



**UvA-DARE (Digital Academic Repository)**

**The measurement of welfare and well-being: the Leyden approach**

van Praag, B.M.S.; Frijters, P.

*Published in:*  
Well-being: the foundations of hedonic psychology

[Link to publication](#)

*Citation for published version (APA):*  
van Praag, B. M. S., & Frijters, P. (1999). The measurement of welfare and well-being: the Leyden approach. In E. Diener, D. Kahneman, & N. Schwarz (Eds.), *Well-being: the foundations of hedonic psychology* (pp. 413-433). Russell Sage Foundation.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<http://dare.uva.nl>)*

# The measurement of welfare and well-being

## *The Leyden approach*

Bernard M.S. Van Praag and Paul Frijters

### Contents

1	From economics to psychology . . . . .	1
2	General approaches to cardinal utility taken by economists . . . . .	6
3	Utility measurement based on verbal qualifiers . . . . .	10
	3.1 The Leyden approach . . . . .	10
	3.2 The shape of the welfare function . . . . .	14
	3.3 The definition of income . . . . .	18
4	The explanation of the welfare function . . . . .	18
5	Dynamics . . . . .	23
6	Methodological discussion . . . . .	26
	6.1 Individual norms and general standards . . . . .	29
7	Future directions: well-being <i>and</i> welfare . . . . .	31
8	Conclusions . . . . .	35

### 1. From economics to psychology

The utility concept is a key concept in economics. It is well-known that modern economics is a discipline with numerous subfields, but nearly all relevant problems have to do with people and people's choice behavior. Individuals have limited resources and opportunities and therefore must choose between alternatives, for example, 1, 2, 3, ...,  $i$ , ... An efficient way to describe the choice problem is to attach a *utility value* to these alternatives, for example,  $U_1, U_2, U_3, \dots, U_i, \dots$  and to postulate that an individual chooses the alternative which has the highest utility value for him. For example, if there is a choice set  $\{1, 2, 3, \dots, i, \dots\}$ , then the choice behavior is described mathematically by

$$\max_{i=1,2,3,..} U_i$$

The implication of this description is that we could predict the individual's choice behavior by knowing his utility values  $U_1, U_2, U_3, ..$ . In empirical reality it is the other way around. We do not know the values of  $U$ , but we may observe the choice process. If an individual consistently chooses alternative 1, economists generally infer that  $U_1$  is larger than  $U_2, U_3, ..$ . If we then remove alternative 1 from the choice set and 2 is chosen consistently we know that  $U_1 > U_2$  and that  $U_2$  is larger than other  $U$  values. In this way, it is possible to find the preference ordering of the alternatives and also to establish inequality relations between the  $U$  values. However, we are unable to say whether  $U_2$  is a *little* smaller than  $U_1$  or if  $U_2$  is *much* smaller than  $U_1$ . In short, by observing choices we get an *ordinal* utility ordering.

The choice model may be extended in two ways. First, we can consider a set of alternatives that is infinite. Alternatives can be described by a continuous variable  $x$  or by more than one variable, for example,  $(x_1, x_2, .., x_n) = \mathbf{x}$ . Then the utility values are denoted by the ordinal utility function  $U(\mathbf{x})$ . Secondly, we may assume that each decision maker  $z$  has his own utility ordering. In that case, the ordinal utility function reads  $U(\mathbf{x}; z)$  where  $z$  may incorporate individually varying parameters such as age, gender, income, social class, etc. We notice that this ordinal function is of the *decision-utility* type in the terminology of Kahneman, Wakker and Sarin (1997). It is needed to make decisions and is empirically established by observations of choice decisions<sup>1</sup>.

The traditional example of choice behavior in economics is the *purchasing* behavior of consumers. The model starts from a utility function

$$U(x; z)$$

where  $x$  stands for quantities of commodities purchased and  $z$  for characteristics describing the individual's circumstances (e.g. age, gender). The consumer is faced by prices  $p_1, .., p_n$  for goods  $x_1$  to  $x_n$ . If he has income  $y$ , his choice set is described by:

$$p_1x_1 + .. + p_nx_n \leq y$$

---

<sup>1</sup>Decision utility in the sense of Kahneman et al. may be ordinal or cardinal.

Any commodity bundle  $(x_1, x_2, \dots, x_n)$  violating the constraint is too expensive for him. The behavioral model explains behavior by assuming that individuals maximize  $U(\cdot; z)$  with respect to the feasible commodity bundle  $x$  subject to the freedom given by the choice set.

Edgeworth (1881) called  $U(\cdot)$  the utility function; Pareto (1904) called it the ophelimity function. Edgeworth more or less implicitly assumed that  $U$  could be measured in a direct way. Samuelson (1945) therefore stated that "Edgeworth considered utility to be as real as his morning jam." Edgeworth interpreted  $U$  as *experienced utility*, that is, a cardinal measure of the joy which the individual derives from the commodity bundle. Pareto became aware of the fact that it could be difficult to establish the individual's utility function over goods. For the description of the consumer choice process, an *experienced utility* function appeared to be unnecessary. Actually, it is a choice between alternatives which may be described by an *ordinal* utility function as described above. If  $U(\cdot)$  is an ordinal utility function, any other utility function which assigns the *same ordering* of utility to the alternatives is also a utility function describing that same choice process. For example, if  $U_1 > U_2 > U_3 > 0$  describes the choice process between alternatives 1,2,3 then  $\tilde{U}_1 = \sqrt{U_1} > \tilde{U}_2 = \sqrt{U_2} > \tilde{U}_3 = \sqrt{U_3}$  will describe the same process. Hence there is a whole equivalence class of *ordinal* utility functions describing the same preference structure.

It is an error to assume that Pareto denied the existence of meaningful cardinal utility measurement or the possibility of measuring it, but he pointed to the fact that utility in the cardinal sense could not be measured by observing consumer behavior, and moreover that it was unnecessary to do so for consumer studies.

Robbins (1932), who had a tremendous influence on economics, was the first to proclaim that utility was immeasurable and that it was more or less a scientific folly to endeavor to measure it. At the very least, it should be left to psychologists.

Other economists such as Pigou (1948), and the Nobel laureates Tinbergen (1991) and Frisch (1932) were certainly of a different opinion.

However, the ordinal line has been continued by Arrow (1951) and Debreu (1959) who were able to include decisions over time and/or under uncertainty in this ordinal framework. They assume a preference ranking described by a utility function on the dated commodity space. Behavior is subject to a budget constraint where the consumption of goods and the prices of those goods are differentiated according to the date of consumption.

Similarly, they incorporate uncertainty by distinguishing states of nature  $s$  varying over  $S$  and commodities available only if  $s$  prevails. Commodities are

then available *contingent* on the status of nature, which is *a priori* not known to an individual. It can be shown that the model describes consumer choice behavior, but it is also clear that this model leads to a decision problem with an unworkable number of dimensions. Its realism as a positive behavioral model is not significant and it has never been used according to our knowledge in empirical work, except in very simplified versions.

In practice, economists are frequently confronted by problems where more is needed than the ordinal concept (see also Ng (1997)).

We think of decisions under *uncertainty*, the basis of insurance theory, investment and saving behavior. Decisions which have to do with different *time* periods such as saving and investment decisions need more than the ordinal concept. The objective function in such models is usually simplified to an additive form such as  $\sum_t w_t U_t$  or  $\sum p_s U_s$  where  $U_t$  stands for *instantaneous* period utilities and  $w_t$  time-discounting weights, and where  $U_s$  stands for state-contingent utility and  $p_s$  for the (real or perceived) chance that state  $s$  occurs. Evidently, time-state mixtures and continuous generalizations are easy to think of.

There are two points of interest in these objective functions. The basic ingredient is a utility function  $U$  which is no longer ordinal. We cannot change the individual form at will according to a monotonous transformation. More specifically, maximizing  $\sum_t w_t \varphi(U_t)$  will yield an optimum which varies with  $\varphi(\cdot)$ , except if  $\varphi$  is a positive linear transformation (i.e.,  $\varphi(\cdot) = \alpha U + \beta$  with  $\alpha > 0$ ). The utility concept in these kinds of problems is what economists call a *cardinal* utility function. It is a much smaller class which only allows for positive linear transformations.

Most mainstream economists have a very uneasy feeling about cardinal utility functions. This uneasiness seems to be based on the Anglo-American dogmatism against cardinality instilled by Robbins. However, most actual studies conducted by economists start with very general "ordinal formulations", but after a while they present a structural specification which in nine times out of ten turns out to be of the cardinal type (see also Van Praag (1968)). These cardinal utility functions are still of the "decision-utility" type. They are instrumental to the description of *decision* processes.

There is a second class of problems for which economists need cardinal utility functions: *normative problems*. The first example of such problems arises if we try to look for optimal (re)distributions. Notably in income taxation, a progressive tax schedule (richer individuals pay relatively more tax than poorer individuals)

is advocated so that that the rich man suffers as much as the poor man. Such comparisons are impossible without a cardinal and interpersonally comparable utility function. Obviously, these utility functions are of the *experienced* utility type.

A second example is provided by equity measures: the concepts of a just income distribution and poverty and the evaluation of income inequality. It is evident that nearly all of these measures are based on a cardinal concept of experienced income utility, though this is rarely mentioned explicitly (c.f. Atkinson (1970)).

A third field where interpersonally comparable and cardinal utility is needed concerns all types of cost-benefit analyses, where specific measures such as building a bridge, deregulation of markets, specific health insurance programs, noise pollution by an airport, etc., have to be evaluated. In these cases, some citizens will profit and others will lose. Those benefits and costs may be partially translated in monetary amounts, but money means different things for different people. For example, when a policy means a loss of \$100 to a poor man and a gain of \$10,000 to a rich man, it is not at all evident that the policy should be realized. The only way to make a decision is to create a balance in terms of comparable utility gains and losses.

The situation in economics is succinctly and wittily summarised by Wansbeek and Kapteyn.

”Utility seems to be to economists what the Lord is to theologians. Economists talk about utility all the time, but do not seem to have hope of ever observing it this side of heaven. In micro-economic theory, almost every model is built on utility functions of some kind. In empirical work little attempt is made to measure this all-pervasive concept. The concept is considered to be so esoteric as to defy direct measurement by mortals. Still, in a different role, viz., of non-economists, the same mortals are the sole possessors of utility functions and can do incredible things with it.” (Kyklos, 36, pp. 249-269, 1983).

By detaching economics from the psychology of ”feelings”, economists have found it difficult to have anything relevant to say on a whole range of issues. In the second part of this paper, we will review an attempt made by economists to measure utility functions using the evaluations given by individuals themselves. Before we do so however, we will first discuss the approaches taken in general to utility functions in the economic literature. We will divide the approaches

that have been taken concerning the problem of utility functions in five distinct approaches.

## 2. General approaches to cardinal utility taken by economists

The first approach to cardinal utility, which is by far the most popular in the economic profession, is not to measure utility at all, but to simply assume a functional form of the utility function for the theoretical or empirical problem at hand. We will ignore this approach in the remainder of this chapter.

Economists who use the second approach, of which perhaps the best-known are Christensen et al. (1985) and Jorgenson and Slesnick (1984), have taken an axiomatic approach to utility functions. They specify the conditions they believe a utility function should satisfy and then derive (a shape of) the utility function which fits these requirements.<sup>2</sup> They then infer the level of utility that individuals enjoy from their observed behavior, which they then use to make normative statements. As utility levels are not directly measured in this approach, but essentially assumed, this approach is not elaborated here. Moreover, if this method has validity, it yields a cardinal decision utility.

Economists who take the third approach use subjective and objective indicators of the work and living conditions of individuals to define a measure of utility. This large group is subdivided into three groups: one group is concerned with poor individuals, another with the quality of life of nations, another with the quality of life of individuals.

The empirical literature on poverty centers around the material resources available to individuals (Townsend (1979, 1993), Sen (1987), Ravallion (1994)). The standard approach is to define households to be poor if their household income falls below a certain cut-off point. This cut-off point can be defined in several ways. For instance, in the "basic needs" approach, the cut-off point is calculated from the expenditures needed to buy a basket of commodities that the researcher considers vital for individuals. In the relative needs approach, the cut-off point is defined to be a certain percentage of the average or median income in a country. It is clear that neither approach, which together form the bulk of the poverty literature, actually measures utility functions, but that they are based on the *assumption* that the utility of individuals whose income is below the cut-off point

---

<sup>2</sup>See also Van Praag (1968) for an attempt to find a functional form of the utility function with the use of axioms and secondary assumptions.

is in some sense "low". See Callan and Nolan (1991) and Frijters and Van Praag (1996) for a more detailed review of the normative issues involved in poverty measurement.

Other literature examines the "quality of life" of nations. In this literature (Kurian (1984), Nussbaum and Sen (1992), Sen (1987), Maasoumi (1989)), economists attempt to rank countries with respect to the quality of life.<sup>3</sup> The quality of life is usually defined as a weighted average of specific country statistics. The statistics used include, for instance, the literacy level of the entire population, the literacy level of women, infant mortality rates, income levels per head, life expectancies of men and women, indicators of political stability, energy consumption per capita, average household size, the number of persons per physician, levels of civil liberties, etc. It is clear that these variables may be very important for the utility levels of individuals and nations, however the utility levels themselves are *not* measured by these variables. An obvious problem is then how one should weigh these statistics: does the quality of life increase more when the female literacy level increases by 1% or when the civil liberty index improves by 1%? It will be clear that if one does not want to use the evaluations of individuals themselves as a weighting method, the opinions of the researcher become the deciding criterium. The problem of how to weigh these different variables into a composite quality of life index is, not surprisingly, the main source of disputes in this literature. For an empirical analysis of some of the weighting methods employed, see Hirschberg et al. (1991).

Some of the works of Clark and Oswald (e.g. Clark and Oswald (1994)) also belong to the third category. In their 1994 paper, Clark and Oswald define "unhappiness" by aggregating the answers to the following 12 questions:

1. Been able to concentrate on whatever you are doing?
2. Lost much sleep over worry?
3. Felt that you are playing a useful part in things?
4. Felt capable of making decisions about things?
5. Felt constantly under strain?
6. Felt you couldn't overcome your difficulties?
7. Been able to enjoy your normal day-to-day activities?

---

<sup>3</sup>The "quality of life" concept is very broad and interpreted by some to mean the same thing as happiness (e.g. Veenhoven (1996)), or average satisfaction (e.g. Dow and Juster (1985)). We discuss here the interpretation we believe most *economists* in this field use.



8. Been able to face your problems?
9. Been feeling unhappy and depressed?
10. Been losing confidence in yourself?
11. Been thinking of yourself as a worthless person?
12. Been feeling reasonably happy, all things considered?

The variable "unhappiness" ranges from 0 to 12, whereby 12 denotes the maximum level of unhappiness and 0 a complete lack of unhappiness. Although some of these questions could arguably be seen as a measure of utility, such as questions 9 and 12, the simple aggregate of all 12 questions cannot be seen as a direct measurement of utility: utility is an evaluation of an individual of his circumstances. Although "losing a lot of sleep" or "being under strain" may affect utility or may be affected *by* utility, they do not directly measure a utility level for they are not an evaluation of "losing sleep" or "being under strain". This measure of happiness may correlate perfectly with the experienced utility of individuals and may hence be as useful as any other measure of experienced utility. Nevertheless, it remains an *indirect* measure of experienced utility which is useful *only* if "losing sleep" and "being under strain" correlate with experienced utility (which seems very likely). Hence, it is a measure of the quality of life entirely on its own. Clark and Oswald seem to acknowledge this by arguing that the individual scores are "more accurately" described as "mental stress" scores. Other individual measures of the quality of life of an individual which are based on aggregations of individual circumstances also fall into this category.

A fourth approach was to estimate decision utility functions by performing probability-choice experiments on individuals. When individuals must choose between either a certain outcome  $Y$  or a lottery in which fate decides whether they will receive an outcome less than  $Y$  or an outcome greater than  $Y$ , individuals will reveal the relative attractiveness of the sure  $Y$  versus the proposed lottery. The main problem in this line of research has been that individuals are not good at using probabilities: they overestimate small probabilities and underestimate large probabilities, as was first demonstrated by the Allais paradox (see Allais and Hagen (1979)). This means that the choice of an individual for a lottery is the result of a combination of the individual's valuation of the outcomes and of the individual's perception of the probability of the outcomes. Following the theoretical advances by Kahneman and Tversky (1979) and Wakker and Tversky (1993), Kahneman et al. (1991) have managed to isolate the effect of gains and

Figure 2.1:

losses on individual's evaluation of outcomes. We will ignore the results on probabilities and focus on the value function they find. The shape of the value function suggested by the choice experiments of Kahneman et al. (1991) is sketched in figure 1.

The main feature of this function is that losses are found to have a greater impact than gains. A second characteristic of this value function is that it levels off at either end of the loss or gain scale, which implies decreasing marginal value of losses and gains. It implies a convex-concave shape, also suggested by Markowitz (1952) and Van Praag (1968).

Finally, the fifth approach taken is to assume that individuals are able to describe their utility level by means of *verbal* qualifications. The rest of this chapter will be devoted to the attempts of economists who belong to the Leyden group where this approach was initiated. There are of course also other economists who use verbal qualifiers as measures of experienced utility (e.g. Clark (1996), Clark and Oswald (1996), Heywood et al. (1997), Levy-Garboua and Montmarquette (1997), Easterlin (1974, 1995), Dow and Juster (1985), and Gershuny and Halpin (1995)).

### 3. Utility measurement based on verbal qualifiers

#### 3.1. The Leyden approach

In this section we will discuss an economic line of research which tries to operationalize the concept of experienced utility. It originated at Leyden University in the Netherlands in the early 1970s. Its main contributors are Van Praag, Kapteyn, Wansbeek, Hageraars, Van der Sar, Plug, and Frijters. It is known in the literature as the Leyden approach (or school). For psychologists, the ideas in this approach may not appear alien, but for most economists they were and still are. Most economists still believe that cardinal experienced utility is unmeasurable and that any measurement should be based on observed decision behavior. This meant that the Leyden approach met with stiff opposition, disbelief and outright hostility. The most outspoken example of this attitude is found in an article by Seidl (1994) in the *European Economic Review* where he criticises Van Praag (1968)<sup>4</sup>.

Although Van Praag (1968) served as a theoretical basis, the ensuing literature on the Leyden approach started with Van Praag (1971) and is mainly empirical and data-oriented. The Leyden approach focuses primarily on the evaluation of *income*, although in later work the focus was extended. We also speak of utility of income, income satisfaction, or, in other words, economic *welfare*. We drop the adjective economic from now on, but when we use the term welfare, we have welfare derived from income in mind. This concept is narrower than the concept of well-being which includes feelings associated with factors unconnected to income or purchasing power. In Section 7 we shall consider well-being and its relationship with welfare in greater detail.

The Leyden approach is based on two assumptions. The first is that individuals are able to evaluate income levels in general and their own income in particular in terms of "good", "bad", "sufficient", etc. We call these terms *verbal qualifiers*. The second assumption is that verbal labels can be translated in a meaningful way into a numerical evaluation on a bounded scale, for example  $[0,1]$ . We shall consider both steps of the measurement procedure in detail.

If we are interested in how a specific income level is evaluated, there are two ways to gather information. The first and most natural way is to propose a sequence of income levels and to ask for their verbal qualifications. An example of this type of question follows:

---

<sup>4</sup>A reply was given by Van Praag and Kapteyn (1995).

”Here is a list of income levels per month, after tax: please evaluate these amounts using verbal qualifications, such as ”very bad”, ”bad”, ”insufficient”, ”sufficient”, ”good”, ”very good”:

- \$2000
- \$4000
- \$6000
- \$8000
- \$10,000 ”

It is obvious that someone who earns \$20,000 a month would be unable to make a distinction between most of these levels. All the incomes are insufficient or worse for him. Therefore, instead of starting with income levels, we can also supply the verbal qualifications as *stimuli* and ask the individual respondent which income level corresponds with the verbal label. This leads to the so-called Income Evaluation Question:

”while keeping prices constant, what after-tax total monthly income would you consider for your family as: ”

- very bad*,..... \$.....
- bad*, ..... \$.....
- sufficient*,.....\$.....
- insufficient*,.....\$.....
- good*, ..... \$.....
- very good*, ..... \$.....

This question appears to have been succesful in anonymous mail-questionnaires, although it has also been posed orally with success. Theoretically, finding a continuous relationship between income and utility would require an infinite number of levels, but, in practice, between four to nine levels have been and can be used. We will discuss here the format used most often, the six-level format.

The question is now how we derive a welfare function from the answers to this question. Or more precisely, how we translate the verbal labels into numbers

on a  $[0,1]$  scale. Following Van Praag (1971), we make an assumption about the way individuals fill out the question. We assume that respondents try to provide information to the interviewer about the shape of their welfare function. The most accurate way for individuals to provide information then depends on the accuracy criterion. Van Praag (1971) and Kapteyn (1977) show that, for a broad class of intuitively plausible criterion functions, the best way for a respondent to provide information is to choose the answers in such a way that each of the six levels corresponds to a jump of  $1/6$ . This is the so-called *equal quantile assumption* (EQA). It implies that

$$\begin{aligned}
 U(\text{very bad}) &= U(\text{first interval}) = 1/12 \\
 U(\text{bad}) &= U(\text{second interval}) = 3/12 \\
 &\quad \dots \\
 &\quad \dots \\
 U(\text{very good}) &= U(\text{last interval}) = 11/12
 \end{aligned}$$

It may be surmised that, even if the verbal descriptions are somewhat vague, the respondent will tend to interpret the question as if it were an equal partition. Only if the verbal labels are ambiguous, practically equal or strongly suggest an unequal partition, we should no longer expect this effect.

If the number of verbal labels is  $k$ , the general formula for the welfare corresponding to the  $i^{\text{th}}$  verbal label is obviously  $\frac{2i-1}{2k}$ . This reasoning and the EQA assumption are very similar to the thesis developed by Parducci (see e.g. Parducci (1995)). It is obvious that this translation of verbal labels into numbers is a linking pin in this measurement procedure. It has been subject to criticism by some economists, while on the other hand experimental psychologists do not find much to criticize: it is a type of Thurstonian measurement. If we do not accept this or any translation into figures, it is obvious that a meaningful analysis of the response is severely hampered, although not impossible (see later).

In Van Praag (1991), an experiment is described in which five labels were supplied and 364 respondents were asked to "translate" these verbal labels onto a  $[0,100]$ -scale. Similarly, the same labels had to be linked with line segments. Both the numbers between  $[0,100]$  and the lengths of the line segments were re-scaled onto a  $[0,1]$  mapping. We present the average results for 364 respondents in Table 1.

Table 1: Translation into numbers and line segments

Numbers:	Emprical mean	St. dev.	Theor. pred.
very bad	$\bar{v}_1 = 0.0892$	0.0927	0.1
bad	$\bar{v}_2 = 0.2013$	0.1234	0.3
not bad, not good	$\bar{v}_3 = 0.4719$	0.1117	0.5
good	$\bar{v}_4 = 0.6682$	0.1169	0.7
very good	$\bar{v}_5 = 0.8655$	0.0941	0.9
Line segments			
very bad	$\bar{w}_1 = 0.0734$	0.0556	0.1
bad	$\bar{w}_2 = 0.1799$	0.0934	0.3
not good, not bad	$\bar{w}_3 = 0.4008$	0.1056	0.5
good	$\bar{w}_4 = 0.5980$	0.1158	0.7
very good	$\bar{w}_5 = 0.8230$	0.1195	0.9
Source: Van Praag (1991)			

For the "numbers" case, one can see that all averages fall within a one  $\sigma$ -interval of their theoretical prediction. This also holds for all levels for the line segments, except one. It is intriguing that the averages are all *below* their theoretical prediction. Perhaps this is due to the order in which the verbal labels were supplied. We think, but do not know, that the bias would have been the other way around if the order in which the verbal labels were supplied was reversed. When we regress the translation of the verbal labels into numbers by individual  $i$ , say  $v_{i,n}$ , onto the translation of the verbal label into a line-segment, say  $w_{i,n}$ , we find

$$v_{i,n} = 0.056 + 0.974w_{i,n}$$

$$(0.005) (0.010)$$

$$R^2 = 0.848$$

for 364\*5 observations, where we did not account for the fact that the five level disturbances per individual will be strongly correlated. The fit is however remarkably good. From Table 1 and this regression, we can draw some tentative conclusions:

1. A verbal label sequence seems to be understood in a similar way by different respondents, irrespective of the context of the individual respondent.
2. A verbal label sequence may be translated on a numerical scale or on a line scale: in both cases the translations are uniform over individuals.

3. Translations via various translation mechanisms (lines and figures) are consistent with each other. That is, we seem to be measuring the same thing, irrespective of whether we use line segments or numbers.
4. The verbal labels are translated on a bounded scale roughly in accordance with the Equal Quantile Assumption.

An interesting point is that these results were found in a context-free setting, that is, the respondents did not know which concept they were evaluating.

A final point of critique is whether the verbal labels "good", "bad", etc., convey the same feeling to every respondent. If not, we falsely assume that individuals derive the same degree of joy from their income, when describing the same verbal label. Actually, this is a question of psycho-linguistics. Generally, the basic idea of language is that frequently used words will have the same meaning and emotional connotation for the members of a language community. It is the main tool of communication between people. Hence, we must assume that verbal labels like "good", "bad", etc., mean approximately the same thing to all respondents sharing the same language.

### **3.2. The shape of the welfare function**

For each respondent we now have six income levels connected to six utility levels. The shape of the function can be inferred from these six combinations. Many functions can be fitted using these six points. In Van Praag (1968), it was argued on theoretical reasons that it would be a lognormal distribution function. The reason to use a distribution function is that we assume boundedness of the utility function: there is a worst and a best position in terms of welfare (satisfaction). It is also known that the Von Neumann-Morgenstern model requires a bounded utility function (see Savage (1954)).

Van Herwaarden and Kapteyn (1981) showed that the points of the welfare function, which were found empirically, best fitted a lognormal curve within the class of distribution functions. The logarithmic function did slightly better, but it is not bounded. Also, the logarithmic function is not borne out by the choice experiments of Kahneman et al. (1991) and others: the marginal effect of greater losses is found to decrease, whereas the logarithmic function would imply that they should increase.

The lognormal function is defined as

Figure 3.1:

$$\Lambda(y; \mu, \sigma) = N(\ln y; \mu, \sigma) = N\left(\frac{\ln y - \mu}{\sigma}; 0, 1\right)$$

where  $N(\cdot; 0, 1)$  stands for the standard lognormal distribution function. The log-normal function is sketched in Figure 2. Notice the resemblance to the shape suggested by the experiments of Kahneman et al. (1991): in both cases the function is S-shaped. Also, it is generally the case that losses to an individual have a greater effect than gains.<sup>5</sup>

The parameter  $\mu$  is interpreted by realizing that  $\Lambda(e^\mu; \mu, \sigma) = 0.5$ . Hence, the income level  $e^\mu$  is halfway between the worst and the best situation.

There are two interesting aspects about this function. First, the function is not concave for all income levels, but convex for low incomes. This runs counter to mainstream economic assumptions. In economics, it is conventional wisdom that the utility function of income is always concave. This is known as the so-called Law of Decreasing Marginal Utility, also known as Gossen's first law. It has always been based on introspection. Concavity implies that individuals are risk-averse, but scientific experiments with insurance or gambling behavior show

---

<sup>5</sup>One particular feature of the value function found by Kahneman et al. (1991) cannot be replicated: Kahneman et al. find a value function which changes direction abruptly at the reference position. The number of levels used in our measure is simply too small to find such a jump in direction.



Figure 3.2:

that this is not always true; it therefore follows that a utility function may be convex in certain regions.<sup>6</sup>

The second point of interest about the lognormal utility function consists of the two parameters  $\mu$  and  $\sigma$  which may *vary individually*. Two functions with different  $\mu$  and equal  $\sigma$  are sketched in Figure 3a. In Figure 3b two functions are sketched with different  $\sigma$  and equal  $\mu$ .

One can see that as  $\mu$  increases, the individual needs more income to reach the same welfare level. For instance, in order to reach the welfare level 0.5, the person A with  $\mu_A = \ln(4000)$  needs \$4000 per month, while B needs \$6000 per month to reach the same welfare level. If the welfare levels of individuals A and B are to be equal for other welfare levels (if  $\sigma$  is equal for both persons), it should hold that

$$\ln y_A - \mu_A = \ln y_B - \mu_B$$

Hence, for any welfare level, income levels are equivalent to A and B if

---

<sup>6</sup>A variable of much economic interest, Pratt's (1964) measure of relative risk aversion (or Frisch flexibility) can be directly calculated as

$$\frac{\partial \ln u}{\partial \ln y} = -\frac{1}{2\sigma^2}(\ln y - \mu) - 1$$

It varies from highly positive for small  $y$  to very negative for large  $y$ .

Figure 3.3:

$$\ln \frac{y_A}{y_B} = \mu_A - \mu_B$$

and therefore

$$\frac{y_A}{y_B} = e^{\mu_A - \mu_B} = \frac{4000}{6000}$$

Hence, a change in  $\mu$  implies a proportional shift of the welfare function. One of our main preoccupations in the remaining section is to discover why individual's  $\mu$ -values differ.

The parameter  $\sigma$  defines the slope of the welfare function.

In Figure 3b, two functions are sketched with  $\sigma_A < \sigma_B$ . If  $\sigma = 0$  we get the limiting case where individuals are completely unsatisfied with any income until their income reaches  $e^\mu$ , and where they are completely satisfied if income exceeds  $e^\mu$ . It is the welfare function of a Hermit. The parameter  $\sigma$  is called the *welfare sensitivity* of the individual.

The parameters  $\mu$  and  $\sigma$  are estimated for each individual by

$$\hat{\mu} = \frac{1}{6} \sum \ln c_j \quad \text{and} \quad \hat{\sigma} = \frac{1}{5} \sum (\ln c_j - \mu)^2$$

where  $c_1, \dots, c_6$  stand for the six income levels reported in the IEQ.

### 3.3. The definition of income

In the usual IEQ version, the income concept is after-tax *monthly* household income. In some versions income *per year* has also been used and/or *before tax* income (see Dubnoff et al. (1981)). The choice of the definition should be adapted to what is well-known to the individual. Hence, an entrepreneur who knows his annual income better than his monthly income should be questioned in terms of his annual income, while a civil servant who is paid monthly should be approached in terms of his monthly income.

## 4. The explanation of the welfare function

In mainstream literature, it is always assumed that the utility function of income is the same for all individuals. A major finding of our empirical research, although intuitively completely plausible, is that individual welfare functions differ between individuals. When differences are found, the imminent question is whether such differences are structural and can be correlated with observable variables. In our case, this means that we try to "explain" the variable  $\mu$  by other factors, varying per individual and/or environment. In the studies, it appeared that  $\mu$  could be explained to a large extent<sup>7</sup>. The parameter  $\sigma$  posed much more of a problem. We shall therefore concentrate on the explanation of  $\mu$  and assume that  $\sigma$  is constant.

We recall that  $\mu$  determines the position of  $U(y)$ . If  $\mu$  increases, the individual becomes less satisfied with the same amount of income. In other words  $U(y;\mu)$  is decreasing in  $\mu$ . The first determinant that naturally comes to mind is the size of the family to be supported from the income. Income needs are probably also determined by the actual circumstances of the individual, for instance as reflected by the individual's current income  $y_c$ . We therefore expect that needs will increase with family size (denoted by  $fs$ ) and with current income  $y_c$ . Hence,  $fs$  and  $y_c$  are parameters in the individual welfare function. In Van Praag (1971) and Van Praag and Kapteyn (1973), the following simple relation has been found

$$\mu_i = constant + \beta_1 \ln fs_i + \beta_2 \ln y_{i,c} \quad (4.1)$$

In Van Praag and Kapteyn (1973), the following (approximate values) were found  $\beta_1 = 0.1$  and  $\beta_2 = 0.6$ ,  $R^2 = 0.6$ , where  $fs_i$  denotes the number of individuals living in the household of respondent  $i$ , and  $y_{i,c}$  denotes the current household

---

<sup>7</sup>An explanation does not necessarily mean a one-way causal relationship.

income of  $i$ .

Since then, the IEQ has been posed in many countries, and similar results have been found. We give an example drawn from a study on poverty by Van Praag, Hagenaars and Van Weeren (1982), based on a 1979 EUROSTAT survey of eight European countries. Moreover, we add values for Russia estimated by Frijters and Van Praag (1996). In Table 2, we present the regression estimates for the nine countries using the equation

$$\mu_i = \beta_0 + \beta_1 \ln fs_i + \beta_2 \ln y_{i,c} + f(X_i) + u_i$$

where  $X$  denotes a number of variables used in the regression which we do not show (including age, education, employment levels and gender), and  $u_i$  denotes the normally distributed error term.

Table 2: Estimates of welfare parameters for 9 countries  
Countries

	$\beta_1$	$\beta_2$	N	R <sup>2</sup>
Belgium	0.097	0.433	1272	0.695
Denmark	0.075	0.631	1972	0.829
France	0.059	0.505	2052	0.676
W. Germany	0.112	0.583	1574	0.693
Great Br.	0.115	0.364	1183	0.575
Ireland	0.169	0.455	1733	0.636
Italy	0.156	0.381	1911	0.510
The Netherlands	0.100	0.537	1933	0.664
Russia 1995	0.250	0.501	1444	0.501

*Source* : Van Praag, Hagenaars and Van Weeren (1982)

All coefficients are highly significant.

The variables vary over the nine countries but not dramatically. The value of  $\beta_1$  of course depends on the national family allowance system. If the family allowance is high and compensates for the additional child costs, we may expect a  $\beta_1$  of about zero. On the other hand, in poor countries with a less liberal system,  $\beta_1$  may be rather high. This is indeed what we observe: the highest coefficient of  $\beta_1$  is for Russia in 1995 where family allowances and child support are virtually non-existent.

It is not surprising that the satisfaction derived from a specific income level depends on the size of the household. Somewhat more surprising, especially for most economists, is that income satisfaction for any income level, not only for an individual's own current income, depends on an individual's own current income.

Figure 4.1:

It implies that two individual A and B with current incomes  $y_{A,c}$  and  $y_{B,c}$  will evaluate any income differently. More precisely, the following is usually true:

$$U(y_B; f s, y_B) \neq U(y_B; f s, y_A)$$

That is, B evaluates his own income differently than A would evaluate the income of B. It is obvious that this fact is very relevant for the evaluation of social inequality, for the theory of a fair income distribution and for the evaluation of social welfare. The outcomes of such normative evaluations depend on the income norm of the evaluations. Actually,  $U_A(y; f s_A, y_{A,c})$  describes the norms of A with respect to what equals a "bad", "...", "good" income and all levels in between.

A person's income may increase, for example from  $y_c^{(1)}$  to  $y_c^{(2)}$ . The evaluation of this change will be evaluated *differently* before the change and after the change, or, as economists say, *ex ante* and *ex post*. The ex ante evaluation of future income is  $U_A(y_c^{(2)}; f s, y_c^{(1)})$ , while the ex post evaluation is  $U_A(y_c^{(2)}; f s, y_c^{(2)})$ . We sketch the difference between the ex ante and ex post welfare function in figure 4.

Due to the fact that  $\mu$  increases with the income change, the welfare function shifts to the right. The effect of this is that the ex post evaluation of both  $y_c^{(1)}$  and  $y_c^{(2)}$  falls compared to the corresponding ex ante evaluations. It can be seen, and also shown, that the ex ante welfare gain is larger than the ex post gain. As a consequence, the ex ante evaluation is exaggerated when reconsidered later on, or

to put it differently, the income increase will be a disappointment in retrospect. The value of the coefficient  $\beta_2$  is crucial in this context. If  $\beta_2 = 0$ , the curve will not shift to the right and the whole income increase will be translated as a welfare increase. In that case ex ante and ex post evaluations are equal.

On the other hand, if  $\beta_2 = 1$ , perceived welfare will not increase at all. This can be seen by examining

$$\ln y_c - \mu = \ln y_c - \beta_0 - \beta_1 \ln fs - 1.00 \ln y_c = -\beta_0 - \beta_1 \ln fs$$

In this case, the subjective ex post welfare evaluation does not depend on actual income. Evidently, this is a pathological case which has not been found in reality. The anticipated welfare increase would end with a complete deception.

The phenomenon of a shifting welfare function arising from a partial adaptation of income norms to changing current incomes, is what Brickman and Campbell (1971) called the hedonic "treadmill". Van Praag (1971) introduced the term "preference drift" for the same phenomenon.

If all individuals have their own norms with regard to income levels, which depend on their own circumstances, the question is justified whether it is possible to construct *social standards* with respect to what is a "good" income, a "bad" income, etc. This is possible in a certain sense. We define a social standard for a "good" income, say  $\tilde{y}_{good}$ , as that level of income which is evaluated to be "good" by an individual with that current income. If "good" income corresponds with a welfare value of 0.7 on a [0,1] scale, it implies that  $\tilde{y}_{good}$  is the solution to the equation

$$U(\tilde{y}_{good}; fs, \tilde{y}_{good}) = 0.7$$

Using lognormality and our estimate of  $\mu$ , it is possible to show that

$$U(\tilde{y}_{good}; fs, \tilde{y}_{good}) = \Lambda(y_{good}; \frac{\beta_0 + \beta_1 \ln fs}{1 - \beta_2}, \frac{\sigma}{1 - \beta_2})$$

Similarly, we can obtain a social standard income for each possible welfare level, sketched in the next figure.

We call the ensuing welfare function of the social standard income levels, which is also lognormal, a *social standard function*. We know that someone with  $\tilde{y}_{0.4}^*$  current income will evaluate his own income by 0.4. This analysis is frequently used to define a *subjective poverty* line as  $\tilde{y}_{0.4}^*$  for poverty and  $\tilde{y}_{0.5}^*$  as near-poverty. Notice that this line varies as a function of family size. Hence, there is a two-person

Figure 4.2:

household poverty line, a three-person poverty line, etc. The *social standard* function is an obvious tool for social policy and the evaluation of income redistribution and tax policy.

From a social-psychological viewpoint, it is very interesting to compare the welfare sensitivity of the *individual* welfare function  $\sigma$  with the corresponding slope parameter of the social standard function  $\frac{\sigma}{1-\beta_2}$ . If  $0 < \beta_2 < 1$ , the latter function is less steep than its individual counterpart. In other words, the larger the preference drift  $\beta_2$ , or in psychological terms, the stronger the working of the hedonic treadmill, the flatter the social standard curve will be compared to the individual welfare function.

Obviously, the difference between the two functions explains why a person with moderate income, for instance \$40,000 a year, thinks that someone with \$100,000 is rich while the rich person himself with \$100,000 does not perceive himself to be rich. In the same way, people with \$20,000 do not feel as poor as the observer earning \$40,000 thinks they would.

The explanation of  $\mu$  by individual variables and the stability of these explanations over samples (see also Van Praag and Van der Sar (1988)) may be seen as indirect evidence for the validity of the Welfare Function of Income. The measured concept may be explained to a certain extent by individual circumstances in a plausible way. One of the more recent additions is a quadratic part in age.

It is seen that financial needs are greatest at the age of about 40.

However, the explanation of  $\mu$  may be useful for *policy* purposes as well. If we find that the welfare derived from income depends on family size, this gives a natural clue to the question what *family allowance* would keep the family at the same household level if family size is increased from two to three by having a child: the welfare a household derives from income is

$$U_{ind} = \Lambda(y_{i,c}; \text{constant} + 0.1 \ln fs_i + 0.6 \ln y_{i,c}, \sigma)$$

In order to keep welfare constant if fs increases from 2 to 3, we should add  $0.1 \ln(\frac{3}{2})$  to  $\ln y_{i,c}$  or multiply  $y_{i,c}$  by  $(\frac{3}{2})^{0.1}$

However, in the long run, this increase will not be enough to compensate the family for an increase in family size, as current income  $y_{i,c}$  increases and hence  $\mu$ . Therefore, we need a second increase of  $0.1 \ln(\frac{3}{2}) * 0.6$  and so forth. The total increase necessary to compensate the household equals

$$0.1 \ln(\frac{3}{2}) [1 + 0.6 + 0.6^2 + ..] = \frac{0.1 \ln(\frac{3}{2})}{1 - 0.6}$$

and this is precisely what the social standard welfare function would prescribe. Here we encounter a dynamic aspect, viz., that the individual welfare function is *anchored* on own current income. That is the meaning of preference drift. People adapt their norms to the present situation.

Parts of this analysis are also possible on the separate  $c_i$  levels without any reference to a cardinal utility function (see Van Praag, Van der Sar (1988)).

## 5. Dynamics

In the previous section, we described how the need parameter  $\mu$  could be explained by variables such as family size and current income  $y_c$ . The latter effect is now refined by supposing that  $\mu$  depends not only on *present* income but also on income in the past and income that is anticipated in the future. It follows that in the  $\mu$ -equation we replace  $y_c$  by  $..., y_{-2}, y_{-1}, y_0, \hat{y}_1, \hat{y}_2, ...$  whereby  $y_0$  denotes current income,  $y_{-2}, y_{-1}$ , for incomes one or two years in the past and where  $\hat{y}_1$  stands for anticipated future income in one year's time. All experienced and anticipated incomes contribute to the formation of our present norm on incomes. In its simplest form, the  $\mu$ -equation looks like

$$\mu_i = \beta_0 + \beta_1 \ln fs_i + \beta_2 \left( \sum_{t=-\infty}^{+\infty} w_t \ln y_{i,t} \right)$$



where  $i$  refers to respondent  $i$ .

The coefficients  $w_{-2}, w_{-1}, w_0, w_1, w_2, \dots$  are weights which add up to one, whereby the weight  $w_0$  denotes the weight of the present income, while  $w_p = \sum_{t=-\infty}^{-1} w_t$  and  $w_f = \sum_{t=1}^{\infty} w_t$  denote the weight of all past incomes and anticipated future incomes respectively. Van Praag and Van Weeren (1983,1988) estimated the parameters of this model on Dutch panel data. The main question concerns how the distribution of time weights will look. They regressed  $\mu_i$  on the incomes of the three years in which the panel was held. For the second wave they found

$$\begin{aligned}\mu_i &= 3.04 + 0.10 \ln f s_i + 0.68(0.16 \ln y_{i,t-1} + 0.75 \ln y_{i,t} + 0.09 \ln y_{i,t+1}) \\ \bar{R}^2 &= 0.69 \\ N &= 645\end{aligned}$$

where all coefficients are significant. The results tell us that current income has the greatest time-weight, which implies that the time-weight distribution peaks near the present. Also, incomes in the past carry more time weight than incomes in the future, which suggests that on aggregate the time-weight distribution peaks just before the present. Of course, this is an aggregate relationship which will differ for individuals of different ages and education profiles. For a more complete analysis, more incomes than the three available were needed. Therefore, Van Praag and Van Weeren (1988) used econometric techniques to *estimate* the incomes which were further back than one year, i.e.  $\dots, y_{-3}, y_{-2}$ . They also estimated incomes further than one year in the future, i.e.  $y_2, y_3, \dots$ . With the use of this complete income stream, they looked somewhat further at the shape of the time-weight distribution.<sup>8</sup> In general, they found the time-weight distribution to have the shape of a normal curve. More specifically, the time-weight distribution may be characterized by a mode parameter,  $\mu_\tau$ , and a dispersion parameter,  $\sigma_\tau$ . The empirically estimated shapes of the time weight distribution are presented in Figure 5.1 for three age brackets, viz., at 30, 50 and 70.

The most interesting points are that:

- The time weight distribution varies for different ages.
- The distribution is not symmetric around the present.

---

<sup>8</sup>The likely result of using estimates for some incomes is that the effect of income different from the present income will be underestimated. The qualitative results should, however, remain the same.

Figure 5.1:

- The time weights of the past are greatest for young and old people.
- The middle age bracket derives its norm mostly from the present and the anticipated future.
- The dispersion of the distribution varies considerably over different ages. In midlife, the time weights are extremely concentrated.

The mode and symmetry point of the time weight distribution is at  $\mu_\tau$ . We call it the *time focus* of the individual. It shifts from more than one year in the past, (-1.3) at 20, to almost half a year in the future, (0.45) at 50, while it shifts back to the past, (-0.43) for the age of 70.

The change of  $\sigma_\tau$  is also interesting. We call  $\sigma_\tau$  the *time span* of the individual. It is rather long for young individuals and shortens when people approach midlife. The time span is intimately related to the *velocity of time* as it is perceived by the individual. The midlife has a narrow time horizon which implies that the individuals then live "by the day". The velocity of life is high. For young and old people, the time horizon is wider and hence the velocity of time is lower. We call the reciprocal of  $\sigma_\tau$ , i.e.  $\frac{1}{\sigma_\tau}$  the *subjective velocity of time*.

In Table 3, we present the relevant figures for several age classes. It is seen that the subjective velocity of time  $\frac{1}{\sigma_\tau}$  increases by a factor  $\frac{1.44}{0.09} \approx 15$  up to midlife,

and then falls by a factor 6 at age 70, and still more at later ages.

Table 3: Values of  $\mu_\tau, \sigma_\tau, w_P, w_0, w_F$

age	$\mu_\tau$	$\sigma_\tau$	$w_P$	$w_0$	$w_F$
20	-1.32	1.44	0.72	0.18	0.10
30	-0.32	0.71	0.40	0.48	0.12
40	0.27	0.26	0.00	0.81	0.19
50	0.45	0.09	0.00	0.70	0.30
60	0.22	0.21	0.00	0.91	0.09
70	-0.43	0.62	0.46	0.48	0.07

The time weight distribution is clearly important for individuals because it determines the speed of adaptation of the income norms of the individual when faced with changing circumstances. This may be the case for individuals who become jobless and then become dependent on social benefits. The adaptation process may be a reason to smooth the path of the income reduction over time, in order to smooth the decline in welfare.

Another potential application is to evaluate the impact of inflation and accelerating inflation on the income norms and the satisfaction level derived from income. These applications are discussed in greater detail in Van Praag and van Weeren (1988).

The research on the time weight distribution has not been repeated since 1988. Therefore, this must be seen as a first attempt, where the results have to be considered with care. It may be that other models would yield other results. This method of obtaining time weights is based on a simple household survey and is very cheap compared to experimental laboratory experiments.

The estimates of time weights are exclusively based on the analysis of income norms. The memory and anticipation weights from norms on other subjects, for instance on fashion, housing, ethics, may be determined by other variables and have different time weight distributions. There is a great need for more research in this area and cooperation with psychologists.

## 6. Methodological discussion

The approach outlined above for measuring individual norms on income has been expanded to other aspects such as *age* and *education* by Van Praag, Dubnoff and

Van der Sar (1988). More specifically, individuals were asked to connect age levels to subjective labels in the following Age Evaluation Question:

”When I think of other adults, I consider people to be

*young*, if they are younger than..... years old

*somewhat young*, if they are about..... years old

*middle aged*, if they are about..... years old

*somewhat old*, if they are about..... years old

*old*, if they are older than..... years old”

Similar to the analysis of the IEQ, it is possible to analyse the age norms of respondents, for example, by explaining the answers by means of regression analysis. In Van Praag et al. (1988) this is done level by level for the Boston data set. Let  $a_i$  ( $i=1..5$ ) be the respondent’s age levels, then they consider the equation

$$\ln a_i = \alpha_{0,i} + \alpha_{1,i} \ln age + \alpha_{2,i} \ln schooling + \alpha_{3,i} \ln fs + \alpha_{4,i} D_{gender}$$

where they assume that what is considered ”young” or ”old” depends on the age of the respondent, the number of years of schooling, the size of the family and the gender of the respondent. The results of the regressions are presented in Table 4.

From a statistical point of view, most coefficients are significant and follow a definite pattern. Our evaluation is that there is a strong systematic pattern which indicates that there is no confusion in connotation. The fraction of the variance explained, as measured by  $R^2$ , is poor in comparison to that of the IEQ, but is certainly not below standard for samples of micro-data of this size ( $\approx 500$ ). However, it implies that there are more individual factors which were not covered in the survey, which must be added to the systematic structure than in the case of income standards.

With respect to the interpretation of the coefficients, we make the following observations. The older the respondent is, the higher his age standards. It follows that if A is 10% older than B, he will have an age standard for ”younger” that

Table 4: Regression equations for the age standards (N=538)

	constant	age	education	family size	gender	R <sup>2</sup>
young	1.414 (0.270)	0.319 (0.043)	0.180 (0.067)	0.069 (0.026)	0.027 (0.030)	0.091
somewhat young	2.329 (0.183)	0.266 (0.029)	0.045 (0.045)	0.056 (0.018)	0.019 (0.020)	0.135
middle-aged	3.160 (0.115)	0.177 (0.018)	0.014 (0.028)	0.016 (0.011)	0.048 (0.013)	0.163
somewhat old	3.740 (0.095)	0.117 (0.015)	0.018 (0.023)	0.003 (0.009)	0.047 (0.011)	0.132
old	4.243 (0.099)	0.058 (0.016)	0.067 (0.025)	0.003 (0.010)	0.048 (0.011)	0.071

standard deviations in parentheses

Source: Van Praag et al. (1988)

is about 3% higher (0.319\*10%). Or in other words, if someone aged 20 finds himself "somewhat young", an older person will find him still "young". For the age standards of "old", there is much less divergence between respondents of different ages.

We see that schooling has a strong impact on the definition of "young"; people tend to stay "young" longer. The impact of a large family on age standards is also evident. In such families, youngsters are considered to be children longer.

The social implications of these tendencies are not imminent. However, culturally it might be of interest that, in western countries, where the level of education has been increasing for decades, the concept of adulthood has become identified with an increase in age. Finally, the gender of the respondent also plays a role. If the respondent is female, the age standards are somewhat higher than for males, which implies that females tend to diminish the impact of age slightly. Except for the female tendency to stay and look young as long as possible, it is in conformity with the longevity of women compared to men.

Again we see that "young" does not mean the same thing to young people as to old people. We can derive a *general age*  $a_i^*$  *standard* by setting  $a_i^* = \text{age}$ , yielding

$$\ln a_i^* = \frac{1}{1 - \alpha_{1,i}} [\alpha_{0,i} + \alpha_{2,i} \ln \text{schooling} + \alpha_{3,i} \ln fs + \alpha_{4,i} D_{\text{gender}}]$$

The resulting age standards are tabulated in Table 5.

Similarly to the age evaluation question, individuals were asked which education level they thought was "very educated", "uneducated", etc. By explaining the answers to this Education Evaluation Question (EEQ), Van der Sar (1991)

Table 5: General age standards

General standards		
	male respondents	female respondents
young	17.69	18.41
somewhat young	30.16	30.95
middle-aged	49.54	52.50
somewhat old	65.73	69.31
old	75.06	78.91

Source: Van Praag et al. (1988)

was able to measure an individual norm on education as well. The interested reader is referred to Van der Sar (1991) for a full discussion of the AEQ, EEQ, and related questions.

### 6.1. Individual norms and general standards

The evidence described above suggests that people have *subjective norms* concerning various concepts. These norms will differ among individuals. They are measured by questions such as the IEQ, AEQ and EEQ, which supply us with numerical levels related to verbal labels or other symbols.

These questions may be posed theoretically in two ways: one may supply the label as stimulus and ask for an amount as a reply. Alternatively, one may supply an amount as stimulus and ask a label as a reply. The first way has been selected as the most practical when there are many different respondents with differing norms. It is also somewhat more informative, as people can space their answers.

In addition, we have *evaluations* by individuals of their *own situation*. This is done by fitting their own situation on their own norm. For instance an individual  $i$  with current income  $y_c$  evaluates his own income by  $U_i(y_c; y_c, fs)$ .

A final point is whether we may in some sense speak of general or social objective standards in contrast to individual subjective norms. Each individual may have an idea about what he thinks is a "bad" or a "good" income, but is there also a way to give content to a social norm with respect to what is a "good" income and what is a "bad" income? This question is especially pertinent when we think of a socially acceptable definition of poverty, or eligibility for social assistance. A second example is the general standard for being "old", which is relevant for fixing the retirement age. A general standard may be derived from the individual standards by calculating the income level, age, education, etc., where

people evaluate their *own* income, age, etc., as "bad", "good" etc. or "young" and "old".

Political applications of the IEQ are not extensively dealt with in this chapter, but we mention applications to poverty measurement (Goedhart et al. (1977), Van Praag et al. (1982), Hageaars (1986), Plug, Krause, Wagner, Van Praag (1997), Van Praag and Flik (1992), Frijters and Van Praag (1996), Colasanto, Kapteyn and Van der Gaag (1983), Stanovnik (1992)). Further applications concern income inequality (Van Praag (1977)), household equivalence scales (Kapteyn and Van Praag (1976), Van Praag and Warneer (1997)), and climate equivalence scales (Van Praag (1988), Frijters and Van Praag (1997)).

In this method, there is a strong anchor effect. The answer of the respondent depends very strongly on his own situation. One may attempt to avoid this, for example, by asking

"Thinking about an *average* family with two children, what does it need per month for an adequate living?" (cf. Rainwater (1971))

While avoiding the anchor effect of one's own situation, it introduces a new problem: what should be regarded as an *average* family, which will depend on the reference weighting system of the respondent. We can at least deal with the anchor effect of the individual's situation, as we know the own situation of the respondent, but we do not know what the respondent considers to be an average family. The usefulness of this question thus depends on whether it is reasonable to assume that there is common agreement about what constitutes an average family. In heterogeneous populations, such agreement will be absent.

Obviously, the method works only to evaluate one-dimensional situations where numbers may be assigned and where a natural ordering is manifest.

A problem where the described IEQ method breaks down is when the society is only partly monetized. In that case, welfare cannot be characterised on the one-dimensional income scale. An ingenious way out has been suggested by Pradhan and Ravallion (1997). Their approach is to ask for evaluations of consumption levels instead of evaluations of income levels.

At present, welfare functions have been measured in nearly all EC countries, the USA, Hungary, Slovenia, Poland and Russia. In almost all cases, except in the USA, the samples were fairly large scale, ranging from 1000 respondents to over 20,000. Panel data are scarce; the Dutch Socio-Economic Panel carried the question for a number of years, while at present, a Russian large-scale household panel includes the question as well (see Frijters and Van Praag (1996)).

## 7. Future directions: well-being *and* welfare

Traditionally, economists identify welfare (or even happiness) with income. However, it is well known and also fully recognised by other disciplines that there is more between heaven and earth than income and everything that can be bought with income.

This calls for an operational distinction between economic *welfare* and *well-being*. Welfare is the evaluation assigned by the individual to income or, more generally, to the contribution to our well-being from those goods and services that we can buy with money.

Next to material resources, we have other aspects which determine the quality of our life. We can think of our health, the relationship with our partner and family and friends, the quality of our work (job satisfaction), our political freedom, our physical environment, etc. We shall call this comprehensive concept *well-being* or quality of life (see Nussbaum and Sen (eds., 1992) for philosophical discussions about this concept).

It is empirically possible for most individuals to evaluate their life as a whole. A well-known example is the following question devised by Cantril (1965):

”Here is a ladder with ten steps which denotes the ”ladder of life”. The bottom step stands for the worst possible life. If you climb up and arrive at the tenth step, you arrive at the best possible life. Can you indicate where you are at the moment?”

Other questions which are very similar to Cantril’s question ask individuals to denote how *satisfied* or how *happy* they are with their life as a whole. The concept of well-being is thus very similar to that of life-satisfaction or happiness, and we will not discuss the differences.

These questions are a standard module in many psycho-sociological surveys and respondents have no difficulty responding. See Veenhoven (1996) and Diener and Suh (1997) for reviews of the psychological literature on well-being. It is also obvious that responding to these questions is tantamount to evaluating one’s life situation on a bounded numerical scale between zero and ten.

In fact, we have here a measurement method which *defines* the well-being concept in an operational way. We notice that what we measure is an evaluation of the individual’s actual situation. Hence it is not an ”individual norm”, as measured by the IEQ, where six qualitative labels are linked to income levels, yielding



an "income norm". The Cantril question provides us with a *social standard* on *well-being*.

We assume again, as is always done implicitly, that the respondent's answers are comparable, in the sense that individuals evaluating their life with the same grade, such as a 5 or an 8, are equally unhappy or happy with their life. The main questions are:

- a. What determines well-being?
- b. What are the differences between welfare and well-being?

Plug and Van Praag (1995), Plug (1997), analysed these two questions on a large sample (1991) of about 6000 Dutch married couples, with the husband younger than 65 years of age. They hypothesised that well-being, to be denoted by  $W$ , depends on various contributing factors and determinants. Some of these factors are objective ones, like family size, income, age, and religion. Other variables are called "problem intensities". They relate to the intensity with which an individual "has problems" with his health, job, marriage, physical environment, etc.<sup>9</sup> Formally we write

$$W = W(P; z)$$

where  $P$  stands for a vector of problem areas and  $z$  for a number of objective variables. Problem intensities, such as  $P_{health}$ ,  $P_{job}$ , ..., are operationalized by questions asking individuals "many/few/no problem" questions. An example of such a question is:

"Have you had problems in the last three months with your health?  
No / a little / some / serious?"

The outcomes are on a numerical scale<sup>10</sup>. Plug and Van Praag (1995) found

---

<sup>9</sup>In the article, Plug and Van Praag name the extent to which individuals are free of a problem, a "partial satisfaction". However, given that this term may be confusing, we use here the term problem intensities.

<sup>10</sup>For estimation purposes, they prefer to transform  $W$  and  $P$  from their bounded scale into  $(-\infty, +\infty)$ . It might be feared that people will center in the middle and that extreme answers will be rare. To solve both problems, the *empirical* distribution functions  $\tilde{F}$  of the  $W$  and  $P$  values are calculated, and the value  $\tilde{F}$  is assigned to the various levels instead of the original ones. Then they transform  $\tilde{F}$  again by taking the inverse standard normal, which means that instead of  $W$  and  $P$   $\tilde{W} = N^{-1}(\tilde{F}_w(W); 0, 1)$  and  $\tilde{P} = N^{-1}(\tilde{F}_p(P); 0, 1)$  are utilized. The transformations do no intrinsic harm, but they are only used to get more response differentiation and a stretching on  $(-\infty, +\infty)$ . From now on we will drop the tildes.

the estimates presented in Table 6.

The first column refers to the explanation of well-being, while the second column refers to the explanation of  $\mu$  by the *same* variables. The first nine variables stand for "problems with health", "problems with partner", "with the job", "with sleep", "with alcohol and drugs", "with the family", "sexual problems", "problems with parents" and the evaluation of the neighborhood. These variables reach their highest value when there are no problems. "Religion" stands for the intensity of religious feelings where the highest value corresponds to "non-religious". The IEQ measures welfare, while the Cantril question measures the broader concept of life satisfaction or well-being.

The main difference between welfare and well-being is that "problem" variables hardly affect the evaluation of income but do, however, affect well-being.

The second question is also answered at the same time. Welfare and well-being are different concepts, where welfare is only dependent on a small subset of the set of variables which influence well-being. The size and sign of the effect is also different.

To illustrate the relevance of the results, Plug and Van Praag estimated the optimum number of children, as family size appears quadratically in  $W$ . This optimum number of children depends on such factors as income. Table 7 was derived for the Netherlands. It shows that the optimum number of children is zero for a family with an annual after-tax income of NLG 20,418. For a family with an annual income of NLG 51,451 two children are the optimum family size.

An especially promising path is the *combination* of more than one satisfaction measure. We will explain this idea with a specific example from Van Praag and Plug (1995).

We consider again the social standard function of income as derived in section 4, which we denote by  $\tilde{U}(y;fs)$  and where we ignore other variables. We know already that  $\tilde{U}$  decreases with the number of children. More specifically, it is possible to assess the monetary value of the "welfare cost of additional children". Assuming that a representative couple has two children and an annual income of \$50,000, their welfare will be  $\tilde{U}(50;4)$ . Assume now that the couple has another child, which causes  $\tilde{U}$  to change to  $\tilde{U}(50;5)$ . The welfare decline may be compensated by an income increase  $\Delta^U y$  such that

$$\tilde{U}(50 + \Delta^U y; 5) = \tilde{U}(50; 4)$$

We call  $\Delta^U y$  the *shadow price* of the additional child, which is the monetary

amount needed to keep a household on the same welfare level<sup>11</sup>. Notice that this depends on the rank order of the child and that the shadow price will depend on income.

On the other hand, we have the well-being measure  $W$ , based on the Cantril question yielding a well-being function

$$W(y, fs)$$

again ignoring all other variables. Given our estimates of the previous table,  $W$  is strongly quadratic in  $fs$ , implying that well-being initially increases with family size up to a certain point, whereafter it falls with an increasing family size. This *non-monotonic* behavior points to the fact that an extra child may be wanted for its non-economic contribution. It is a gain for well-being while at the same time a loss in terms of welfare. Hence there are *non-monetary* benefits of having children and monetary costs. The  $W$  function captures both and increases in family size if the benefits outweigh the costs, and decreases if the costs outweigh the benefits. There is equality at the optimum family size.

Considering  $W(y, fs)$  we may calculate the *shadow well-being* price of an additional child  $\Delta^W y$ , by solving

$$W(y, fs) = W(y + \Delta^W y, fs + \Delta fs)$$

Here  $\Delta^W y$  is the monetary counter value of this difference:

$$\Delta^W y = \text{Benefits} - \text{Costs}$$

We call it the shadow price with respect to well-being. It is positive if we welcome a child and negative if the opposite holds. From the calculation on  $U$ , we obtained an estimate of the cost

$$\Delta^U y = \text{Costs}$$

Addition yields  $\Delta^W y + \Delta^U y =$  monetary value of benefits.

The benefits may be considerable, as is witnessed by the fact that childless couples try to adopt children or are willing to undergo expensive medical treatment. From Van Praag and Plug, we show the following Table 8, where the benefits of children are calculated for Dutch families in 1991.

---

<sup>11</sup>The "cost of children" will at least include the expenditures on an additional child. Future research looks at whether it also includes the monetary shadow-value of time spent on a child.

We see that the value of the (non-monetary) benefits of the first child is negative at a low income level. The benefit of the first child becomes positive for incomes above NLG 20,000. For the second child, the benefits remain negative until an income of about NLG 40,000. For the third and fourth child, the benefits remain negative for even longer. With respect to costs, there is no ambiguity; costs are always positive. However, the cost of each additional child decreases. As we can see, these non-monetary benefits are substantial and increase strongly with income.

At this stage, we warn that the study is in the beginning stage and that not too much value should be assigned to this or other results without replication. However, the path seems promising. A future step is to estimate the substitution and trade-off between variables and the calculation of monetary values of health increases, family increases, education, marriage quality, etc.

See Frijters (1998) for other applications and extensions of Leyden methodology on welfare and satisfaction.

## 8. Conclusions

The work originating from the Leyden School has tried to operationalize the concepts of welfare, well-being, etc., which are considered immeasurable and esoteric by most of the economic profession. With rather simple and inexpensive questions in large-scale surveys, considerable information has been found on feelings. At least the feelings of welfare and well-being may be "explained" by objectively measurable variables and by partial satisfaction measures with respect to aspects of life. The information is helpful for quantifying memory and anticipation weights.<sup>12</sup> The potential policy applications are plentiful. We briefly described its use to calculate family equivalence scales. A rather recent development is the combination of the welfare and well-being measurement, which makes it possible to identify the cost and benefits of various choices. We demonstrated this for the option of choosing to have children.

The apparatus developed thus far is not typically restricted to economic problems, but can also be used by psychologists, sociologists, and political scientists. Its use in health economics seems straightforward.

---

<sup>12</sup>In Van Praag (1981), Kapteyn (1977), Van der Sar (1991), Van de Stadt et al. (1985), Kapteyn et al. (1976), the IEQ was also used for the extraction of information on the social reference mechanism, as the answers to the IEQ are influenced by social reference groups. This application was not dealt with in this chapter.

The story is hopefully not finished but only in the early stages. The main empirical restriction is that the data sets are scattered and almost never contain the IEQ, sound economic information (consumption, income, job characteristics) and at the same time "soft" information on feelings on several aspects of life, such as the Cantril question. In this respect, the USA, where so much effort is given to research, is conspicuously absent.

Keywords: measurement of welfare, well-being, social filter, interpersonal comparison, social reference group, individual welfare functions, utility.

## Literature

1. Allais, M., Hagen, O. (1979), *Expected utility hypothesis and the Allais paradox*, Reidel, Dordrecht; the Netherlands
2. Arrow, K. (1951), *Social choice and individual values*, New York: Wiley & Son.
3. Atkinson, A.B. (1970), "On the measurement of inequality", *Journal of Economic Theory*, 2, pp. 244-263.
4. Brickman, P., Cambell, D.T. (1971), "Hedonic relativism and planning the good society", in Apley, M.H. (ed.), *Adaptation-level theory: a symposium*, New York: Academic Press.
5. Callan, T., Nolan, B. (1991), "Concepts of poverty and the poverty line", *Journal of Economic Surveys*, no. 5, pp. 243-261.
6. Cantril, H. (1965), *The pattern of human concerns*, New Brunswick, Rutgers University Press.
7. Christensen, L.R., Jorgenson, D.W., Lau, L.J. (1975), "Trancendental logarithmic utility functions", *American Economic Review*, pp. 367-383.
8. Clark, A.E. (1996), "Job satisfaction in Britain", *British Journal of Industrial Relations*, 34, 189-217.
9. Clark, A.E., Oswald, A.J. (1994), "Unhappiness and unemployment", *Economic Journal*, pp. 648-659.

10. Clark, A.E., Oswald, A.J. (1996), "Satisfaction and comparison income", *Journal of Public Economics*, 61, pp. 359-381.
11. Colosanto, D., Kapteyn, A., Van der Gaag, J. (1983), "Two subjective definitions of poverty: results from the Wisconsin basic needs study", *Journal of Human Resources*, 28, pp. 127-138.
12. Debreu, G. (1959), *Theory of value*. New Haven: Yale University Press.
13. Diener, E., Suh, E. (1997), "Measuring the quality of life: economic, social and subjective indicators", *Social Indicators Research*, forthcoming.
14. Dow, G.K., Juster, F.T. (1985), "Goods, time and well-being: the joint dependency problem", in F.T. Juster and F.P. Stafford (eds.), *Time, goods and well-being*. Ann Arbor, Michigan: Institute of Social Research.
15. Dubnoff, D., Vaughan, D., Lancaster, C. (1981), "Income satisfaction measures in equivalence scale applications", *Proceedings of the social statistics section, American Statistical Association*, 348-352.
16. Easterlin, R. (1974), "Does economic growth improve the human lot? Some empirical evidence" in David, P. & Reeder, R., eds., *Nations and households in economic growth: essays in honor of Moses Abramovitz*, New York, Academic Press.
17. Easterlin, R. (1995), "Will raising the incomes of all increase the happiness of all?", *Journal of Economic behavior and organization*, 27, pp. 35-47.
18. Edgeworth, F.Y. (1881), *Mathematical Psychics*, London: Kegan Paul.
19. Frijters, P. (1998), *Explorations of welfare and satisfaction*, Ph.D. thesis, forthcoming.
20. Frijters, P., Van Praag, B.M.S. (1995), "Estimates of poverty ratios and equivalence scales for Russia and parts of the former USSR", *Tinbergen Discussion Papers* 95-149.
21. Frijters, P., Van Praag, B.M.S. (1997), "Climate equivalence scales and the effect of climate change on Russian welfare and well-being", forthcoming in *Climate Change*.
22. Frisch, R. (1932), *New methods of measuring marginal utility*, Tübingen, Mohr.
23. Gershuny, J., Halpin, B. (1995), "Time use, quality of life, and process benefit", in Offer, A. (ed), *In pursuit of the quality of life*, pp. 188-210, Oxford: Calrendon Press.

24. Goedhart, T., Halberstadt, V., Kapteyn, A., Van Praag, B.M.S. (1977), "The poverty line: concepts and measurement", *Journal of Human Resources*, no. 12, pp. 503-520.
25. Hagenaars, A.J.M., (1986), *The perception of poverty*, North-Holland, Amsterdam.
26. Heywood, J.S., Siebert, W.S, Wei, X. (1997), "Are union jobs worse? Are government jobs better?", mimeo.
27. Hirschberg, J.G., Maasoumi, A., Slottje, D.J. (1991), "Cluster analysis for measuring welfare and quality of life across countries", *Journal of Econometrics*, 50, pp. 131-150.
28. Houthakker, H.S. (1950), "Revealed preference and the utility function", *Economica*, 17, pp. 159-174.
29. Jorgenson, D.W., Slesnick, D.T. (1984), "Aggregate consumer behavior and the measurement of inequality", *Review of Economic Studies*, 166, pp. 369-392.
30. Kahneman, D, Knetsch, J., Thaler, R. (1991), "The endowment effect, loss aversion, and status quo bias", *Journal of Economic Perspectives*, 5, 193-206.
31. Kahneman, D, Tversky, A. (1979), "Prospect theory: an analysis of decision under risk", *Econometrica*, 47, pp. 263-291.
32. Kahneman, D, Wakker, P.P., Sarin, R. (1997), "Back to Bentham? Explorations of experienced utility", *Quarterly Journal of Economics*, May, pp. 375-405.
33. Kapteyn, A. (1977), *A theory of preference formation*, Ph.D. Leyden University, Leyden.
34. Kapteyn, A. (1994), "The measurement of household cost functions: revealed preference versus subjective measures", *Journal of Population Economics*, 7, pp. 333-350.
35. Kapteyn, A., Van Praag, B.M.S. (1976), "A new approach to the construction of equivalence scales", *European Economic Review*, 7, pp. 313-335.
36. Kapteyn, A., Van Praag, B.M.S., Van Heerwaarden, F.G. (1976), "Individual welfare functions and social reference spaces", Report 76.01, Economic Institute, *Leyden University*
37. Kurian, G.T. (1984), *The new book of world rankings*, New York, NY.
38. Levy-Garboua, L., Montmarquette, L.C. (1997), "Reported job-satisfaction: what does it mean?", Cahier de Recherche 1, *University of Paris*.

39. Maasuomi, E. (1989), "Composite indices of income and other developmental indicators: a general approach", *Research on Economic Inequality*, 1, pp. 269-286.
40. Markowitz, H.M. (1952), "The utility of wealth", *Journal of Political Economy*, 60, pp. 51-58
41. Ng, Y.K (1997), "A case for happiness, cardinalism, and interpersonal comparability", *Economic Journal*, forthcoming.
42. Nussbaum, M., Sen, A.K. (1992, eds), *The quality of life*, Oxford: Clarendon Press.
43. Parducci, A. (1995), *Happiness, pleasure and judgment, the contextual theory and its applications*, Erlbaum Associates, Mahwah, New York.
44. Pareto, V. (1904), *Manuel d'economic politique*, Girard, Paris.
45. Pigou, A.C. (1948), *The economics of welfare*, fourth edition, London: Macmillan.
46. Plug, E.J.S. (1997), *Leyden welfare and beyond*, Ph.D. thesis, Thesis Publishers, Amsterdam.
47. Plug, E.J.S., Van Praag, B.M.S. (1995), "Family equivalence scales within a narrow and broad welfare context", *Journal of Income Distribution*, 4, pp. 171-186.
48. Plug, E.J.S., Krausse, P., Van Praag, B.M.S., Wagner, G.G. (1996), "The measurement of poverty; Exemplified by the German case", in Ott, N., Wagner, G.G. (eds.), *Income inequality and poverty in Eastern and Western Europe*. Springer Verlag.
49. Pratt, J.W. (1964), "Risk aversion in the small and in the large", *Econometrica*, 32, pp.122-126.
50. Rainwater, L. (1971), "Interim report on explorations of social status, living standards and family life styles" (unpublished paper), Cambridge MA: Joint Center for Urban Studies of the Massachusetts Institute of Technology and Harvard University.
51. Ravallion, M. (1994), *Poverty comparisons*. Fundamentals in pure and applied economics, 56, Chur, Switzerland: Harwood Academic Press.
52. Robbins, K. (1932), *An essay on the nature and significance of economic science*, London: Macmillan.



53. Samuelson, P.A. (1945), *Foundations of Economic Analysis*, Harvard University Press, Cambridge MA.
54. Savage, L.J. (1954), *The foundations of statistics*, New York: J. Wiley and Sons.
55. Seidl, C. (1994), "How sensible is the Leyden individual welfare function of income?", *European Economic Review*, 38, pp. 1633-1659.
56. Sen, A.K. (1987), *The standard of living*, Cambridge University Press, The Tanner Lectures, pg. 49.
57. Stanovnik, T. (1992), "Perception of poverty and income satisfaction", *Journal of Economic Psychology*, 13, pp. 57-69.
58. Tinbergen, J. (1991), "On the measurement of welfare", *Journal of Econometrics*, 50, pp 7-13.
59. Townsend, P. (1979), *Poverty in the United Kingdom*, Harmondsworth, Penguin.
60. Townsend, P. (1993), *The analysis of poverty*, Harvester/Wheatsheaf.
61. Van Herwaarden, F.G, Kapteyn, A. (1981), "Empirical comparisons of the shape of welfare functions", *European economic Review*, pp. 261-286, nr. 15.
62. Van der Sar, N.L. (1991), *Applied Utility Analysis*, Ph.D. Thesis, Alblasterdam: Haveka
63. Van de Stadt, H., Kapteyn, A., Van de Geer, S. (1985), "The relativity of utility: evidence from panel data", *The Review of Economics and Statistics*, 67, pp. 179-187
64. Van Herwaarden , F.G., Kapteyn, A. (1981), "Empirical comparison of the shape of welfare functions", *European Economic Review*, 15, pp. 261-286.
65. Van Praag, B.M.S. (1968), *Individual welfare functions and consumer behavior. A theory of rational irrationality*, Ph.D.-thesis, North-Holland Publishing Company, Amsterdam.
66. Van Praag, B.M.S. (1971), "The welfare function of income in Belgium: an empirical investigation", *European Economic Review*, nr. 2, pp. 337-369.
67. Van Praag, B.M.S. (1977), "The perception of welfare inequality", *European Economic Review*, no. 10, pp. 189-207.

68. 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*.
69. Van Praag, B.M.S. (1988), "Climate equivalence scales: an application of a general method", *European Economic Review*, 32 (4), pp. 1019-1024.
70. Van Praag, B.M.S. (1991), "Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept", *Journal of Econometrics*, 50, pp. 69-89.
71. Van Praag, B.M.S. (1994), "Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept", in Blundell, R., Preston, I. Walker, I. (eds.), *The measurement of household welfare*, Cambridge University Press, pp. 86-110.
72. Van Praag, B.M.S., Dubnoff, S., Van der Sar, N.L. (1988), "On the measurement and explanation of standards with respect to income, age and education", *Journal of Economic Psychology*, 9, pp. 481-498.
73. Van Praag, B.M.S., Flik, R.J. (1992), "Poverty lines and equivalence scales. A theoretical and empirical investigation", *Poverty measurement for economies in transition in Eastern Europe*, International Scientific Conference, Warsaw, 7-9 October, Polish Statistical Association, Central Statistical Office.
74. Van Praag, B.M.S., Hagenaars, A., Van Weeren, J. (1982), "Poverty in Europe", *Review of Income and Wealth*, 28, pp. 345-359.
75. Van Praag, B.M.S., Kapteyn, A., (1973), "Further evidence on the individual welfare function of income: an empirical investigation in the Netherlands", *European Economic Review*, nr. 4, pp. 33-62.
76. Van Praag, B.M.S., Kapteyn, A., (1994) "How sensible is the Leyden individual welfare function of income? A reply", *European Economic Review*, 38, pp. 1817-1825.
77. Van Praag, B.M.S., Kapteyn, A., Herwaarden, F.G. van (1979), "The definition and measurement of social reference spaces", *The Netherlands Journal of Sociology*, 15, pp. 13-25.
78. Van Praag, B.M.S., Plug, E.J.S. (1995), "New developments in the measurement of welfare and well-being", *Tinbergen Discussion Paper 95-60*, University of Amsterdam.

79. Van Praag, B.M.S., Van der Sar, N.L. (1988), "Household cost functions and equivalence scales", *Journal of Human Resources*, 23, pp. 193-210.
80. Van Praag, B.M.S., Van Weeren, J. (1983), "Some panel-data evidence on the time-discounting mechanism in the formation of value judgments on income with applications to social security and income policy", Report no. 83.22 of the *Center for Research in Public Economics*, Leyden University.
81. Van Praag, B.M.S., Van Weeren, J. (1988), "Memory and anticipation processes and their significance for social security and income inequality", in Maital S (ed.), *Applied behavioral Economics*, Brighton: Wheatsheaf Books.
82. Van Praag, B.M.S., Warnaar, M. (1997), "The cost of children and the effect of demographic variables on consumer demand", in Rosenzweig, M.R., Stark, O. (eds.), *The Handbook of Population and Family Economics*. Amsterdam: North Holland.
83. Veenhoven, R. (1996), "Happy life-expectancy: a comprehensive measure of quality of life in nations", *Social Indicators Research*, 39, pp. 1-58.
84. Wakker, P.P., Tversky, A. (1993), "An axiomatization of cumulative prospect theory", *Journal of Risk and Uncertainty*, 7, 147-76.
85. Wansbeek, T., Kapteyn, A. (1983), "Tackling hard questions by means of soft methods: the use of individual welfare functions in socio-economic policy", *Kyklos*, 36, pp. 249-269.

Table 6: Estimation results of  $w$  and  $\mu$ 

	$w$	$\mu$
Health	0.08 (11.43)	-0.00 (-1.11)
Partner	0.04 (3.62)	-0.01 (-1.13)
Job	0.07 (9.57)	-0.01 (-1.67)
Sleep	0.07 (8.90)	0.00 (0.55)
Alcohol/Drugs	0.04 (4.27)	-0.01 (-1.22)
Family	0.07 (7.92)	-0.01 (-2.61)
Sexuality	0.03 (3.50)	0.00 (0.86)
Parents	0.05 (6.53)	-0.01 (-2.16)
Neighbourhood	0.08 (13.61)	0.00 (0.26)
Religion	0.02 (4.00)	-0.01 (-3.54)
$\ln y$	0.12 (5.13)	0.55 (41.49)
$\ln fs$	-0.81 (-3.60)	-0.34 (-2.66)
$\ln y \ln fs$	0.09 (4.13)	0.03 (2.44)
$\ln^2 fs$	-0.06 (-3.36)	0.03 (2.49)
$\ln age$	-2.14 (-5.88)	1.10 (5.27)
$\ln^2 age$	0.30 (4.13)	-0.14 (-4.82)
Dummy-Job	-0.10 (-5.07)	0.01 (-0.57)
Constant	3.10 (4.72)	2.24 (5.93)
$R^2$	0.24	0.61

Source: Van Praag and Plug(1995)

Table 7: The optimum family size for specific income levels

fs=1	fs=2	fs=3	fs=4	fs=5	fs=6
8,103	20,418	35,060	51,451	69,280	88,346

note: family income is measured in Dutch guilders (2 guilders is about 1 dollar)

## Table 8

MONEY VALUE OF NON-MONETARY CHILD BENEFITS  
ONE BREADWINNER

Income	1 <sup>th</sup> child	2 <sup>th</sup> child	3 <sup>th</sup> child	4 <sup>th</sup> child
20,000	-262	-838	-1005	-1039
30,000	1114	-236	-748	-959
40,000	2911	651	-279	-713
50,000	5023	1749	348	-341
60,000	7383	3018	1100	130

TWO BREADWINNERS

Income	1 <sup>th</sup> child	2 <sup>th</sup> child	3 <sup>th</sup> child	4 <sup>th</sup> child
20,000	-726	-1153	-1240	-1223
30,000	419	-708	-1100	-1236
40,000	1983	22	-747	-1082
50,000	3871	964	-237	-802
60,000	5990	2074	399	-423