Journal of Economic Psychology 6 (1985) 333-363
North-Holland

THE INDIVIDUAL WELFARE FUNCTION

A Review *

Arie KAPTEYN
Tilburg University, The Netherlands

Tom WANSBEEK
University of Groningen, The Netherlands

Received October 31, 1983; accepted December 11, 1984

The Individual Welfare Function (IWF), introduced by Van Praag (1968), is a cardinal utility function. It can be measured by means of survey questions. Since its introduction, the IWF has been used extensively in both theoretical and empirical research. This research is reviewed, with an emphasis on policy applications.

1. Introduction

Since the days of Edgeworth, who allegedly considered utility to be 'as real as his morning jam' (Samuelson 1945: 206), utility theory has grown into a highly abstract field of research in economics. In the course of this development, the focus of attention has shifted from utility functions per se to the representation of the underlying preferences. Since, usually a preference ordering can be represented by more than just one utility function, the empirical status of the concept has eroded dramatically.

* The research reported here is part of the Leyden Income Evaluation Project. The paper was written while T. Wansbeek was with the Netherlands Central Bureau of Statistics (CBS); however, the views expressed in this paper are those of the authors and do not necessarily reflect the policies of the CBS.

The authors thank Peter Kooreman, Jan Tinbergen, Bernard van Praag and two anonymous referees for their helpful comments.

Mailing address: A. Kapteyn, Dept. of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

This has two consequences. The non-uniqueness of utility functions as representations of preference orderings complicates measurement in that measurement takes place *indirectly* via the revealed preference approach. Secondly, the assumed non-cardinality and non-interpersonal comparability of utility functions limits their applicability in policy matters. It is now generally accepted that for purposes of social policy a certain amount of interpersonal comparability is indispensable, cf., e.g., Sen (1974).

Presumably, the early writers were a little overconfident regarding the intersubjective nature and measurability of utility functions. But, one wonders whether modern economists' retreat to an almost agnostic position has really been forced upon them by outcomes of empirical observations, or whether to some extent the restraint is selfimposed. For example, in a static setting (under conditions of perfect information, etc.) indifference curves, and hence ordinal utility, are sufficient to explain choice behavior. This obviates the assumption of cardinal utility and, invoking Occam's razor, it has been dropped. In this process *empirical observations play no role*. In fact, one can easily think of empirical evidence that would support cardinality, like individuals being able to order utility differences. For example: an individual may be able to state that a move from one sandwich to two sandwiches adds more utility than a move from two to three. It is well-known that in such a case preferences have to be described by a cardinal utility function (e.g., Fishburn 1970: ch. 6).

If it is true that economists' retreat from measurable cardinal utility is not based on convincing empirical evidence, but on a desire to economize on assumptions, the question arises whether something has been lost by giving up the possibility of a cardinally measurable utility function. Are there empirical or theoretical problems that can not (or only with considerably more difficulty) be tackled without a cardinally measurable utility concept? If the answer to this question is in the affirmative we also have to face the question to what extent cardinality is a testable assumption.

In this article we consider these questions by reviewing results regarding a cardinal utility concept that has been subject to research over the last fifteen years or so. This utility function is the so-called *Individual Welfare Function* (IWF), which was introduced by Van Praag (1968). Since its inception, the concept has been applied and extended in various directions. An earlier review of results is given by Van Praag

(1976). Much of the work on the concept of the IWF has a direct bearing on economic theory and on socio-economic policy. In our review we will pay special attention to its potential for policy analysis.

Since both authors of this paper have devoted a fair amount of their time to research involving IWFs, the reader should be warned that our discussion will presumably be biased in its favor. Furthermore, the vast majority of papers reviewed here has been written by a small number of researchers. To contain the length of the review within reasonable bounds, we do not try to discuss in any detail related research by others – we trust that that has been done sufficiently in the papers referred to. In essence, our review is a plea to enhance the empirical status of utility theory, without necessarily going all the way back to the morning jam.

There are a number of outstanding features of the IWF that will be discussed consecutively:

(1) The IWF is a cardinal utility function, with function values ranging from zero (worst case) to one (best case). Under certain assumptions its functional form can be derived. In section 2 the functional form is discussed, whereas section 3 gives an interpretation of its parameters. There are different types of IWFs and these are also discussed.
(2) IWFs differ between individuals and can be measured by means of survey questions. The measurement method is introduced in section 4. Tests of cardinality and functional form are reported.
(3) There is a simple theory explaining why and how individual welfare functions differ between individuals. This 'theory of preference formation' is explained in section 5 and further discussed in section 6.

After this exposition of the basic features we discuss empirical evidence regarding the preference formation theory and mention various applications and implications in sections 7, 8 and 9. The applications include the construction of family equivalence scales and poverty lines and an analysis of the welfare effects of economic growth and income redistribution. Section 10 concludes.

2. The individual welfare function

The individual welfare function is the outcome of a rather elaborate theoretical structure. To keep our exposition simple, we only give a brief sketch of the underlying theory, omitting as many details as possible. For the latter, the reader is referred to Van Praag (1968, 1975).

A central role in the theory is played by the notion of a *commodity group*. A commodity group is a set of one or more commodities that, in any combination of quantities, can be represented by the same (finite or infinite) set of characteristics (cf. Gorman 1956, Lancaster 1971). The number of commodity groups that one can distinguish obviously depends on the way consumers define characteristics. A large part of Van Praag's theory is concerned with psychological assumptions on the way people define characteristics and how characteristics enter the utility function. The set-up chosen is one in which *utility theory becomes formally isomorphic with probability theory*. The isomorphism is used to apply the Central Limit Theorem from probability theory to utility theory. The Central Limit Theorem comes in when considering commodity groups that are *broad*, i.e., that are described by a large number of characteristics. The bare essentials of Van Praag's theory can be sketched as follows.

Consider a commodity group represented by n characteristics. The welfare that a consumer derives from a particular combination of characteristics $x_1,\dots,x_n$ is denoted as $W(x_1,\dots,x_n)$, i.e., W is a utility function defined on characteristics. Regarding the utility function W, Van Praag assumes that it has the same mathematical properties as a probability distribution function (it assumes values between zero and one, it is non-decreasing in each x_i, etc.) and that it has the following separable structure:

$$W(x_1,\dots,x_n)=W_1(x_1)\cdot W_2(x_2)\dots W_n(x_n), \tag{1}$$

where W_i represents the utility attached to x_i. Thus the welfare derived from the combination $(x_1,\dots,x_n)$ is the product of the amounts of welfare that each characteristic provides separately. In standard economic theory, W would be considered to be *ordinal*, so that for example W could be replaced by $\ln W$, because W and $\ln W$ imply the same *ordering* of combinations of characteristics. If we take logarithms on both sides of (1), the right hand side becomes a sum of the $\ln W_i$. In terms of conventional economic theory we would then say that the utility function W is *additive*, an assumption which is often made. In Van Praag's theory such a transformation is not permitted, i.e., he takes W to be a *cardinal* concept.

In order to be able to consume a certain combination of characteristics $x_1, \ldots, x_n$ the consumer has to spend money on the commodity groups. Conversely, given that he [1] is willing to spend a certain amount y on the commodity group, there are probably many different combinations of characteristics that he can acquire. Which combinations these are depends on the market. Or to put it in a slightly different perspective, which combinations the consumer believes to be obtainable by spending an amount y depends on his perception of the market of consumption goods. Van Praag assumes that the consumer believes that by spending an amount y, all combinations of characteristics $(x_1, \ldots, x_n)$ are available that satisfy

$$f_1(x_1) \cdot f_2(x_2) \ldots f_n(x_n) \leqslant y. \tag{2}$$

or equivalently,

$$\ln f_1(x_1) + \ln f_2(x_2) + \ldots + \ln f_n(x_n) \leqslant \ln y, \tag{2a}$$

where $f_1, f_2, \ldots, f_n$ are functions that are left unspecified, although they are assumed to be continuous and positive-valued. An example of inequality (2) is given in fig. 1, for the case $n = 2$. Given that he spends an amount y on the commodity group under consideration, the consumer believes that each point in the area under the curve is attainable. The form of the area depends of course on the mathematical form of the functions f_1 and f_2.

The area under the curve in fig. 1 bears a close resemblance to the concept of a *budget set* in economics. Usually, however, a budget set describes a collection of *commodities* that is available to the consumer, rather than *characteristics* as in fig. 1. Furthermore, the area in fig. 1 only represents the *perception* of the consumer of what he could acquire if he spent an amount y. There is no presumption that this perception is correct.

Next, Van Praag considers the following question. Given that the utility of characteristics is given by eq. (1) and that eq. (2) delimits the combinations of characteristics that can be obtained by spending y, what welfare will the consumer attach to y?

There are at least two plausible answers to this question. The first

[1] Throughout the paper, the terms 'he' and 'she' are used indiscriminately.

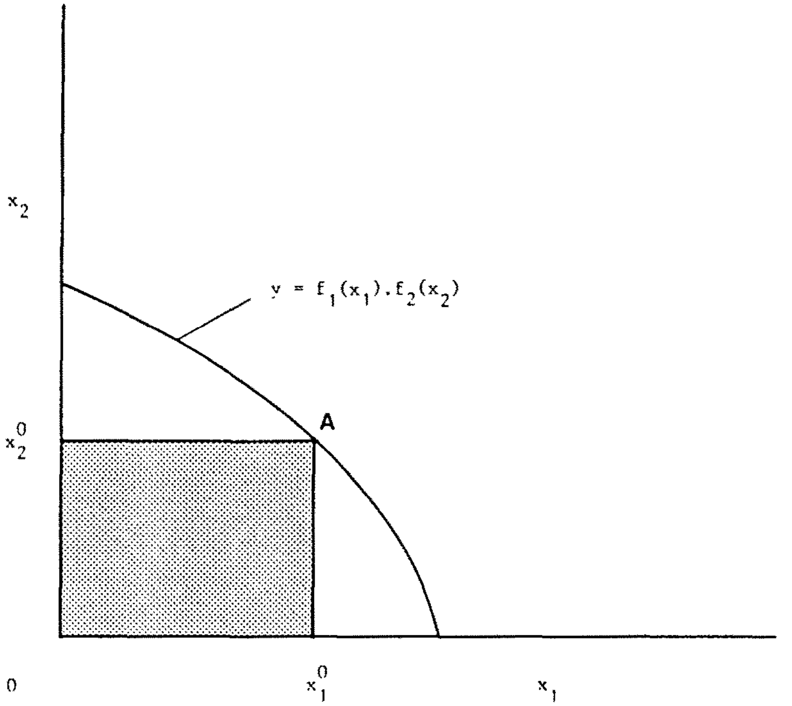

Fig. 1. The set of characteristics obtainable by spending an amount y on the commodity group.

one is a rational one. The consumer should decide which point in the area under the curve he prefers. Let us say that this is point A in fig. 1. According to his welfare function W, the evaluation (the 'utility') of the combination of characteristics (x_1^0, x_2^0) is equal to $W(x_1^0, x_2^0)$. Given that W is formally a probability distribution function we can write

$$W(x_1^0, x_2^0) = \int_0^{x_1^0}\int_0^{x_2^0} \mathrm{d}W(\zeta_1, \zeta_2), \tag{3}$$

where $\mathrm{d}W(\zeta_1, \zeta_2)$ is simply the (small) increase in welfare that results if we move from the point (ζ_1, ζ_2) to the point $(\zeta_1 + \mathrm{d}\zeta_1, \zeta_2 + \mathrm{d}\zeta_2)$. Formally, we can identify $\mathrm{d}W$ with a probability density function. Van Praag calls $\mathrm{d}W(\zeta_1, \zeta_2)$ the *welfare mass* concentrated in the point (ζ_1, ζ_2). Thus, we can paraphrase (3) by stating that the evaluation of (x_1^0, x_2^0), and hence of y, is obtained by integrating the welfare mass over the shaded area.

There is a second plausible, but less rational, answer to Van Praag's question. Since we are dealing with the evaluation of y *before* it is spent, the consumer considers each point under the curve to be feasible. And rather than going through the painstaking exercise of having to decide what his preferred point may be, he may evaluate y by identifying it with the total welfare mass under the curve. That is, the evaluation of y, denoted by $U(y)$, is

$$U(y)=\iint_{\ln f_1(\zeta_1)+\ln f_2(\zeta_2)\leq \ln y} \mathrm{d}W(\zeta_1, \zeta_2). \tag{4}$$

Eq. (4) represents Van Praag's preferred answer to his own question.

From comparison of (4) and (3), it is evident that (4) represents an overestimation of the welfare that the consumer can derive from spending an amount y on the commodity group. As a result, actually spending an amount y on the commodity group will lead to a slight disappointment. Although this incorrect anticipation by the consumer may be less than rational, introspection suggests that it is nevertheless realistic.

Notice that formally $U(y)$, given in (4), represents the distribution function of a sum of two random variables $\ln f_1(\zeta_1)+\ln f_2(\zeta_2)$. Also formally, these random variables are independently distributed because of (1). Generally, we are dealing with n of those variables. Under some further technical assumptions and assuming that n is large (i.e., the commodity group is broad) we can then infer that the evaluation of $\ln y$ has approximately the mathematical form of a normal distribution. Consequently, the evaluation of y, $U(y)$, has approximately the mathematical form of a lognormal distribution function. We write this as

$$U(y)\approx N(\ln y;\ \mu, \sigma)=\Lambda(y;\ \mu, \sigma), \tag{5}$$

where $N(\cdot;\ \mu, \sigma)$ is the normal distribution function with mean μ and variance σ^2, and $\Lambda(\cdot;\ \mu, \sigma)$ is the lognormal distribution function with parameters μ and σ. Of course the parameters μ and σ will differ across commodity groups and across individuals. The function U is called the *Partial Welfare Function* (PWF) of the commodity group under consideration.

How broad a commodity group has to be in order to make the approximation (5) sufficiently close is an empirical matter. It appears from empirical work (e.g., Kapteyn et al. 1979, 1980) that durables, like a refrigerator or a car, are represented by a sufficiently large number of characteristics to make the corresponding PWFs approximately lognormal. The best approximation may be expected when we consider the broadest possible commodity group, to wit total expenditures, or, taking savings as postponed expenditures, total income. When y refers to income, (5) is called the individual *Welfare Function of Income* (WFI). In the sequel we shall often illustrate aspects of the theory by means of the WFI.

The result (5) can be generalized to the evaluation of money amounts spent on a number of commodity groups simultaneously. A vector $(y_1, \ldots, y_k)$ of money amounts spent on k broad commodity groups is approximately evaluated according to an k-variate lognornal distribution function.

3. An interpretation of the WFI parameters

For a good understanding of what follows it is useful to first take a closer look at the economic and psychological interpretation of the *welfare parameters* μ and σ of the WFI. Similar interpretations pertain to the parameters of PWFs.

We start with μ. In fig. 2a, a few lognormal distribution functions have been drawn for various values of μ. The parameter values have been taken from a survey in The Netherlands, conducted in 1975. It appears that the WFI shifts to the right when μ increases. The quantity e^{μ} is the income level evaluated by 0.5. The larger an individual's μ is, the larger the income is he needs in order to evaluate it by 0.5; hence, e^{μ} can be seen as a want parameter.

In fig. 2b, lognormal distribution functions have been drawn for various values of σ. It can be seen that the WFI becomes flatter when σ increases. The larger an individual's σ, the broader the range of incomes that are evaluated substantially different from zero or one. If σ is very small, the WFI becomes almost a step function at e^{μ}. An individual with such a small σ will be quite happy if his income is slightly above e^{μ} and quite unhappy if his income is slightly below e^{μ}. Notice that e^{μ}, and hence μ, depends on the unit chosen, whereas σ is dimensionless. Note also that the WFI (and IWFs in general) is s-shaped, being first

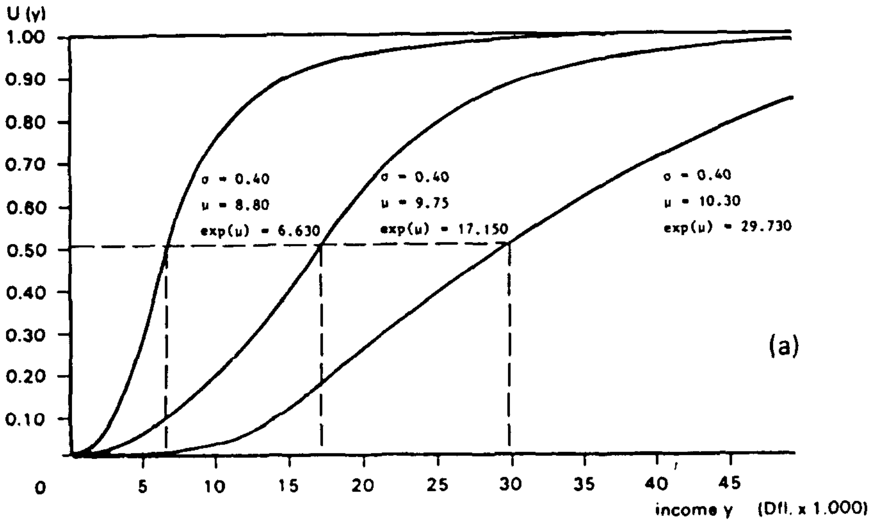

Fig. 2a. The WFI for various values of μ.

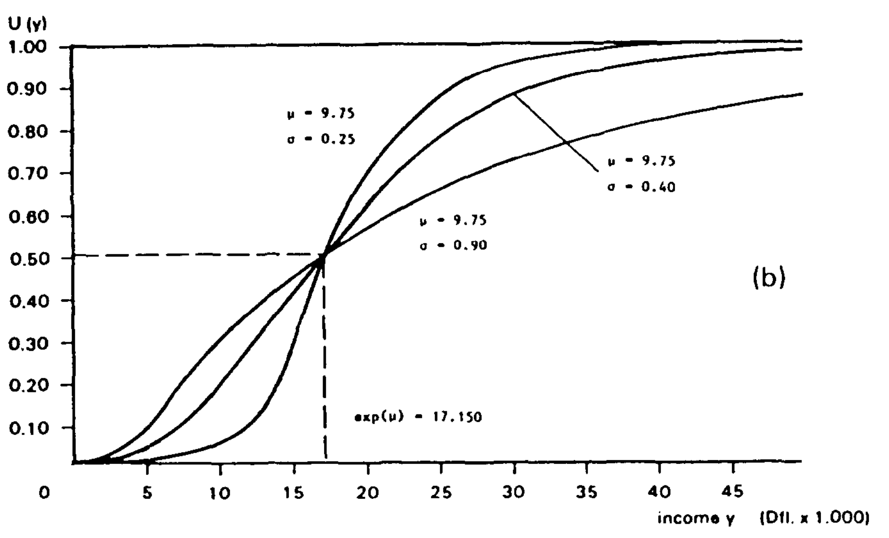

Fig. 2b. The WFI for various values of σ.

convex from 0 up to $\exp(\mu - \sigma^2)$ and concave thereafter, thus violating Gossen's first law (decreasing marginal utility with increasing income) for lower values of y.

Although Gossen's first law is often mentioned in textbooks, we are not aware of any empirical tests of it. Of course, the law can only refer to cardinal utility functions, because an ordinal utility function can be transformed at will to make it concave. Within a cardinal framework, increasing marginal utility at the lower income range seems quite a plausible property for a utility function of income. If income is below subsistence level, each extra dollar brings one closer to the point where survival is possible and hence each extra dollar carries a higher marginal utility. According to this reasoning, the inflection point $\exp(\mu - \sigma^2)$ would be a prime candidate for a definition of a poverty line. Indeed, in related research on poverty (e.g., Goedhart et al. 1977), it is found that the income level at which people report to be just able to make ends meet is approximately equal to $\exp(\mu - \sigma^2)$. Hitherto, this phenomenon has not been investigated in more depth, however.

In practice, the phenomenon of initially increasing marginal utility is of minor importance as it turns out that almost all observations correspond to the concave part of the function (most people are able to make ends meet), thereby maintaining Gossen's first law in practice. In the context of utility maximization using the multivariate lognormal function (see the last paragraph of section 2), the s-shape does not pose problems as the multivariate lognormal is quasi-concave, hence the indifference curves are convex towards the origin as usual.

4. The measurement of μ and σ

One of the major features of the theory of the individual welfare function is the possibility to measure PWFs and WFIs directly on the basis of survey questions. As PWFs and WFIs are determined by two parameters, measuring them amounts to measuring their parameters.

We concentrate on a description of the measurement of WFIs. They are measured by asking respondents to a survey of the so-called *Income Evaluation Question* (IEQ) (see table 1) [2]. The labels 'excellent', 'good', etc. are translated into numbers between zero and one in the following simple way. Each label is supposed to 'carve out' one ninth of the [0,1]-interval; thus, 'excellent' corresponds to the interval $(\frac{8}{9},1]$,

[2] The precise wording and the number of levels differ somewhat between surveys. Often a simpler version is used, where the respondent is invited to associate with a verbal label a single income level, rather than an income range.

'good' to $(\frac{7}{9},\frac{8}{9}]$, etc. The original motivation for this so-called equal interval assumption, by Van Praag (1971), was based on an information maximization argument, which was later generalized by Kapteyn (1977). The argument is roughly as follows. The answers to the income evaluation question furnish a division of the income range into income brackets $[y_0, y_1), [y_1, y_2), \ldots, [y_n, y_{n+1})$, where $y_0 = 0$ and $y_{n+1} = \infty$. The division differs from individual to individual, but certainly the division is not given in a random way. It is not unreasonable to assume that the individual tries to inform us as exactly as possible about his welfare function. He attempts to maximize the information value of his answer. How can we define the information value?

Consider the answer in table 1. The welfare evaluation of an income in the bracket [25000, 30000), labelled 'sufficient', is on average

$$\tfrac{1}{2}[U(25{,}000) + U(30{,}000)] \equiv U(\bar{y}_5), \tag{6}$$

with $\bar{y}_5$ defined implicitly. However, we cannot say that all income levels in [25000, 30000) are evaluated by $U(\bar{y}_5)$. The average inaccuracy

Table 1

Taking into account your own situation with respect to family and job you would call your *net-income* (including fringe benefits and after subtraction of social security premiums):*

per { week A, month B, year C }

excellent	*if it were above*	*Dfl.* 45,000		
good	*if it were between*	*Dfl.* 35,000	*and*	*Dfl.* 45,000
amply sufficient	*if it were between*	*Dfl.* 30,000	*and*	*Dfl.* 35,000
sufficient	*if it were between*	*Dfl.* 25,000	*and*	*Dfl.* 30,000
barely sufficient	*if it were between*	*Dfl.* 22,000	*and*	*Dfl.* 25,000
insufficient	*if it were between*	*Dfl.* 20,000	*and*	*Dfl.* 22,000
very insufficient	*if it were between*	*Dfl.* 17,000	*and*	*Dfl.* 20,000
bad	*if it were between*	*Dfl.* 12,000	*and*	*Dfl.* 17,000
very bad	*if it were below*	*Dfl.* 12,000		

* Encircle your reference period.

Note: The money amounts entered (in roman) are the response of an arbitrary chosen respondent of a 1971 survey in The Netherlands.

of evaluating the income levels in [25000, 30000) by $U(\bar{y}_5)$ may be defined as

$$\int_{25.000}^{30.000} [U(y) - U(\bar{y}_5)]^2 \mathrm{d}U(y). \tag{7}$$

When we have a partition $[0, y_1), [y_1, y_2), \ldots, [y_n, \infty)$, the total average inaccuracy of this partition is defined by

$$\sum_{i=0}^{n} \int_{y_i}^{y_{i+1}} [U(y) - U(\bar{y}_i)]^2 \mathrm{d}U(y). \tag{8}$$

Now it is evident that the separate integrals increase with the variation of the U-function on $[y_i, y_{i+1})$ and with the interval length $(y_{i+1} - y_i)$. Hence, the individual selects narrow brackets where the U-function is steep, and wide brackets where U increases slowly. Mathematically, the individual attempts to choose the y_i-values in such a way that (8) is minimized. It is a matter of some algebra to show that the solution is $U(y_i) = i/(n+1)$, which justifies the equal interval assumption.

Until recently, the procedure was untested and almost all measurements of WFIs and of PWFs have been based on this assumption. Some recent tests were carried out by Buyze (1982) and Antonides et al. (1980). Buyze (1982) builds a statistical model in which the equal interval assumption is tested by means of an analysis of variance technique. Her conclusion is that the assumption may hold true approximately, but not exactly. Antonides et al. (1980) come to a similar conclusion on the basis of a direct test (see below).

Given a procedure to translate the labels in the IEQ into numbers between zero and one, estimation of the WFI-parameters μ and σ of the individual who answered the IEQ amounts to a simple problem of curve fitting. As one sees from the question, generally there are eight different income responses. This yields a scatter of eight points from which this individual's μ and σ can be estimated. By plotting the eight points on lognormal paper the s-shaped scatter turns into a linear scatter and μ and σ are estimated by a linear regression with eight points. The estimation method is very simple and cheap to apply, also in large scale surveys. For results of this method, see for example Van Praag (1971), Van Praag and Kapteyn (1973), Van Herwaarden et al. (1977).

Since the estimation method was developed, some 70,000 WFIs have been measured for individuals in ten European countries and the U.S. In addition, about 11,000 PWFs have been measured in two surveys in The Netherlands [3]. The fit of the regression to estimate μ and σ per individual is, on average, very good. The typical average correlation coefficient (unsquared and uncorrected for the number of degrees of freedom) in the various samples equals 0.98 [4].

The measurement method outlined above does not hinge upon the lognormality of the WFIs or PWFs. It can be applied equally well to other functional specifications that are restricted to the [0,1]-interval. This allows for a simple test of the lognormal specification vis-à-vis other two-parameter alternatives. Such a test has been performed by Van Herwaarden and Kapteyn (1981), for about 9,000 PWFs and 15,000 WFIs, using 12 alternative specifications like the normal, logistic, and Weibull distribution functions, and the logarithm and the straight line, both truncated from above and below to restrict their range to the [0,1]-interval. It appears that the lognormal and logarithmic functions outperform the other ones. The logarithm seems to fit slightly better than the lognormal. Results obtained by Antonides et al. (1980) suggest that this may be due to the equal interval assumption. These authors use data from a survey in which respondents have been asked to attach numerical values to the verbal labels used in the income evaluation question. Thus, they can investigate the validity of the equal interval assumption directly. It turns out that the equal interval assumption holds true approximately, but at the lower end of the scale the equal interval assumption attaches numbers to the verbal labels that are too low. As a result the graph of the WFI is somewhat distorted by the assumption, in such a way that the scatter of points becomes too concave. Although further research into these matters is certainly

[3] PWFs are measured by a question analogous to the PWF: the respondent is invited to report, for one or more durables he is planning to buy, the expenditure on each durable that he associates with a perfectly satisfactory purchase, a rather satisfactory purchase, etc.

[4] See, e.g., Van Herwaarden et al. (1977). In itself a good fit is not surprising since the reported income levels or expenditure amounts are bound to be monotonous, which virtually guarantees a high correlation coefficient. Wierenga (1978) has investigated the significance of the correlation coefficient by simulating pseudo-respondents whose 'answers' were obtained by randomly drawing from a rectangular distribution. He claims that the random data give as good a fit as the real data. In a rejoinder (Van Praag et al. 1978) it is shown, however, that his results depend crucially on the way he generates his random data, and that a random data generating process closer in line with the data actually observed gives a significantly worse fit than the real data.

warranted, for the moment the general conclusion may be drawn that the lognormal provides a fairly good approximation to the true functional form of WFIs and PWFs.

The method described to measure the parameters of a utility function is, of course, unusual in economics, where measurement usually takes place via observed behavior. Kapteyn et al. (1979) discuss the difference between both measurement methods (*direct* versus *indirect* measurement), and the implications of this difference for the testing of models of consumer behavior.

5. A theory of preference formation

Different individuals will have different PWFs for the same commodity group. In this section we discuss a theory which explains the formation of individual preferences, developed by Kapteyn (1977). This development was spawned by earlier empirical results obtained in inter alia Van Praag (1971), Van Praag and Kapteyn (1973) and Kapteyn et al. (1976). Again, our exposition is sketchy and attempts to be didactic rather than formally correct. We once more focus attention on the broadest commodity group, income, and discuss the formation of an individual's WFI.

Consider one of the N individuals in a society. Let us call him n. By various kinds of social interaction, he is likely to have a more or less exact perception of the incomes in society. Probably the perception is selective: he will attach much weight to some incomes, like those of relatives, colleagues and other members of his 'peer group', and little weight to other incomes, like those of individuals with a different occupation, a different education level and a different geographical location.

We represent this subjective perception of other individuals' incomes by a set of *reference weights* w_{nk}, $k = 1, \ldots, N$. These weights are assumed to be non-negative and are normalized to add up to unity. Note that w_{nn} is one of the reference weights, because individuals evidently also perceive their own income.

We define individual n's *perceived contemporary income distribution* as the income distribution in which all incomes have been weighted by these reference weights. It would be equal to the actual income distribution if all w_{nk} would be equal, i.e., $w_{nk} = 1/N$; then individual n would attach the same importance to each and every income in society. Intuitively, this seems to be an unlikely occurrence.

The notion of a perceived contemporary income distribution can be clarified by a simple example. Consider a society consisting of three individuals ($n = 1, 2, 3$). Let the reference weights and incomes be as in table 2; so, individual 1, who has an \$8,000 income, refers for 50% to individual 2, and for the remaining 50% to himself. The income of individual 3 is irrelevant for him. Individual 2, who is in the middle of the income distribution, refers to individual 1 for 1/6 and to individual 3 for 1/3. The weights assigned by individual 3, who has the highest income, are given in the last line of the table. Note that, in this example, all w_{nn}'s are taken to be the same ($= 1/2$). Whether or not this is true is an empirical matter.

The three perceived contemporary income distributions are visualized in fig. 3a. In addition, the actual income distribution is also sketched in the figure as a dotted line. The latter distribution is, of course, equal to a perceived contemporary income distribution for an individual with reference weights (1/3, 1/3, 1/3).

The graphs depicted in fig. 3a are step-functions. To aid the intuition behind the preference formation theory to be introduced below, fig. 3b contains an adaptation of fig. 3a for the situation when society consists of about ten individuals rather than three. Without giving a detailed numerical example, it is clear that the various distributions will now be represented by finer step-functions. (The figure only depicts the perceived contemporary income distribution of a subpopulation, viz., those of individuals 1, 2 and 3, plus the actual distribution.) The final step is taken in fig. 3c: the number of individuals in society is now so large that optically the step-functions have become smooth curves.

An individual's perception of the income distribution may change over time, either because the actual income distribution changes or because his reference weights change. It is unlikely that when a new contemporary income distribution emerges, the individual immediately forgets the previous one. Rather, it seems plausible that previous

Table 2
Reference weights and incomes in an imaginary three-person society.

Individual	Income (in \$)	Weights attached to		
		1	2	3
1	8,000	1/2	1/2	0
2	10,000	1/6	1/2	1/3
3	15,000	1/8	3/8	1/2

income distributions keep lingering in his memory. To clarify the issue, we introduce the notion of time a little more formally. Time is measured in discrete periods: $t = 0, -1, -2, \ldots,$ where $t = 0$ denotes the present. Any individual is assumed to weigh experiences in different

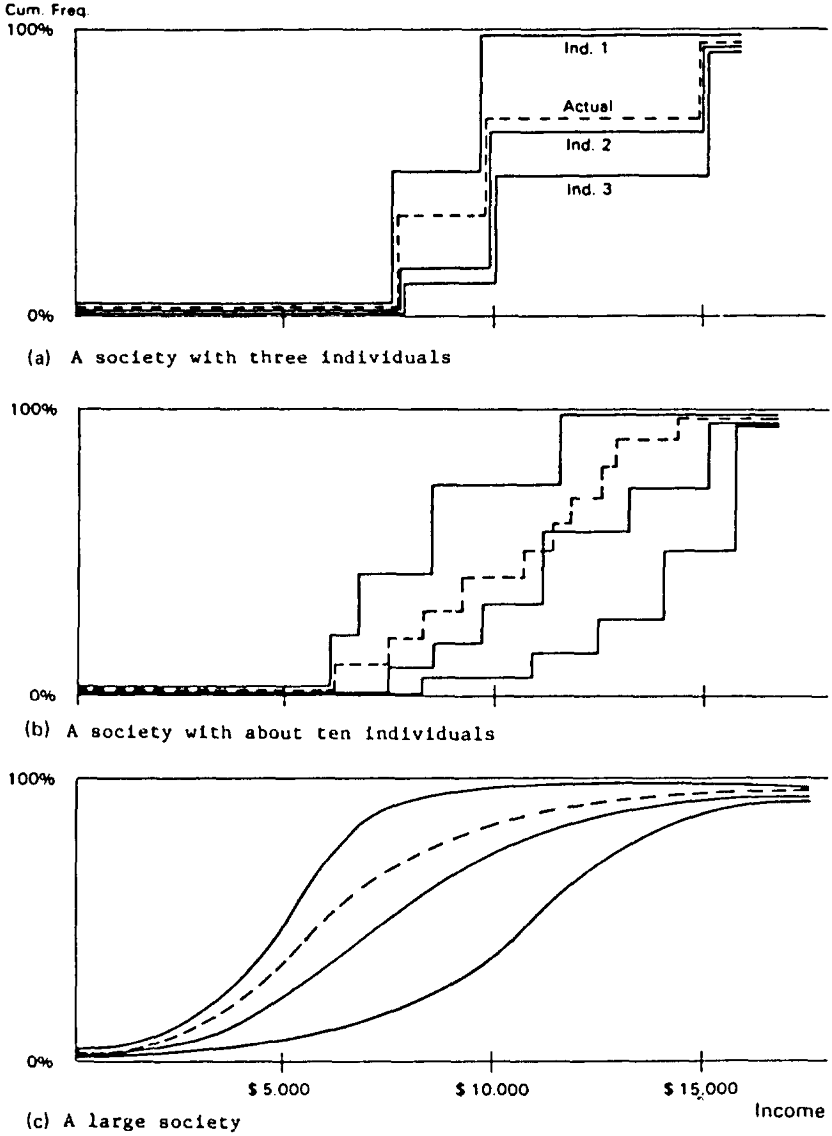

Fig. 3. Actual income distribution and perceived contemporary income distributions.

periods by *memory weights*, a_{n0}, $a_{n,-1}$, $a_{n,-2}$, adding up to unity. Presumably, a_{nt} is monotonically increasing in t, as recent experiences will be perceived stronger than experiences longer ago.

Now, we can introduce the concept of an individual's overall *Perceived Income Distribution* (PID) as the memory weighted sum of the perceived contemporary income distributions of all separate periods. Thus the perceived contemporary income distributions are building blocks of the overall perceived income distribution. Casual observation suggests that an individual's WFI will be related to his PID. In fig. 3a, for instance, we would expect individual 3's WFI to be located more to the right than the WFI of individual 1. From the infinity of possible relations between the WFI and the PID, Kapteyn (1977, 1979) has proposed the simplest one. His theory of preference formation implies that an individual's WFI is *identical* to his PID.

One of the basic tenets in the lognormal framework was the isomorphism between a WFI and a probability distribution function, and now this simile has been extended to claim that a WFI *is* actually a perceived income distribution function. Analogously, when considering the expenditures on a commodity group rather than income, e.g., expenditures on cars, the theory of preference formation states that the PWF of cars is identical to the perceived distribution of expenditures on cars.

An implication of the theory is, for example, that an individual whose income is in the 80th percentile of his perceived income distribution will evaluate his income by 0.80, an individual who is in the 60th percentile will evaluate his income by 0.60, etc. The idea behind the theory of preference formation is almost implausibly simple. But it appears to be rich in its possibilities to explain social and economic phenomena and in its implications for empirical research.

6. Implications of the preference formation theory

Some implications of the preference formation theory deserve attention. The main theme of the theory is the *relativity* of the evaluation of income or, in the case of a PWF, of expenditures on a certain commodity group. (In the latter case, the theory implies, for example, that I evaluate my car by 0.80 if my car is at least as expensive as 80% of the cars in my reference group.) It is the relative position in the income distribution (or distribution of expenditures), appropriately reweighted, that determines the individual's utility level; not the absolute level of an individual's income.

Reconsider the example given in table 2. According to the theory, and momentarily ignoring the dynamic aspect, individual 1 evaluates his income by 0.50, since 50% of the population (viz., he himself, subjectively weighted) has an income smaller than or equal to his own income. Individual 2 evaluates his income by 0.67 (= 1/6 + 1/2), and individual 3 by 1 (= 1/8 + 3/8 + 1/2). The latter individual would not very much appreciate the $8,000 income that individual 1 has: he would evaluate that by approximately 0.13 (= 1/8). Neither would he be enthusiastic about the idea of an income reduction to individual 2's level: he evaluates a $10,000 income by 1/8 + 3/8 = 1/2.

What happens when individual 1 gets his income increased to $18,000? From table 2, it directly follows that the individuals' evaluations of their own incomes now become 1, 1/2 and 7/8, respectively. That is, individual 2 and 3 have become worse off due to individual 1's income increase. Their welfare losses, however, are not very dramatic as each of them attaches a low weight to individual 1's income.

These numbers are only an illustration of the basic points, a more complete discussion being postponed to later sections. What they do suggest, however, is the importance of the distributional aspect of income (expenditure, etc.) evaluation. The preference formation theory is reminiscent of a statement by Scitovsky (1976: 199): '(W)hen people seek status (...) in a general token, like income, (...) the seeking of status becomes a zero sum game'. By and large, the preference theory subscribes to this point of view, be it that the sum need not be exactly zero, depending on the pattern of reference weights involved.

The relative nature of evaluations is well documented in the sociology and psychology literature. This is not the place to review that literature. A number of instances from this literature are dealt with in Kapteyn and Wansbeek (1982).

Now we turn to a discussion of the more formal aspects of the preference formation theory, WFI = PID. As both the WFI and PID are, formally, distribution functions, their equality implies that their log-moments must be equal. The equality of their first log-moments (the log-means) can be shown to imply (see, e.g., Kapteyn 1977, 1979):

$$\mu_n = \sum_{t=-\infty}^{0} a_{nt} \sum_{k=1}^{N} w_{nk,t} \ln y_{kt}, \tag{9}$$

where $w_{nk,t}$ is the reference weight assigned by individual n to individual k at time t. In words, (9) states that an individual's μ is a weighted average of log-incomes, the weights being memory weights, a_{nt}, and reference weights, $w_{nk,t}$.

In principle, (9) is straightforward to test by means of a regression type of analysis. The observed dependent variable μ_n depends linearly on observable explanatory variables, the $\ln y_{kt}$. This idea underlies the empirical work to be described below. Before that, we briefly discuss two extensions of the simple model (9).

Firstly, the development so far has been in terms of individuals rather than families, and especially for empirical work the distinction is important. Let f_{nt} be the number of *equivalent adults* in family n at time t, then the preference formation theory can be reformulated in terms of *per capita income* (y_{nt}/f_{nt}). This implies (after some algebra) that, in (9), μ_n should be replaced by $\mu_n - \ln f_n$ [5], and y_{kt} by y_{kt}/f_{kt}; see Kapteyn et al. (1980). So (9) has to be replaced by

$$\mu_n = \ln f_n + \sum_{t=-\infty}^{0} a_{nt} \sum_{k=1}^{N} w_{nk,t} \ln y_{kt} - \sum_{t=-\infty}^{0} a_{nt} \sum_{k=1}^{N} w_{nk,t} \ln f_{kt}. \qquad (10)$$

Secondly, the equality of the second log-moments of the WFI and the PID (the log-variances) implies, analogous to (9):

$$\sigma_n^2 = \sum_{t=-\infty}^{0} a_{nt} \sum_{k=1}^{N} w_{nk,t}(\ln y_{kt} - \mu_n)^2. \qquad (11)$$

We thus obtain a second relation between observable quantities, σ_n^2 and $(\ln y_{kt} - \mu_n)^2$. Also this relation can be extended to account for family size effects, in much the same way as (9). The resulting expression is a bit more complicated than (10) and will not be given here (see, e.g., Kapteyn et al. 1980). The above analysis also applies to the explanation of PWF parameters. The parameters μ_n and σ_n then refer to the PWF of a commodity group under consideration and y_{kt} is the amount spent on the commodity group in period t by individual k.

[5] We adopt the convention that subscripts equal to zero are omitted. So, for example, f_{n0} is written as f_n and $w_{nk,0}$ is written as w_{nk}.

7. Empirical evidence

Eqs. (10) and (11) (the latter after adaptation for family size effects) show how the preference formation theory implies relations between observable quantities. After adding disturbance terms to the right-hand sides of (10) and (11), one obtains a system of two regression equations that state how μ_n and σ_n^2 depend on $\ln y_{kt}$ and $\ln f_{kt}$, with k running over all individuals in society, and t running over all time periods from minus infinity to zero. In principle, this system can be estimated from panel data. Since this type of data has become available only very recently, no testing of the full model (10) and (11) has yet been carried out. Rather, pieces of the model have been estimated on the basis of various data sets. Part of these results were obtained on the basis of simpler models that were specified before the preference formation theory was formulated. These results can be reinterpreted in terms of (10)–(11) and provide evidence for the theory's plausibility. More recently, models have been specified and tested that are explicitly derived from (10), but still ignore (11).

In this section we sketch some of the empirical results obtained. The first problem dealt with is the specification of f_{nt}. Secondly, we look at the specification of the reference weights $w_{nk,t}$, and thirdly at a_{nt}. Most of the exposition will again take place in terms of WFIs, but at the end of the section we also consider PWFs.

7.1. *Family income equivalence scales*

The function f_{nt} represents the number of equivalent adults in family n at time t. Since only cross-section data have been available until recently, we shall only pay attention to $f_n \equiv f_{n0}$, i.e., the number of equivalent adults in family n at the time the data were collected. The specification of f_n amounts to the construction of *family equivalence scales*, a problem with a long history; see, e.g., Deaton and Muellbauer (1980) or Pollak and Wales (1982). The theory developed so far does not give guidance as to the specification of f_n, nor does the traditional literature. A possible crude specification is

$$\ln f_n = \beta_0 + \beta_1 \ln fs_n, \qquad (12)$$

where fs_n is the number of persons in family n. The main advantage of this formulation (used in Van Praag (1971) and Van Praag and Kapteyn (1973)) is its simplicity. In Kapteyn and Van Praag (1976), a more sophisticated specification is used by specifying fs_n as:

$$fs_n = \sum_{j=1}^{m} \alpha_j g(l_{nj}), \tag{13}$$

where the summation is over the m members of family, α_j is a *rank weight* assigned to the j-th family member (the members are ranked in order of decreasing age) and g is a monotonically non-decreasing function of the age, l_{nj}, of the j-th family member. The function g contains three unknown parameters, α_j is a two-parameter function in j. The parameters are estimated on the basis of a cross-section of about 3,000 members of the Dutch Consumer Union, taken in 1971, to which the following model is applied:

$$\mu_n = \beta_0 + \beta_1 \ln fs_n + \beta_2 \ln y_n + \epsilon_n, \tag{14}$$

with ϵ_n a random disturbance term and fs_n defined by (13). Essentially, (14) is obtained from (10) by omitting all terms except $\ln f_n$ (for which the right-hand side of (12) is substituted), and $a_n w_{nn} \ln y_n$ (with $a_n w_{nn}$ replaced by β_2). In other words, all variables pertaining to the past or to individuals in the reference group are omitted. The omitted terms are represented by ϵ_n. Presumably, the omitted terms correlate with $\ln fs_n$ and $\ln y_n$, which introduces specification errors. Kapteyn and Van Praag (1976) ignore these errors, partly because the preference formation theory had not been formulated yet and partly because the availability of cross-section data only forced the omission of the bulk of the terms in (10) anyhow. These specification errors detract from the validity of the empirical results. However, similar specification errors are made in all traditional investigations on family equivalence scales based on demand studies.

The key empirical findings with respect to (13) are: the cost of family members (i.e., their contribution to the total number of equivalent adults in the family) decreases sharply with their rank number and rises with age between about 25 and 50 years. Below 25 and above 50 the cost is approximately constant. Estimation of (14) on the basis of two Belgian samples, reported in Kapteyn (1977), yields similar results.

7.2. Social reference groups

In order to do more justice to the preference formation theory as represented by (10) and (11), one has to model the reference weights. The specification of these reference weights poses the thorniest econometric problem in the estimation of (10) and (11). The main problem is that there are too many of them, viz., $N(N-1)$ for each t; that is, N^2 minus N as, for all n, $\sum_{k=1}^{N} w_{nk,t} = 1$. Evidently, restrictions are needed to be able to estimate the reference weights al all.

Kapteyn et al. (1976) and Van Praag et al. (1979) set out to estimate the reference weights pertaining to year zero, $w_{nk,0}$ (abbreviated to w_{nk}), in (10) and (11), omitting all terms that pertain to the past. They specify the w_{nk} as functions of the similarity in social characteristics of individuals n and k. The core of the parametrization is the introduction of six matrices of 'partial reference weights', one for each of the six characteristics distinguished: Education (4 levels), Sector of employment (government, industry, or not employed), Job-type (5 types), Degree of urbanization (2 categories), Age (4 brackets) and Place of residence (western part of The Netherlands, remainder). An entry in such a matrix (the (i, j)-th, say) represents the average weight given by an individual in category i of the characteristic concerned to an individual in category j. Weights are normalized to sum to 1 per row. The overall reference weight w_{nk} is expressed as a function of the product of the six partial reference weights involved.

Since the number of categories per characteristic is 4, 3, 5, 2, 4, and 2, respectively, there are in principle still 68 parameters involved. By further parametrization, this number was reduced to 26. These 26 parameters have been estimated, together with β_0, β_1 and the parameters implicit in the definition of fs_n.

In general, it appears that (i) estimation of the reference weights is possible, although complicated, thereby providing the first empirical quantification of preference interdependence in economics, (ii) the inclusion of a reference group term yields a significant improvement of the explanation of μ, compared to the model (14), (iii) about 16% of the variation in σ^2 is explained by the reference group model.

As to the detailed results the most interesting aspects concern the partial reference weight matrices. They yield insight into the social process of 'who looks at whom' when evaluating own income. So it appears, for example, that individuals give no weight to incomes of other individuals whose educational attainment is lower; individuals only give weight to incomes of others of the same or higher level of education, with the exception of those with only primary education: they only consider 'peer' incomes. As to age, individuals appear to refer to individuals of at most the same age, not to older ones. For numerical values and for results on the other four characteristics, the reader is referred to the papers mentioned above.

7.3. Habit formation

Recently, panel data (the first two years of a longitudinal study) collected by the Netherlands Central Bureau of Statistics have been used by Van de Stadt et al. (1985) to estimate (10), ignoring (11) for reasons of simplicity. This neglect inflicts a loss of efficiency but introduces no specification error. They assume a geometric pattern for the a_{nt} (i.e., $a_{nt} = (1-a)a^t$). The reference weights have been modelled a lot simpler than before; only three characteristics are considered (education, employment status and age) and the weights w_{nk} ($k \neq n$) can only take on two values, say κ and γ. The value κ is assumed if individuals n and k have the same level of education, the same employment status (self-employed, employee or unemployed) and are in the same age bracket; otherwise $w_{nk} = \gamma$. Furthermore, w_{nn} is assumed to be the same for all n. Thus, rather than having to estimate 26 parameters, as in Kapteyn et al. (1976), only the three parameters κ, γ and w_{nn} remain to be estimated. Of course, the assumptions made are not correct, but the resulting inaccuracy is modelled explicitly.

The geometric specification of a_{nt} makes it possible to rewrite (10) in lagged form, so that one obtains a relation explaining this year's μ_n on the basis of incomes and family sizes in this year and the value of μ_n one year ago. Some results are (standard errors in parentheses): $\hat{a} = 0.83$ (0.15), $\hat{w}_{nn} = 0.66$ (0.19).

The estimate for w_{nn} indicates that an individual's preferences regarding income are about half as much influenced by the incomes in his social reference group (past and present) than by his own (past and present) income. The value of $\hat{a}$ indicates that the incomes in the previous eight years (both own income and incomes in the social reference group) determine 80% of the individual's μ, i.e., $(1-a)(1+a+\ldots+a^6+a^7) \approx 0.80$, if $a = 0.83$.

A previous study (Kapteyn et al. 1980), pertaining to holiday expenditures rather than income, exhibited a lower value of a ($\hat{a} = 0.57$) and w_{nn} ($\hat{w}_{nn} = 0.13$), indicating a shorter memory span and a higher degree of preference interdependence. These differences may be due to the differences in the object studied (holiday expenditures are very conspicuous), but may also be due to the larger imprecision of the holiday expenditures results, where some strong assumptions had to be made in order to estimate a dynamic process from a single cross-section.

7.4. Further evidence

Rather than estimating more or less complicated models, one can also try to corroborate or refute the preference formation theory on the basis of qualitative relationships. We shall give a few examples relating to WFIs.

Since a WFI is equal to a perceived income distribution function, we predict σ's of individuals in countries with a relatively unequal income distribution to be larger than in countries with a more equal distribution. This prediction is borne out by a comparison of ten European countries (Van Praag et al. 1980). If one's income has varied a lot over time, σ_n^2 will be large, according to (11). Indeed, it turns out that individuals with widely fluctuating incomes have high σ's (Van Herwaarden et al. 1977). Another type of test rests on an analysis of the specification error that is introduced if one estimates a relation like (14). It can be made plausible that, if (9) is the correct expression explaining μ, the estimate of β_2 will be smaller if incomes fluctuate more over time. This effect is found in various samples (Kapteyn 1977).

8. Some theoretical exercises

After having discussed empirical evidence to support the preference formation theory, let us have a closer look at some of its implications. We shall discuss these implications in terms of WFIs, the exposition for PWFs being analogous. To simplify the exposition, it is convenient to carry out the analysis in terms of income 'per equivalent adult', or, what amounts to the same thing, to assume that all families in society are of equal size. Moreover, we ignore the lags implied by the memory function, i.e., we assume that each change in preferences due to a change in circumstances takes place instantaneously. (An equivalent procedure would be to say that we only consider long-term effects.)

Let us consider what happens when individual n, say, gets a wage raise. If individual n is the only one to get a wage raise, we expect him to be better off after the raise, because he will attain a higher relative position in his perceived income distribution. If, however, all individuals in society get the same percentage raise, nothing happens to individual n's relative position in the income distribution and he will evaluate his new income by the same number as his old income. This was already indicated in section 6. Employing the preference formation theory, the welfare effect of the wage raise can be studied in some more detail.

Consider fig. 4. Let curve I be individual n's WFI before the raise. He evaluates his income y_{I} by 0.6 (point A). When he is promised a raise to y_{II} he expects to evaluate the new income by 0.9 (point B). We call this his *ex ante* evaluation of y_{II}. If he actually gets the raise (and no one else gets a raise), (10) implies that μ will rise and the WFI will shift to the dashed line. Thus, *ex post* the new income is only evaluated by 0.8 (point C). This evaluation is higher than the evaluation of y_{I}, reflecting the fact that y_{II} is higher in the perceived income distribution than y_{I} was. But it is not as high as was anticipated *ex ante*. This is basically due to the fact that in individual n's perceived income

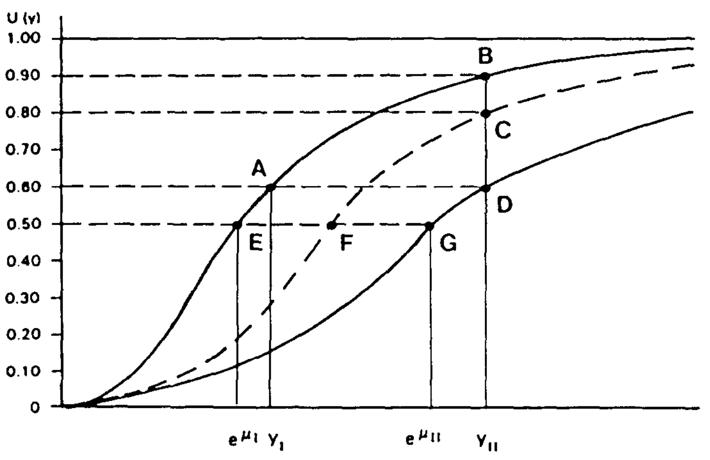

Fig. 4. The effect of a wage raise.

distribution his own income plays a prominent role, so that with each change of his income the perceived income distribution shifts as well.

If everybody else (or at least everyone in individual n's reference group) gets the same percentage increase in income, (10) implies that e^{μ} will rise by that percentage, so that individual n's WFI shifts to position II. Now y_{II} is evaluated by 0.6 (point D), equal to the evaluation of y_I before the wage increase. As indicated above, this stems from the fact that individual n's relative position in his perceived income distribution has not changed.

There is a slightly different way to look at the same phenomenon. It is a well-known property of the normal and lognormal distribution function that

$$N(\ln y;\ \mu,\ \sigma) = N(\ln y - \mu;\ 0,\ \sigma) = \Lambda(y/e^{\mu};\ 0,\ \sigma). \tag{15}$$

(This property was also used in section 3.) Hence, for a given σ, the welfare evaluation of an income y is determined entirely by the ratio y/e^{μ}. Thus, if individual n's income increases from y_I to y_{II} and the WFI shifts from position I to position II, then the effect of the income increase is 'eaten up' by an equal relative increase in e^{μ}. This latter increase consists of two parts. In fig. 4 the ratio EF/EG represents the proportion that is eaten up by a shift in e^{μ} due to a change in the individual's own income. The ratio FG/EG is the proportion eaten up by the shift in e^{μ} due to the income increase in the social reference group. The ratio EF/EG has been called the *preference drift rate* (Van Praag 1971), whereas FG/EG has been called the *reference drift rate* (Kapteyn 1977). The empirical results quoted in section 7.3 suggest that the preference drift phenomenon and the reference drift phenomenon each eat up about one half of an across the board income increase.

This discussion suggests again that income evaluation is a zero sum game: an income increase for one individual implies a welfare level reduction for all others who attach a positive reference weight to this individual. Although there is ample evidence that by and large the evaluation of income is entirely relativistic (cf., e.g., Easterlin 1974; Duncan 1975), there are at least two reasons why income evaluation is not entirely a zero sum game. First of all, the pattern of reference weights may be such that at least some individuals' incomes may be raised without thereby lowering the income evaluation of others. This is corroborated in Kapteyn and Van Herwaarden (1980), who use previously estimated reference weights (see section 7.3), and find that the 1971 incomes in The Netherlands could have been distributed such that average welfare would have improved. Secondly, the dynamic aspect is ignored by assuming instantaneous adjustment. When it is recognized that the adjustment is not instantaneous but takes a number of years to become fully effective, overall wage raises are seen to have a short-term welfare effect, although the income distribution remains unaffected. It takes some time before each individual's μ has adapted to the raise, eventually erasing the welfare effect.

9. Policy applications

Directly measured welfare functions are not only convenient vehicles to investigate certain questions in positive economics, like preference formation, they can also be used to address various policy questions. To conclude, we briefly indicate a number of policy fields where the IWF can be or has been fruitfully applied. Considerations of space rule out a more extensive exposition. For a more elaborate discussion, see Wansbeek and Kapteyn (1983).

9.1. Family equivalence scales

A family equivalence scale measures the relative cost of living of families of different composition or, in our earlier terminology, it measures the number of equivalent adults in different families. The relevance of such measurement to income maintenance policies is obvious. Within our framework, this measurement is straightforward: Find a reasonable specification of f_n and apply (10) to panel data. The research discussed in section 7.1 is an instance of such research. In Kapteyn and Van Praag (1980) the differences between and communalities with other, more traditional methods are discussed.

9.2. Definition of a poverty line

Two approaches to the problem of determining a poverty line have been followed in the framework of the WFI: One is to add to the IEQ a direct question as to the minimum income a respondent feels he needs in his circumstances to make ends meet. The other one is to assume that politicians are willing to impose a certain welfare level (0.5, say) as a minimum for all members of society and to use the WFI to translate this level into a minimum income level, differentiated by family size (and possibly other socio-economic characteristics). The first approach leads to a new intersubjective definition of a poverty line, cf. Goedhart et al. (1977), Kapteyn and Halberstadt (1980), Kapteyn et al. (1985), Colasanto et al. (1984), Danziger et al. (1984). The second approach is explored in, e.g., Goedhart et al. (1977), Van Praag et al. (1982a, b), Hagenaars and Van Praag (1985). The latter studies are based on surveys taken in the member countries of the European Community.

9.3. Optimal income redistribution

A brief discussion was given in section 8. A more extensive analysis of the same points is provided in Kapteyn (1977). Van Praag (1977, 1978) uses the dependency of μ_n on own income (cf. (14)) to have another look at income inequality. Since individuals do not have identical WFI's, a certain income change is translated into different welfare changes by different people. As a result, income inequality is perceived differently by different individuals and policy measures that are considered egalitarian by some may be perceived as increasing the inequality of incomes by others.

9.4. Municipal welfare functions

In a social welfare kind of setting IWFs can be extended to apply to various levels of government, like municipalities. Van Praag and Linthorst (1976) analyze the response to a survey of Dutch municipalities in which officials of local governments answer IEQ-type of questions pertaining to municipal expenditures in a number of fields. Such an analysis may, for instance, be used to design an optimal (in some sense) allocation of block grants to local governments.

9.5. Public goods

Dagenais (1977) describes an experimental survey where a bivariate IWF is measured, pertaining to 'income' and 'air quality'. Such information can, for example, be used to assess the distributional effects of an air pollution abatement project.

10. Conclusion

The empirical results discussed are in many ways preliminary. Mainly due to data limitations, the specification and estimation of the various models leave something to be desired. That applies both to the measurement of WFIs and to the various relationships between their parameters and other variables. Still, it appears that some questions posed in the Introduction can be answered with confidence. Let us mention a few of them.

Economists' retreat to ordinal utility does not appear to be based on strong empirical evidence, since WFIs and PWFs turn out to have cardinal properties (cf. section 4).

The results obtained with respect to the preference formation theory suggest that at the very least a cardinal, individually measurable utility concept substantially facilitates the investigation of a number of empirical and theoretical problems. In addition, new light is shed on a number of policy issues. Remember, for example, the use of WFIs to construct family income equivalence scales.

The preference formation theory, which is a natural extension of the probability-like nature of the individual welfare function, should have dramatic implications for various parts of economic theory where, hitherto, utility functions are taken as constant and independent. Some of these implications were sketched in some detail by Layard (1980).

Hitherto, most attention has been directed towards the measurement and explanation of IWFs. No attention has been given to the use of WFIs and PWFs as predictors of behavior. The only attempt in this direction is by Kapteyn et al. (1979) were cardinally measured IWFs have been used to test alternative behavioral hypotheses.

Most of the policy issues discussed in section 9 have been or can be dealt with by means of revealed preference approaches. The theoretical basis for these approaches is given by utility theory. The revealed preference approach is then used to obtain empirical evidence on utility functions and, next, to base policy analysis on that. We feel that it may be equally justified, and often easier to implement, to measure utility directly, and use the results in policy analysis. There does not appear to be any reason why indirectly measured utility functions, via the revealed preference approach, would be a more solid base for policy than directly measured utility functions like IWFs (Wansbeek and Kapteyn 1983). Ideally, of course, one would hope that both modes of measurement lead to the same conclusion.

References

Antonides, G., A. Kapteyn and T.J. Wansbeek, 1980. Reliability and validity assessment of ten methods for the measurement of individual welfare functions of income. Working paper, Modelling Research Group, University of Southern California, Los Angeles.

Buyze, J., 1982. The estimation of welfare levels of a cardinal utility function. European Economic Review 17, 325–332.

Colasanto, D., A. Kapteyn and J. van der Gaag, 1984. Two subjective definitions of poverty: results from the Wisconsin Basic Needs Study. Journal of Human Resources 19, 127–138.

Dagenais, D.L., 1977. Evaluating public goods from individual welfare functions. European Economic Review 9, 123–149.

Danziger, S., J. van der Gaag, E. Smolensky and M. Taussig, 1984. The direct measurement of welfare levels: how much does it cost to make ends meet? The Review of Economics and Statistics 64, 500–505.

Deaton, A. and J. Muellbauer, 1980. Economics and consumer behavior. Cambridge: Cambridge University Press.

Duncan, O.D., 1975. Does money buy satisfaction? Social Indicators Research 2, 267–274.

Easterlin, R.A., 1974. 'Does economic growth improve the human lot? Some empirical evidence'. In: P.A. David and M.W. Reder (eds.), Nations and households in economic growth. Essays in honor of Moses Abramowitz. New York: Academic Press. pp. 89–105.

Fishburn, P.C., 1970. Utility theory for decision making. New York: Wiley.

Goedhart, T., V. Halberstadt, A. Kapteyn and B.M.S. van Praag, 1977. The poverty line: concept and measurement. Journal of Human Resources 12, 503–520.

Gorman, W.M., 1956. The demand for related goods: a possible procedure for analyzing quality differentials in the egg market. Iowa Agricultural Experiment Station.

Hagenaars, A.J.M. and B.M.S. van Praag, 1985. A synthesis of poverty line definitions. Review of Income and Wealth 31, 139–153.

Kapteyn, A., 1977. A theory of preference formation. Dissertation Leyden University, Leyden.

Kapteyn, A., 1979. A theory of preference formation. Modelling Research Group, University of Southern California, Los Angeles.

Kapteyn, A. and V. Halberstadt, 1980. Il concetto di reditto minimo. [The notion of minimum income.] Revista Internazionale di Scienze Sociali, 183–219.

Kapteyn, A. and F.G. van Herwaarden, 1980. Interdependent welfare functions and optimal income distribution. Journal of Public Economics 14, 375–397.

Kapteyn, A. and B.M.S. van Praag, 1976. A new approach to the construction of family equivalence scales. European Economic Review 7, 313–335.

Kapteyn, A. and B.M.S. van Praag, 1980. 'Family composition and family welfare'. In: J.L. Simon and J. daVanzo (eds.), Research in population economics, Vol. II. Greenwich, CT: JAI Press. pp. 77–97.

Kapteyn, A. and T.J. Wansbeek, 1982. Empirical evidence on preference formation. Journal of Economic Psychology 2, 137–154.

Kapteyn, A., B.M.S. van Praag and F.G. van Herwaarden, 1976. Individual welfare functions and social reference spaces. Economic Institute, Leyden University, Leyden. Abstracted in: Economics Letters 1, 173–178.

Kapteyn, A., S.A. van de Geer, H. van de Stadt, 1985. 'The impact of changes in income and family composition on subjective measures of well-being'. In: M. David and T. Smeeding (eds.), Horizontal equity, uncertainty, and economic well-being. Chicago, IL: University of Chicago Press. (Forthcoming.)

Kapteyn, A., T.J. Wansbeek and J. Buyze, 1979. Maximizing or satisficing? The Review of Economics and Statistics 61, 549–563.

Kapteyn, A., T.J. Wansbeek and J. Buyze, 1980. The dynamics of preference formation. Journal of Economic Behavior and Organization 1, 123–157.
Lancaster, K.J., 1971. Consumer demand: a new approach. New York: Columbia University Press.
Layard, R., 1980. Human satisfactions and public policy. Economic Journal 90, 737–749.
Pollak, R.A. and T.J. Wales, 1982. Demographic variables in demand analysis. Econometrica 49, 1533–1551.
Samuelson, P.A., 1945. Foundations of economic analysis. Cambridge, MA: Harvard University Press.
Scitovsky, T., 1976. The joyless economy. Oxford: Oxford University Press.
Sen, A.K., 1974. Informational bases of alternative welfare approaches. Journal of Public Economics 3, 387–403.
Van de Stadt, H., A. Kapteyn and S. van de Geer, 1985. The relativity of utility: evidence from panel data. The Review of Economics and Statistics 67, 179–187.
Van Herwaarden, F.G. and A. Kapteyn, 1981. Empirical comparison of the shape of welfare functions. European Economic Review 15, 261–286.
Van Herwaarden, F.G., A. Kapteyn and B.M.S. van Praag, 1977. Twelve thousand individual welfare functions of income: a comparison of six samples in Belgium and The Netherlands. European Economic Review 9, 283–300.
Van Praag, B.M.S., 1968. Individual welfare functions and consumer behavior. Amsterdam: North-Holland.
Van Praag, B.M.S., 1971. The welfare function of income in Belgium: an empirical investigation. European Economic Review 2, 337–369.
Van Praag, B.M.S., 1975. 'Utility, welfare and probability: an unorthodox economist's view'. In: D. Wendt and C. Vlek (eds.), Utility, probability and human decision making. Dordrecht: Reidel. pp. 279–295.
Van Praag, B.M.S., 1976. 'The individual welfare function and its offspring'. In: J.S. Cramer, A. Heertje, P. Venekamp (eds.), Relevance and precision: essays in honour of Pieter de Wolff. Amsterdam: North-Holland. pp. 263–293.
Van Praag, B.M.S., 1977. The perception of welfare inequality. European Economic Review 10, 189–207.
Van Praag, B.M.S., 1978. 'The perception of income inequality'. In: W. Krelle and A.F. Shorrocks (eds.), Personal income distribution. Amsterdam: North-Holland. pp. 113–136.
Van Praag, B.M.S. and A. Kapteyn, 1973. Further evidence on the individual welfare function of income: an empirical investigation in The Netherlands. European Economic Review 4, 33–62.
Van Praag, B.M.S. and J. Linthorst, 1976. Municipal welfare functions. Regional Science and Urban Economics 6, 51–79.
Van Praag, B.M.S., T. Goedhart and A. Kapteyn, 1980. The poverty line: a pilot survey in Europe. The Review of Economics and Statistics 62, 461–465. (A more extensive version, presenting the result quoted in section 7.4, appeared as vol. 23 in the series Statistical Studies, Staatsuitgeverij, The Hague, 1979.)
Van Praag, B.M.S., A. Kapteyn and F.G. van Herwaarden, 1978. The individual welfare function of income: a lognormal distribution function! European Economic Review 11, 395–402.
Van Praag, B.M.S., A. Kapteyn and F.G. van Herwaarden, 1979. The definition and measurement of social reference spaces. The Netherlands Journal of Sociology 15, 13–25.
Van Praag, B.M.S., A.J.M. Hagenaars and H. van Weeren, 1982a. Poverty in Europe. Review of Income and Wealth 28, 345–359.
Van Praag, B.M.S., J.S. Spit and H. van de Stadt, 1982b. A comparison between the food ratio poverty line and the Leyden poverty line. The Review of Economics and Statistics 64, 691–694.
Wansbeek, T.J. and A. Kapteyn, 1983. Tackling hard questions by means of soft methods: the use of individual welfare functions in socio-economic policy. Kyklos 36, 249–269.
Wierenga, B., 1978. The individual welfare function of income: a lognormal distribution function? European Economic Review 11, 387–393.