Journal of Economic Psychology 2 (1982) 137–154
North-Holland Publishing Company

EMPIRICAL EVIDENCE ON PREFERENCE FORMATION *

Arie KAPTEYN
University of Southern California, USA

Tom WANSBEEK
The Netherlands Central Bureau of Statistics, The Netherlands

Received August 31, 1981; accepted February 3, 1982

To explain the empirical fact that different individuals have different individual welfare functions, a theory of preference formation has been developed by the first author. The theory essentially states that an individual's welfare function is *identical* to the distribution of consumption patterns the individual has observed over time. This includes both his own consumption and the consumption by others in his social reference group. The paper reviews the evidence collected with respect to directly measured individual welfare functions in Europe. Evidence obtained in related research by economists, psychologists, and sociologists is also discussed.

1. Introduction

The Preference Formation (PF) theory to be discussed in this paper is presumably as obvious to sociologists and psychologists as it is alien to economists. Still, to say that it is obvious is not to say it is trivial; to say it is alien to economists is not to say it is irrelevant to economics. These observations provide the two main themes of this paper. First, stating the theory with a reasonable degree of exactness brings out implications that seem to have gone unnoticed hitherto. Secondly, economists' almost universal neglect of preference formation has led them to construct theories that are in certain respects altogether unrealistic.

* The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the Netherlands Central Bureau of Statistics or the views of other CBS staff members. The authors are grateful to Wouter Keller, the referees, and the editor for their comments, and to Richard Easterlin for guiding us to a number of references.

Mailing address: T.J. Wansbeek, Centraal Bureau voor de Statistiek, Hoofdafdeling Statistische Methoden, Postbus 959, 2270 AZ Voorburg, The Netherlands.

0167-4870/82/0000-0000/$02.75

The authors are economists. Consequently, the discussion will be geared towards economic modelling and economic policy. We will not resist the temptation, however, to quote examples from related disciplines that were brought to our attention.

In section 2 we will explain the PF theory in an informal way but yet, hopefully, in sufficient detail to convey the basic notions. In section 3 we quote some evidence from psychology and sociology. Section 4 summarizes quantitative evidence obtained in a more systematic fashion within the Leyden Income Evaluation Project. To illustrate how a quantitative theory yields sharper policy conclusions than a qualitative one, some exercises with the theory are sketched in section 5. To motivate why economists should be interested in this kind of theory, a discussion of its relevance to certain parts of economics follows in section 6. Section 7 concludes the paper with an appraisal of the empirical status of the PF theory and its promises.

2. The theory

"Happiness" is not a very well-defined concept. In practice it is generally defined in an operational way, e.g. by asking people to rate their own happiness and by taking the response as a measure of happiness (cf. Bradburn 1967: Ch. 3). We shall use the terms "happiness" and "well-being" interchangeably. The original meaning of the word "utility" was probably quite close to the concepts of happiness and well-being. Over the years the meaning has been diluted and nowadays economists merely consider utility as an *index* (cf., e.g., Phlips (1974), Deaton and Muellbauer (1980)). Utility has become an ordinal concept, indicating that a number representing the utility of a certain commodity bundle is more or less arbitrary, the only condition being that if the utility of commodity bundle A is given by the number n_A and the utility of commodity bundle B is given by the number n_B, then $n_A > n_B$ if and only if A is preferred to B.

The assumed ordinal nature of utility is *not* based on any empirical evidence implying that individuals are not able to rate commodity bundles on an interval scale. Ordinality has been assumed because the economic theory of choice does not require anything stronger than ordinal utility. Consequently, economists have invoked Occam's razor to dispose of cardinal utility. However, upon introspection it appears

reasonable to suppose that if an individual prefers commodity bundle A to commodity bundle B he will actually be happier when he consumes A than when he would consume B. Since there is no compelling empirical evidence against cardinal utility and because economists discover an increasing number of choice situations where only a cardinal utility concept can describe choices (the most famous one having to do with decisions under uncertainty, cf. Von Neumann and Morgenstern (1944)) we feel free to consider utility, well-being, welfare and happiness as rough equivalents. That is, empirical evidence concerning one of these concepts is taken to have a bearing on the other concepts as well.

In this paper we discuss empirical evidence regarding the determinants of utility. Since, as said, a utility function is a representation of one's preferences, this is equivalent to saying that we are concerned with empirical evidence on preference formation. In this section the theory is introduced by means of some suggestive examples. A more formal statement has been given in Kapteyn (1977).

In a well-known paper, Easterlin (1974) reports on a cross-country comparison of self-ratings of happiness. It turns out that there is no significant correlation between the average rating per country and its per capita national income. Within each country, however, the self-ratings increase with the respondents' incomes. The result has been explained, by Easterlin himself and by others, by hypothesizing that happiness doesn't have so much to do with one's *absolute* income level as with the ratio between one's income level and the average income level in one's reference group.

In slightly more formal terminology:

$$U_n = f\left[\frac{y_n}{\sum_{k=1}^{N} w_{nk} y_k}\right], \tag{1}$$

where U_n is the self-rating of happiness ("well-being") of Individual n ($n = 1, \ldots, N$; i.e. there are N individuals in society), y_n is the income level of Individual n and w_{nk} is the *reference weight* assigned by Individual n to Individual k. The reference weights are normalized to sum to unity; $\Sigma_{k=1}^{N} w_{nk} = 1$, for each n. The function f is presumably non-decreasing. This formulation was suggested by Easterlin (1974:

112), except that we have replaced total consumption C_k by after tax income y_k. Easterlin's formulation is a specialization of the original suggestion by Duesenberry (1949: 35).

Obviously, happiness is a function of more variables than just income (like health, leisure, family relations, equity evaluations, expectations, work conditions). We could represent those by adding arguments to f. For simplicity these complications are mostly ignored. But at the end of this section we will briefly discuss the possible impact of this neglect.

Of course, the hypothesis gives a beautiful explanation of Easterlin's findings, although it is not the only possible explanation (see, e.g., Easterlin (1974), Abramowitz (1979), Hirsch (1977)). For reasons of space we will not discuss the alternative explanations (for one thing, the alternative explanations are not capable of explaining certain other empirical facts presented below). Rather we want to take the basic idea for granted and elaborate upon it.

First, what is the functional form of f? Is f the same for each individual? Notice, incidentally, that for the ordinalists among us these questions do not make much sense. To them, (1) only says that more (relative) income is preferred to less. If it were true that (1) has only ordinal significance, however, then Easterlin's exercise could not have been carried out. There simply would be no such thing as a numerical self-rating of happiness. The mere possibility of Easterlin's study therefore strongly supports a cardinal viewpoint.

Secondly, (1) refers to a particular period of time. Would it not be reasonable to assume that one's utility does not only depend on today's income relative to the mean income in today's reference group, but also on yesterday's mean reference group income, say? Note, incidentally, that w_{nn} is one of the N reference weights, so that the influence of Individual n's own income (both today's and yesterday's) is automatically accounted for.

Thirdly, societies happen to consist of families rather than individuals. One would think that in the computation of the denominator in (1), each income y_k should somehow be deflated by the size of the kth family sharing the income.

The three questions posed here indicate on the one hand that (1) may be a little too simple to capture the determinants of individual utility, whereas on the other hand we would like to be more specific as to the form of the function f. Below we describe a theory which addresses these points.

First consider the function f. To fix ideas, let us take an individual whose income is twice the mean income of his reference group. Now imagine two situations. In the first situation everyone in his reference group has roughly the same income, so that his income ranks at the very top of the income distribution of his reference group. In the second situation we assume that there is a wide disperson of incomes in his reference group. Now his income is in the 70th percentile of the income distribution of the reference group. Comparing the two situations, would we predict the same reported satisfaction with income in both cases? The answer has to be negative, we believe. In the first situation the individual will give higher marks to his income than in the second situation. As a consequence, we have to conclude that f depends on the shape of the income distribution in one's reference group. Since, in principle, different individuals have different reference groups, f should differ between individuals. A very simple way to accommodate this requirement is to assume that an individual's evaluation of his income depends on the *ranking* of the income in the income distribution in the reference group.

Now the element of time. If the individual's evaluation of income depends on the ranking of his income in today's income distribution in his reference group, but also on how today's income compares to past income distributions in his reference group, then we need to combine the income distributions in different periods. Intuitively, it would seem that the income distributions at different times will get different weights, presumably the weight given to recent income distributions being higher than those given to income distributions in a more remote past. Let us call these weights *memory weights*, denoted as a_{nt}. So a_{ni} is the weight given by Individual n to the incomes in period t, where $t = 0$ denotes the present. The memory weights can be used to combine the income distributions in different periods into Individual n's *presently perceived income distribution,* which is nothing else than a memory weighted average of these income distributions. The word "perceived" is added to indicate the subjective nature of the concept, since both the memory weights and the reference weights (defining the reference group) are subjective quantities which cause the presently perceived income distributions to be different for different individuals. An obvious generalization of the earlier idea that an income is evaluated by considering its ranking in the income distribution of the reference group is now to hypothesize that it is the ranking in the presently

perceived income distribution which influences Individual n's evaluation of a given income level.

Finally, let us take up the problem of correcting for differences in family size. It is not income itself which provides utility, it is the consumption possibilities associated with it. Thus, if a family in Individual n's reference group has a high income but also many family members, then the consumption possibilities of those members may still be fairly limited. The consumption possibilities of each family member are represented by the (after tax family) *income per equivalent adult* ("normalized income", for short). Thus it seems that it is not the presently perceived distribution of family incomes that matters, but the presently perceived distribution of normalized incomes. When the individual evaluates the income of his own family with regard to the consumption possibilities it represents, he will again take into account the number of equivalent adults sharing the income [1].

As a consequence, he evaluates his family income by standardizing first and then considering its ranking in the presently perceived distribution of standardized incomes. This does not completely answer the question how f in (1) has to be specified, because we have not said anything yet about the form of the utility function. In Kapteyn (1977), it has been stated that one's satisfaction with income, measured on a [0,1]-scale (i.e. 0 = no satisfaction, 1 = complete satisfaction) is *equal* to the rank order of the normalized income in the perceived normalized distribution. This follows from his Preference Formation (PF) theory. It implies that the utility function on the left hand side must have the mathematical form of a probability distribution function.

There are two arguments for this choice. First, Van Praag's individual welfare function of income (WFI) (1968, 1971) is a utility function with mathematical properties identical to those of a probability distribution function. In view of the simple way in which WFI's can be measured, this formulation of the PF theory lends itself for extensive testing, as we shall see. The second argument is of a more fundamental nature. In economics, utility functions are somewhat mysterious concepts. They describe people's preferences, but the origin of these

[1] One may argue that the family size decision is not entirely exogenous, i.e. that children are a consumption good. However, the difference in length of horizon between family size decisions and other consumption decisions is sufficiently large to take family size as exogenous in the evaluation of income.

preferences is unclear. One may argue, however (Kapteyn (1979)), that utility functions can only exist after a frame of reference has developed. In the present context, the frame of reference is the perceived standardized income distribution. From considerations of scientific parsimony, as well as introspection, it seems most reasonable to assume that the frame of reference is used *directly*, rather than first transformed into something else. In any case, empirical evidence should help to decide the issue. This will be taken up in section 4.

The preceding discussion was cast in terms of incomes. There is no barrier to extend the basic ideas to, for instance, consumption vectors (this was done in Kapteyn (1977)). Neither is there any obvious reason to limit the applicability of the theory to economics, as we will see in the next section.

We started the discussion of Easterlin's findings by saying that 'other' factors influencing welfare would be ignored. We did this in the belief that the theory described here at least captures one of the main mechanisms explaining utility functions. The following sections will indicate to which extent this is true. In evaluating empirical evidence it has to be assumed that the left out factors do not systematically bias the evidence in favor of the theory. Generally, one would expect left out factors to attenuate observed relationships. So, if anything, our striving for simplicity should weaken the empirical evidence rather than strengthen it.

3. Evidence from psychology and sociology

The findings of Easterlin (1974) have been confirmed by Duncan (1975). In two samples of housewives in the Detroit area in 1955 and 1971 respectively, respondents were asked to rate their own satisfaction or dissatisfaction with their standard of living. Although the median family income in the 1971 sample was 42% higher in real terms than in the 1955 sample, the distribution of responses was not significantly different between the two samples. Regression analysis shows that within each sample the satisfaction with the standard of living rises significantly with income. These findings are completely in line with those of Easterlin.

The basic notions underlying the PF theory are more common outside economics. A particularly neat example is provided by Davis

(1966). A national probability sample of graduating college men is analyzed to investigate determinants of career decisions of those graduates. It turns out that grade-point average is a far more important determinant of career plans than the selectivity of the college attended. Thus, for example, grade A students at a prestigious college and a not very selective one have basically the same career plans even although A-students at the latter college would only score C's at the former one. This indicates that students mainly restrict their reference group to other students at their own college, more or less disregarding the ranking of their college nationally.

Research by Bassis (1977) qualifies this outcome somewhat in that he finds that academic self-evaluations of college freshmen also take into account the selectivity of the college they attend, i.e., they extend their reference group to students outside their own college. However, GPA is still more important than college selectivity. Furthermore, one would expect the freshmen's reference weights to evolve over time, so that by the time of graduation they may have reference groups mainly restricted to other students at their own college.

The examples quoted so far are cross sectional in nature. Evaluations appear to depend on one's ranking in a perceived distribution in a reference group. The findings discussed can be grouped under the heading of "relative deprivation". This notion has been developed by Stouffer et al. (1949), Merton and Kitt (1950), Davis (1959), Runciman (1966), among others. An individual's relative deprivation in a relevant dimension depends on which persons he compares himself to and whether these persons score higher or lower on that dimension. For instance, if career progress is the dimension we are interested in, the individual is relatively deprived if most persons he compares himself to (i.e., people in his reference group) make faster promotions than he does. Relative deprivation is a measure of the lack of well-being. Relative deprivation theory appears to be closely connected with the PF theory, given its emphasis on comparisons with others as a determinant of well-being. The dynamic component of the PF theory is mainly lacking in the Relative Deprivation theory [2].

A theory which takes into account both the cross sectional and the dynamic aspects of preference formation (or more generally the elicita-

[2] Runciman's description of relative deprivation (1966: 10) does also include a dynamic component, but in practice this component seems to receive little attention.

tion of judgments) is the so-called Adaptation Level (AL) theory. A definition of the AL concept is given by Helson as follows: "... adaptation level is defined as a weighted geometric mean of all stimuli impinging upon the organism from without and all stimuli affecting behaviour from within" (Helson 1964: 59). The weights in the geometric mean are determined experimentally and vary per context. Among other things, the weights may depend on the amount of time elapsed since a stimulus was perceived. The AL serves as a frame of reference to judge new stimuli.

An example of how AL theory is tested in psychological experiments is provided by Helson and Kozaki (1968) (described by Helson (1971)). Four groups of 5 subjects were shown random patterns of 10, 12, 14, 16, and 18 dots exposed for 0.30 seconds. Before that, one group was shown a random pattern of 4 dots, the second group was shown 13 dots and the third group was shown 32 dots. The fourth group (the control group) was not shown anything in advance. During the experiment the subjects were asked to estimate the number of dots shown to them. It turns out that the first group gives the highest estimates and the third group the lowest, with the remaining two groups in between. The explanation is that the AL, being a weighted geometric mean of previous stimuli (in this case presumably all random patterns of dots the subjects have ever seen over their life-time) is different for the different groups, because the most recent component of the AL (the dots shown just before the experiment) is different.

The weighted mean in equation (1) can obviously be interpreted as an AL (strictly speaking, of course, AL requires a geometric mean, but that does not affect the basic point). So (1) is a special case of AL theory. This clarifies the relation between AL theory and PF theory. On the one hand AL theory is more general, because it extends to all kinds of judgments, not just evaluation of income or consumption. It is also less complete, however, in that only a weighted average of stimuli is used to define the AL. No allowance is made for the *distribution* of previous stimuli, in the way the PF theory emerged as a refinement and generalization of (1).

Relative deprivation theory and AL theory were taken here as examples of sociological and psychological theories with a bearing on preference formation. Obviously, other theories are related as well, like reference group theory (e.g., Hyman and Singer (1968)), dissonance theory (e.g. Festinger (1957)), social comparison theory (e.g. Festinger

(1954)), anchoring effects (e.g. Tversky and Kahneman (1974)), etc. Being economists, we are not able to exactly evaluate the differences and agreements between these theories. A superficial glance suggests that they reflect very similar mechanisms, which may not be understood completely yet, and of which the PF theory is just one representative.

4. Evidence involving individual welfare functions

For simplicity we will almost exclusively discuss evidence with respect to the evaluation of incomes, although also evidence has been obtained regarding the evaluation of expenditures on durables. A couple of individual welfare functions are sketched in figs. 1 and 2. An individual's WFI measures on a [0,1]-scale the satisfaction the individual would derive from any income level between zero and infinity. A WFI is described by two parameters μ and σ, which generally differ between individuals (see, e.g. Van Praag and Kapteyn (1973)). The quantity e^{μ_n} is the income which Individual n would evaluate by 0.5, on a [0,1]-scale. Thus the higher e^{μ_n} (or μ_n for that matter) is, the more income Individual n requires to attain a certain degree of satisfaction. In this respect μ_n is a measure of Individual n's wants. The parameter σ_n measures to which extent Individual n is *sensitive* to income changes or

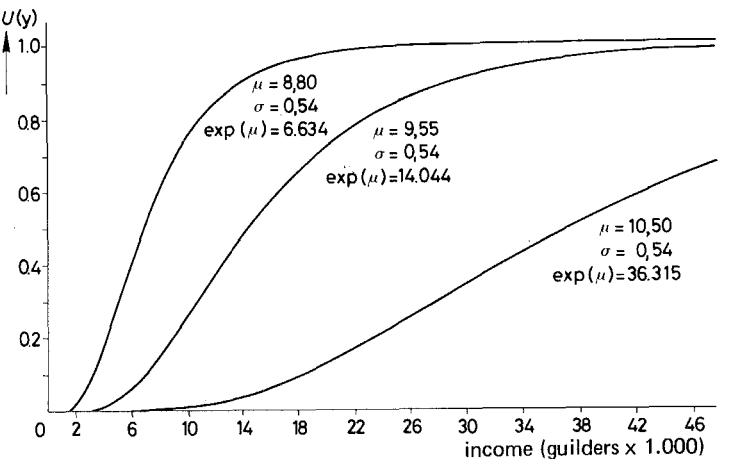

Fig. 1. The WFI for various values of μ.

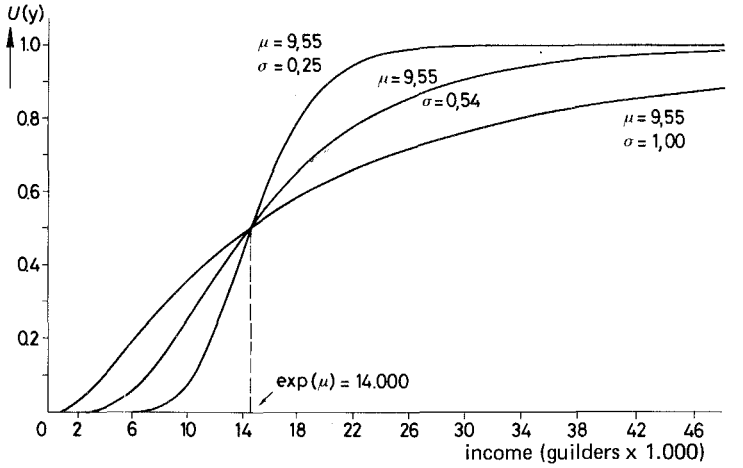

Fig. 2. The WFI for various values of σ.

income differences. If σ_n is small, a modest change of Individual n's income will have a pronounced effect on his evaluation of his income. If σ_n is large, such a change of income has only a minor effect on the evaluation of his income.

This is not the place to describe in detail how the parameters μ and σ are measured per individual (see, e.g., Van Praag (1971), Van Herwaarden and Kapteyn (1981)). Basically an individual is presented a number of verbal descriptions of satisfaction levels and asked to report which income level corresponds to each satisfaction level. This provides a number of points of the individual's WFI from which the parameters μ and σ of his WFI can be estimated by a simple regression.

Given the interpretation of μ_n and σ_n, a number of predictions can be derived from the PF theory and confronted with data:

(1) When asked to evaluate family income, individuals in larger families will tend to exhibit a larger μ, because a larger family requires a larger family income to have the same standardized income (and hence the same ranking in the perceived standardized income distribution) as a smaller family. This finding is confirmed by Van Praag (1971), Van Praag and Kapteyn (1973), Kapteyn and Van Praag (1976), among others. Invoking the economic definition of a true cost of living index, the dependence of μ on family size can be employed to construct family equivalence scales (Kapteyn and Van Praag (1976, 1980)).

(2) *Ceteris paribus,* individuals with reference groups containing relatively large families will have a lower μ than others with smaller families in their reference group. This is so, because the incidence of large families decreases the standardized incomes in the reference group, so one needs a lower income to attain a certain ranking at the perceived distribution of standardized incomes. Indeed, Van de Stadt and Kapteyn (1982) find a significantly negative effect on μ of the median family size in one's reference group.
(3) Individuals with a larger income tend to have a larger μ, because one's own income is part of the presently perceived income distribution. This finding was confirmed in the studies mentioned under (1), but also in Van Herwaarden et al. (1977).
(4) The higher the incomes in one's reference group, the higher one's μ will be. This is confirmed by Kapteyn et al. (1976, 1978).
(5) From a somewhat technical argument it can be seen that the correlation between μ and income, referred to under (3), will be strongest for people whose incomes don't change much over time. This is confirmed by Kapteyn (1977), and Van Herwaarden et al. (1977).
(6) An individual's σ will tend to be large if the dispersion of incomes in his reference group is relatively high. This is confirmed by Kapteyn et al. (1976, 1978). As slightly more indirect evidence, Van Praag et al. (1977, 1980) find in a cross-country comparison within the European Community that, on average, individuals have a significantly larger σ in countries with an unequal distribution of incomes.
(7) An individual's σ will tend to be large if his income has changed considerably over time. This was found to be true by Kapteyn (1977), and Van Herwaarden et al. (1977).

These various effects can be integrated in one quantitative model describing the influence of past and present incomes on individual μ's and σ's. Estimating such a model involves the estimation of reference weights and memory weights. The paper by Kapteyn et al. (1980) has gone farthest in this respect. Their results include an estimate of the memory weights indicating that for any practical purpose it is the events of the last seven years that determine an individual's frame of reference. These results have been replicated, by and large, by Van de Stadt and Kapteyn (1982), although their estimate of the memory

weights indicates a memory of more than seven years. Both studies indicate that about 84% of the variance in μ can be explained by the PF theory. The two papers did not employ the PF theory to explain variation in σ over individuals, so it is not known yet which percentage of its variance can be explained by the PF theory.

Given the availability of a quantitative model explaining an individual's μ and σ, one can also investigate how individual satisfaction with income can be manipulated by changing the income distribution (e.g., by introducing a family allowance system) or the growth rate of incomes. Such exercises are of obvious relevance to socio-economic policy and will be briefly considered in the next section.

5. A few exercises

The exercises sketched here come from Kapteyn (1977) and Kapteyn and Van Herwaarden (1980). We do not aim at a derivation of results, but rather want to state some outcomes and give an intuitive explanation. The most simple exercise is the following one. Imagine a society, consisting only of single person households in which everyone has had the same income for a long time. That income has been growing at a constant rate γ, for the same long time. Except for personal differences, everyone evaluates his income by the same number. How does this number depend on the growth rate γ?

The answer is: Not at all! This result appears counter-intuitive and would have been hard to predict on the basis of qualitative considerations alone. Once the result has been obtained, an intuitive explanation is of course possible: Remember from the previous section that small changes in income will lead to a small σ. If σ is small, a small increase in income each year still has a sizeable positive effect on the evaluation of that income (cf. fig. 2). On the other hand, a high growth rate leads to a large σ. Consequently, the relatively large increase in income from year to year has only a modest upward effect on the evaluation of income. Thus, there are two situations: With a low growth rate, the small σ amplifies the positive welfare effects of the modest increases in income. With a high growth rate, the large σ dampens the positive welfare effects of the sizeable increases in income. As a result the two situations yield the same level of satisfaction of income.

The exercise described here refers to an unrealistic situation, or at

least to a situation that has not materialized yet. That does not make it irrelevant. It is in the nature of policy to aim at the realization of situations that do not yet exist. The exercise illustrates that in such a case intuition may be insufficient to predict what the result of a policy will be. Only a quantitative thoroughly tested model can be of guidance in such situations. That is not to say that the PF theory is such a model. Substantially more work has to be done to assess its empirical status.

The beauty of the example considered is its independence of parameter values. In more general cases in which incomes differ between individuals, families are of different size and growth rates of incomes are possibly different, the effect of a particular income policy on any individual's satisfaction with income depends on the memory weights, reference weights, family size, etc. So then these parameters have to be known with sufficient accuracy to attain specific policy goals.

As an example, Kapteyn and Van Herwaarden (1980) use reference weights estimated by Kapteyn et al. (1976, 1978) to calculate which income distribution in The Netherlands would have maximized an additive social welfare function, which has WFI's as its argument. This analysis ignores dynamic effects as well as any feedback from the income distribution to work effort and should therefore be interpreted with caution. But the results serve to illustrate how the PF theory alters policy conclusions, or, equivalently, how a neglect of preference formation aspects may lead to policies that do not achieve their goals.

The first example given also illustrates how a more specific theory yields extra insights. If one only takes into account the adaptation of μ (which is done in AL theory), then it would seem that the faster economic growth is, the higher individual welfare. This brings out the prospect of a Hedonic treadmill (Brickman and Campbell (1971)). We need economic growth forever to keep people happy. The example illustrates that at least the treadmill doesn't have to go very fast. One may furthermore introduce the possibility of upward sloping age income profiles for every individual without an increase of average income in society. In that case everyone would experience a growth in income which, according to the PF theory, increases the satisfaction with one's income.

As a result a steady state appears to be attainable where people are reasonably happy. Given constraints stemming from the limited availability of natural resources such a policy may be optimal in the long run.

6. What has all this to do with economics?

With relatively few exceptions, economists have ignored preference formation. In quite a few areas, theories are heavily dependent on the assumption that preferences are constant and/or independent of behavior of others. Allowing for preference formation has some significant consequences.

In welfare economics, the Pareto principle (or its sister the Kaldor compensation principle) looses its practical appeal if it can no longer be assumed that an income increase for some, without an income reduction for anyone else, provides an increase in social welfare. In cost benefit analysis distributional aspects move from the periphery of the problem to the core. Results in optimal taxation change likewise. For example, in a simple model where individual utility depends on one's relative position in the income distribution and on number of hours worked, Layard (1980) derives the optimal marginal tax rate to be unity [3]! Earlier, Feldstein (1976: 81) suggested that results in the optimal taxation literature, which imply "surprisingly little redistribution through the tax-transfer process" may be due to the neglect of external effects of consumption (in particular altruism or envy).

If the casual observer (that is everyone who is not an economist) is right in asserting that an individual's consumption behavior is influenced by the consumption of others, then all economic models of consumption are misspecified. It goes without saying that this does not help the quality of econometric forecasts. Similarly, one would expect behavior on the labor market to be influenced by the behavior of others, which also suggests improvements in models of labor supply, etc.

7. Concluding remarks

The evidence on preference formation sketched in this paper is unambiguous. It uniformly supports the PF theory explained in section 2. Still, more testing is possible and should be carried out, but at the very

[3] One of the referees observes that "... any model which derives an optimal marginal tax rate of unity must be wrong!" This may be true, of course, but Layard's analysis shows how sensitive theoretical results are to assumptions on preference formation.

least the existence of the preference formation phenomenon appears to be firmly established. Consequently, economic theories that are based on the assumption of constant preferences should be considered unacceptable [4].

Admittedly, allowing for preference formation does not make economic modelling any easier. Economic models are already complicated, because economists tend to despise the possibility of direct utility measurement. Utility is measured indirectly through models which relate behavior to utility functions. If we allow for shifting utility functions, the shifts have to be incorporated in the behavioral model.

In the words of Duesenberry (1949: 17):

> Ordinarily we try to measure preference parameters (or functions of them) by market behavior, since we cannot observe the preferences directly. With shifting parameters we should be carrying indirect measurement a step farther. We would not only have to measure the preference parameters but the parameters of the relation governing shifts in the preferences.

The use of direct measures of utility, like WFI's, makes life simpler again, because now the preference formation process can be studied independently from a behavioral model, which would relate behavior to preferences. In addition, the data requirements will tend to be substantially less, because utility measures like WFI's can be measured relatively easy, whereas models of consumption, for instance, require records of consumption over an extended period of time. Likewise, in applied welfare economics, the existence of preference formation makes it almost impossible to use traditional indicators of well-being like income, because income no longer bears a simple monotonic relation with well-being. Directly measured individual utility indicators remain as the obvious alternative.

Finally, let us return to the first theme mentioned at the beginning of the paper. It is clear that the main ideas embodied in the PF theory can be found at many places in the psychology and sociology literature. The

[4] Again, as observed by a referee, although overall happiness may increase or decrease, an individual's ordering of alternative consumption bundles may remain unchanged. Since a large part of this evidence discussed deals with well-being and happiness, this would not imply that demand functions are subject to preference formation. Some of the evidence, like that concerning career choice, does deal with multidimensional choices, however, and also the preference formation plays its part. More generally, the evidence discussed in section 4 concerns a theory which is *derived* from a theory dealing with consumption bundles. The latter theory does imply that the ordering of consumption bundles will change under the influence of consumption by others or under the influence of past experiences.

thing that sets the PF theory aside is its exactness. This increases the number of tests that can be performed. In addition, the theory leads to quantitative predictions, which are more informative, both from a scientific and a policy view point.

References

Abramowitz, M., 1979. 'Economic growth and its discontents'. In: M.J. Boskin (ed.), Economics and human welfare. Essays in honor of Tibor Scitovsky. New York: Academic Press. pp. 3–21.

Appley, M.H. (ed.), 1971. Adaptation-level theory: a symposium. New York: Academic Press.

Bassis, M.S., 1977. The campus as a frog pond: a theoretical and empirical reassessment. The American Journal of Sociology 82, 1318–1326.

Bradburn, N.M., 1967. The structure of psychological well-being. Chicago: Aldine.

Brickman, P. and D.T. Campbell, 1971. 'Hedonic relativism and planning the good society'. In: M.H. Appley (ed.). pp. 287–302.

Davis, J.A., 1959. A formal interpretation of the theory of relative deprivation. Sociometry 22, 280–296.

Davis, J.A., 1966. The campus as a frog pond: an application of the theory of relative deprivation to career decisions of college men. The American Journal of Sociology 72, 17–31.

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

Duesenberry, J.S., 1949. Income, saving and the theory of consumer behavior. Cambridge, MA: Harvard 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.

Feldstein, M., 1976. On the theory of tax reform. Journal of Public Economics 6, 77–104.

Festinger, L., 1954. A theory of social comparison processes. Human Relations 7, 117–140.

Festinger, L., 1957. A theory of cognitive dissonance. Evanston, IL: Row, Peterson and Co.

Helson, H., 1964. Adaptation-level theory: an experimental and systematic approach to behavior. New York: Harper.

Helson, H., 1971. 'Adaptation-level theory: 1970 – and after'. In: M.H. Appley (ed.). pp. 5–17.

Helson, H. and A. Kozaki, 1968. Anchor effects using numerical estimates of simple dot patterns. Perception and Psychophysics 4, 163–164.

Hirsch, F., 1977. Social limits to growth. London: Routledge and Kegan Paul.

Hyman, H.H. and E. Singer (eds.), 1968. Readings in reference group theory and research. New York: The Free Press.

Kapteyn, A., 1977. A theory of preference formation. Unpublished Ph.D. thesis, Leyden University.

Kapteyn, A., 1979. A theory of preference formation. MRG Working Paper 7933, University of Southern California.

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. 2. Greenwich, CT: JAI Press. pp. 77–97.

Kapteyn, A., B.M.S. Van Praag and F.G. Van Herwaarden, 1976. Individual welfare functions and social reference spaces. Report 76.01, Economic Institute of Leyden University, Leyden.

Kapteyn, A., B.M.S. Van Praag and F.G. Van Herwaarden, 1978. Individual welfare functions and social reference spaces. Economics Letters 1, 173–178. (An abstracted version of the 1976 publication.)

Kapteyn, A., T.J. Wansbeek and J. Buyze, 1980. The dynamics of preference formation. Journal of Economic Behavior and Organization 1, 123–157.

Layard, R., 1980. Human satisfactions and public policy. The Economic Journal 90, 737–750.

Merton, R.K. and A.S. Kitt, 1950. 'Contributions to the theory of reference group behavior'. In: R.K. Merton and P.F. Lazarsfeld (eds.), Continuities in social research; studies in the scope and method of "The American Soldier". Glencoe, IL: The Free Press. pp. 40–105.

Phlips, L., 1974. Applied consumption analysis. Amsterdam: North-Holland.

Runciman, W.G., 1966. Relative deprivation and social justice. London: Routledge and Kegan Paul.

Stouffer, S.A., E.A. Suchman, L.C. DeVinney, S.A. Star and R.M. Williams, Jr., 1949. The American Soldier: adjustment during army life. Princeton, NJ: Princeton University Press.

Tversky, A. and D. Kahneman, 1974. Judgement under uncertainty: heuristics and biases. Science 185, 1124–1131.

Van de Stadt, H. and A. Kapteyn, 1982. The relativity of utility: evidence from panel data. Working paper, Netherlands Central Bureau of Statistics.

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. 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., 1981. Reflections on the theory of individual welfare functions. Paper presented at the American Statistical Association Meetings, Detroit, August 10–13, 1981.

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., T. Goedhart and A. Kapteyn, 1977. The poverty line. A pilot survey in Europe. Statistical Studies, 23. The Hague: Staatsuitgeverij.

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. (An abbreviated version of the 1977 publication.)

Von Neumann, J. and O. Morgenstern, 1944. Theory of games and economic behavior. Princeton, NJ: Princeton University Press.