by Van Praag and Kapteyn ('Leyden School'). In this chapter we compare their results derived from the Income Evaluation Question with results derived from the Financial Satisfaction Question (FSQ). We found that both methods yield approximately the same trade-off coefficients.
• The FSQ yields an experienced utility function in the terms of Kahneman et al. The IEQ yields a virtual and a true individual welfare function, which concepts coincide with Kahneman et al.'s decision utility and experienced utility functions, respectively. The result of this comparison is that most results derived by WFI-analysis by or in the spirit of the 'Leyden School' could have been derived by analysis of the FSQ as well.
• The FSQ is easier to answer for respondents than the IEQ. Moreover, the IEQ-format does not seem applicable when we ask for Health Satisfaction, Housing Satisfaction, etc., while the FSQ can be used. However, the IEQ yields information on the decision utility function, which the FSQ cannot provide.

References

Edgeworth, F.Y., [1881], 1961, Mathematical Psychics, A.M. Kelley, N.Y.


