European Economic Review 38 (1994) 1817–1825.

How sensible is the Leyden individual welfare function of income? A Reply

Bernard M.S. van Praag*

Foundation for Economic Research of the University of Amsterdam, Amsterdam, The Netherlands

Arie Kapteyn

CentER, Tilburg University, Tilburg, The Netherlands

Received October 1994

In this paper we react to Seidl's (1994) critique on the Leyden approach. Our conclusion is that Seidl's critique is ill-founded as a serious critique on the Leyden approach.

Key words: Utility function of income; Individual welfare function; Leyden welfare function of income

JEL classification: C90, D10, D60

1. Introduction

In a recent issue of this *Review* Seidl (1994) published a critical evaluation on the sensibility of a concept which we and others utilize a great deal, viz., the Leyden Individual Welfare Function of Income (LIFWI). Seidl concludes that 'this edifice is not built on solid ground, neither from the point of economic theory nor of experimental psychology'. It will not be surprising that we read this paper with more than usual interest. Our main conclusion is that his critique is essentially non-substantial and may be discarded.

Seidel's critique may be split up under two headings:

(a) Is the lognormal welfare function a reasonable utility function of income against the background of economic literature and evidence from experimental psychology? (see Seidl, sections 3.2.1, 3.2.2 and 4).
(b) How reasonable is the theory given in Van Praag (1968), vP68 for short, which suggests the lognormal specification? (see sections 3.2.3, 3.2.4, 3.2.5).

Due to space limitations we will focus in this paper on the issues under

Correspondence to: Prof. B.M.S. van Praag, Foundation for Economic Research of the University of Amsterdam, Roeterstraat 11, 1018 BW Amsterdam, The Netherlands.
*We are indebted to Erik Plug for his assistance.

SSDI 0014-2921(94)00089-1

(a). In a more complete version of this rejoinder (available on request from the first author) we deal with the points under (b) as well. With respect to the second set of issues let us say here only that we do not think that Seidl has given any substantial critique, and that, even if his critique of this theory on the genesis of the lognormal welfare function would have been substantial, this would not imply that the incriminated specification may not yield good and plausible results as a descriptor of economic phenomena in the sense that the empirical apparatus yields stable, non-trivial and plausible results for practical problems. After all, all functional specifications in the literature except the LIFWI are not based on an elaborate theory at all but just postulated.

However, before turning to Seidl's critique, it is appropriate to say a few words on the setting of the Leyden approach and on how the lognormal specification was derived in vP68.

2. Preliminary remarks

During the last twenty years there has been a persistent trickle of contributions which make use of a specific questioning approach to gauge the evaluation of income amounts by individuals. This is done by the so-called WFI-approach or Leyden approach. We refer to Seidl's bibliography (until 1988) for general references [for later references see e.g. Van Praag (1991), Van Praag (1994a, b), Kapteyn (1994)].

Actually, the Leyden approach may be split up into three strands:

(a) the formulation of a theory;
(b) the estimation of models suggested by theory;
(c) the applications.

In the Leyden paradigm the construction of a theory was started in vP68, followed by Kapteyn (1977) and later on much theory has been formulated jointly with estimations and applications in e.g. Hagenaars (1986), etc. The stage under (b) begins in Van Praag (1971), Van Praag and Kapteyn (1973) and Kapteyn and Van Praag (1976). In these articles a practical subjective question module is proposed to catch a hitherto unmeasurable concept, viz., a welfare function of income. Here we use the lognormal model and we try to estimate its parameters and to explain the parameters by objectively measurable characteristics.

Under (c) we apply the model, for instance, to the derivation of family equivalence scales [see Van Praag and Kapteyn (1973) and Kapteyn and Van Praag (1976) and Kapteyn (1994)]. the estimation of social reference groups [see Van de Stadt et al. (1985)], climate equivalence scales [see Van Praag (1988)], poverty concepts [see Goedhart et al. (1977), Hagenaars and Van Praag (1985)], etc. There we estimate for instance the equation

$$\mu=\beta_0+\beta_1 \ln fs+\beta_2 \ln y_c \tag{1}$$

where fs and y_c stand for family size and current income of the respondent. The parameter μ is a welfare parameter to be explained later on. We and others found in scores of large-scale samples rather stable coefficients.

3. Theoretical foundations

The main gist of vP68 cannot be deduced from Seidl's fragmentary account. The book which was published in 1968, hence more than 26(!) years ago, was aimed as an attempt to revitalize the (cardinal) utility concept. In those days the utility concept was almost eliminated as a concept to describe consumer behaviour in favour of the (revealed) preference approach. Our point of departure was that nearly all individuals feel able to and do evaluate their own situation or external phenomena in relative terms by positioning the situation somewhere between a 'worst' situation and a 'best' situation, hence on a finite interval scale. An example is the ladder-scale devised by Cantril (1965), by which people are asked to evaluate their own life situation on a zero–ten scale. Experience in hundreds of surveys shows that respondents are able to do this. With respect to phenomena like the quality of a music or ballet performance, the quality of a product or ratings of school performances, numerical evaluation seems a regular and common procedure. Numerical evaluation is performed as a matter of routine and always within the framework of a bounded interval.

Hence, if a utility[1] function $U(x)$ is a descriptor of an evaluation of a commodity bundle x in $\mathbb{R}^n$, it is also an application of this general observation to assume that $U(x)$ is a bounded function on a finite interval, say $[0,1]$. This triggered the idea to see utility as a normed measure W on the commodity space. The utility of a commodity bundle x is then the W measure assigned to the rectangular block $[0,x]$ in $\mathbb{R}^n$. This led to the idea to exploit the isomorphism between the probability model where the probability mass P is spread over $\mathbb{R}^n$ and this welfare measure model where a welfare mass W is spread over the commodity space. In that case the separate commodity dimensions, denoted by $X_1,\ldots,X_n$, take the role of random variables in the probability model and the rectangular block is described as a set $\{X \Leftarrow x\}$ to which we assign a welfare measure $W(\{X \Leftarrow x\})=U(x)$. Notice that in probability the basic set is a point in $\mathbb{R}^n$ with its associated mass density function, while in the commodity context the basic set is the block $\{X \leqq x\}$ with the corresponding welfare value $U(x)$.

One of the results in vP68 is that by employing the analogy with probability theory and more specifically the Central Limit Theorem one can

[1]We shall not discriminate between utility and welfare in this paper.

show that under a specific set of assumptions a money amount y spent on a specific broad commodity group will be evaluated approximately by a lognormal distribution function $\Lambda(y;\mu,\sigma)$ and this result carries over to more dimensions. This result is derived by assuming that commodities according to Lancaster (1966) may be translated into bundles of characteristics, and that those characteristics are defined by the human mind in such a way that they can be evaluated independently of each other. Examples are found in consumer tests like in 'Which' where for instance the performance of a PC is decomposed into k characteristics like keyboard, monitor, storage, fast memory, speed, etc. Assuming that a money amount y stands for specific combinations of characteristics, it is shown that as a first approximation the welfare value of an amount y spent on, say, a PC follows a lognormal distribution function.

In vP68 it is made perfectly clear that this is an approximation, but it is well-known from probability theory that the convergence to the normal distribution is remarkably good even for modest values of k. In vP68 it is also stressed that this preliminary evaluation of money amounts, which is necessary for distributing one's income over various expenditure categories, exaggerates the welfare value of y; later on, while actually spending the amount y, the buyer is either obliged to compromise if one spends y, or to spend more than y in order to get everything expected at the preliminary stage.

These assumptions are of a psychological nature. The interested reader is referred to vP68 rather than to the fragmentary resume of Seidl to get an idea about the generality and the caveats. We observe that just as for the Central Limit Theorem the lognormality result can be derived under various and also more general conditions which are not spelled out in vP68. We also notice that the psychological (basically simplifying) assumptions on individual evaluation behaviour need not to be reflected in the consciousness of the individual. It is here that we may draw the analogy with the profit-maximizing entrepreneur who does not know differential calculus or the motor driver who does not know the laws of mechanics when he makes a turn. Their behaviour (and evaluation is also behaviour) can be described by a model, which the actor himself does not know but which he may apply subconsciously.

4.1 Seidl's criticisms

4.1. Boundedness of the LIWFI from below and from above

The first point which is the subject of Seidl's wrath is the fact that we assume that a utility function of income is bounded both from below and above. This boundedness would not be supported by convincing economic and psychological arguments. As we saw already, from survey practice it is quite clear that evaluation on an unbounded scale by respondents is rarely if ever used. Hence, from general psychology there cannot be much support.

There is also nothing in the ordinal theory of consumer behaviour which excludes a bounded utility function of income.

Furthermore, if we accept that uncertainty is a feature of any economic performance, decisions under uncertainty are analyzed with the help of utility functions. Consider a function with constant *absolute risk aversion*, viz. $U(x)=1\exp(-ax)$. It is bounded between zero and one. The utility function with constant *relative risk aversion*, i.e., $U(x)=1-x^{-a}$ $(x>1,\ a>0)$ is also bounded. These functions are regularly used in the literature on decisions under uncertainty [see Arrow (1971)].

Samuelson (1977) states on boundedness: Menger's analysis has alerted writers like Arrow (1971) to the fact that, in order to be able to have a complete ordering of all probabilistic outcomes of options in terms of expected utilities[2] ... *It turns out, somewhat surprisingly, to imply that the utility function must be bounded: we cannot have utilities which are indefinitely large in either the positive or negative direction.* It appears that professor Seidl in 1994 is still surprised.

Professor Seidl is also confusing boundedness with finite satiation. Functions may tend asymptotically to their upper bound while marginal utility is positive everywhere. Seidl takes refuge to the authority of Debreu (1959), by quoting him that there is no finite satiation possible. However, this is just an assumption in Debreu's general equilibrium model with no empirical underpinning. In later studies on general equilibrium this assumption has been weakened of course.

4.2. The sigmoid shape of the LIWFI

Seidl criticizes the idea that a utility function of money can have a convex–concave shape. His gut-feelings are based on Gossen's First Law which is just an intuitively based proposition. According to ordinal consumer theory a utility function is preferably quasi-concave and this does not preclude that a utility function of money may be convex over a certain income range. Although economists frequently assume Gossen's Law to hold everywhere without discussion, for the extremely poor Gossen's Law might fail as every extra dollar brings them closer to survival. Hence, the utility function would be convex shaped until an individual exceeds the situation of being poor. Another argument for possible convexity comes from Friedman and Savage (1946) who argue that gambling behaviour or risk-lovingness can

[2]Samuelson quotes Arrow.

only be explained by assuming an S-curve [see also Markowitz (1952) and Kahnemann and Tversky (1979)].

4.3. Style

In section 3.2.6 but actually in the whole paper our colleague utilizes rather strong and suggestive language, frequently calling our ideas 'absurd'. He attacks the style in which vP68 has been written as 'extremely tedious reading due to its multifarious elusiveness', 'messy', and he even suggests that 'the reader was to be lulled into the impression that ...'. It is up to the reader to agree or disagree, how readable this book is after 26 years. However, by doing so, Professor Seidl invites us to say a few words on his own style. We are afraid that our colleague utilizes the same style figures which he so abhors in vP68. Especially, when one according to our opinion does not succeed in substantiating rather strong claims, it would have seemed wise to choose a more modest approach.

5. The good empirical fit of the LIWFI

Under this heading professor Seidl throws his final spear at our work. As the discussion would be incomprehensible otherwise, let us first describe in a few sentences the measurement instrument used. As Seidl departs from a now old-fashioned version (1971) and notation we shall also use it here. The so-called Income Evalution Question (IEQ) runs as follows:

> '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/month/year:

excellent	if it were above	\$..........	
good	if it were between	\$..........and	\$..........
amply sufficient	if it were between	\$..........and	\$..........
sufficient	if it were between	\$..........and	\$..........
burely sufficient	if it were between	\$..........and	\$..........
insufficient	if it were between	\$..........and	\$..........
very insufficient	if it were between	\$..........and	\$..........
bad	if it were between	\$..........and	\$..........
very bad	if it were below	\$..........	

The resulting boundaries (answers) are denoted by $y_1, \dots, y_n$ respectively. If we assume that the respondent attempts to inform us as well as possible about the shape of his welfare function, we assume that he minimizes the average inaccuracy of his answers, given that he has to describe a continuous phenomenon in terms of discrete categories. Let the average inaccuracy be described by

$$\sum_{i=0}^{h-1} \int_{y_i}^{y_{i+1}} [U(y)-\tfrac{1}{2}(U(y_i)+U(y_{i+1}))]^2 \, \mathrm{d}U(y) \tag{2}$$

where the inaccuracy contribution of each interval is weighted by the welfare mass in that interval and y_0 equals zero and y_{n+1} reaches infinity. Given that $U(y_0)=0$ and $U(y_{n+1})=1$, minimizing yields the solution known as the Equal Quantile Assumption, which mathematically reads as

$$U(y_{i+1})-U(y_i)=1/n \quad \text{for } i=0,\dots,n-1. \tag{3}$$

This reasoning was first presented in vP68 (p.57) and later on in the empirical context of Van Praag (1971). Kapteyn (1977) showed that other functions symmetric about zero yield the same result. Empirical evidence on the Equal Quantile Assumption was also presented in Van Praag (1991, 1994a, b), where it is described how people were asked to place the verbal labels used in the IEQ on a zero–hundred scale. Not only did different people put the same words roughly at the same position, also it turned out that the equal interval assumption was a reasonable (not perfect) approximation to the observed positions chosen by the experimental subjects.

Section 4 of Seidl's paper is very confusing, but let us try to get the gist out of it. Seidl departs from the idea that every individual evaluates income y by the *same 'true' or proper* welfare function $V(y)$. In contrast to that, the main point of the Leyden approach is that each individual has his/her *own individual* welfare function, which depends on the individual's own circumstances, like family size and current income y_c. This function $U(y;\, y_c,\, fs)$ may be assessed for each individual by the method sketched above. From those estimates we may derive that an individual evaluates his *own current income* by $U(y_c;\, y_c,\, fs)$. Ignoring fs we may write $V(y_c)$ for that function and this seems to be the concept Seidl has in mind as *the welfare function 'proper'*. We notice that $V(.)$ is *functionally specified* as soon as we know $U(.; y_c)$ for all y_c.

Now Seidl starts at the other side and stipulates from psychophysical evidence that the function $V(y)$ should be either logarithmic $\ln(y)$ or a power function. If our estimates $U(y\!: y_c)$ do not conform to the logarithmic or the power function, it would imply according to Seidl that $U(.)$ is a transform, say, $U(.;y_c)\equiv \Psi(V(.);y_c)$, where the transform varies with individual parameters, e.g. y_c.

It is obvious that for any exogenously specified monotonic function $V(.)$ we may find a function Ψ which performs this trick. This is however no proof that $V(.)$ is the 'true' welfare function. In particular, it is no evidence whatsoever that a logarithmic or a power function should be the 'true' welfare function $V(.)$.

On the other hand, as we just showed, it is possible from our empirical estimates to find a unique function $V(y_c)$, which happens to be a lognormal distribution function as well and not a logarithm or a power function.

Is there then support from economics for those specifications? As Arrow's result on bounded utility (also stated by e.g. Von Neumann and Morgenstern and Savage) implies, both functional specifications are inadmissible from the viewpoint of the theory of decision under uncertainty (or the dynamic theories of growth, savings, etc.).

So Seidl's evidence is only the evidence from experimental psychophysics. Let us see what the experiments of his 'good company' are. That evidence is much less firm than Seidl suggests. For instance, it is true that Stevens (1957, 1973) and his 'numerous collaborators' did 'innumerable experiments' on the subjective perception on sound intensity, brightness of light, etc., but their experiments took place in a laboratory setting and applied to a fairly narrow range of stimuli. Their studies allow only for conclusions for those narrow ranges. They found an approximate double-log relationship between subjective evaluations and stimuli for specific regions but their statistical methodology was rather shaky. The same holds for the case where such experiments were applied to the attitude towards money, presented by Galanter (1962). These authors do not present any theory for their experimental observations, nor do they account for individual differences. Although the experimental data in some instances were fitted to other functional specifications, as a rule no standard errors were presented and formal comparisons of goodness of fit between different specifications are conspicuously absent. The only thing is that the logarithm and/or power were proposed as not implausible empirically fitting relationships for a certain range. However, there are no 'laws' at all, only empirical regularities.

It is tempting to quote Samuelson (1977) again on 'classical logarithmic utility'. After saying that Bernoulli proposed the logarithmic utility function Samuelson goes on:

> *'few economists will today swallow this gratuitous postulate, but to novitiates such argumentation apparently carries some plausibility (as with the Weber–Fechner law in experimental psychology)'.*

This does not mean that these experiments in the fifties and sixties are without value. It does only mean that it is not the God-given measuring-rod against which our results can be discarded as trivial and nonsensical or a mere repetition of earlier results.

6. Concluding remarks

In this paper we refuted Seidl's principal points of critique as ill-founded. Obviously, his paper including the 103 footnotes abounds of other minor points which may be addressed as well. We abstain from doing so in order to allow the reader to keep the broader perspective in view. Despite major disagreements with respect to Seidl's arguments, we will emphasize, as empirical scientists, that we are certainly not 100% sure that our results are not prone to criticism or that they will have to be discarded once in favour of a better theory and for more convincing results. In this instance, however, we see no reason at all.

References

Arrow, K.J., 1971, Essays in the theory of risk bearing (North-Holland, Amsterdam).

Cantril, H., 1965, The pattern of human concern (Rutgers University Press, New Brunswick NJ).

Debreu, G., 1959, Theory of value (Wiley, New York).

Friedman, M. and L.J. Savage, 1948, The utility analysis of choices involving risk, Journal of Political Economy 56, 279–304.

Galanter, E., 1962, The direct measurement of utility and subjective probability, American Journal of Psychology 75, 208-220.

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

Hagenaars, Aldi. J.M. 1986, The perception of povery (North-Holland, Amsterdam).

Hagenaars, A.J.M. and B.M.S. van Praag, 1985, A synthesis of poverty line definitions, Review of Income and Wealth 31, 139–154.

Kahneman, D. and A. Tversky 1979, Prospect theory: An analysis of decision under risk, Econometrica, 47, 263–291.

Kapteyn, A., 1977, A theory of preference formation, Ph.D. (Leyden University, Leyden).

Kapteyn, A., 1994, The measurement of household cost functions: Revealed preference versus subjective measures, Journal of Population Economics, 7, forthcoming.

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

Lancaster, K.J., 1966, A new approach to consumer theory, Journal of Political Economy 74, 132–157.

Markowitz, H.M., 1952, The utility of wealth, Journal of Political Economy, 60, 151–158.

Samuelson, P., 1977, St. Petersburg paradoxes: Defanged, dissected and historically described, Journal of Economic Literature, 15, 24–55.

Seidl, Chr., 1994, How sensible is the Leyden individual welfare function of income?, European Economic Review 38, no. 8, 1633–1659.

Stevens, S.S., 1957, On the psychophysical law, Psychological Review 64, 153–181.

Stevens, S.S., 1975, Psychophysics, introduction to its perceptual, neural and social prospects (Wiley, New York).

Van de Stadt, H., A. Kapteyn and S.A. van de Geer, 1985, The relativity of utility: Evidence from panel data, The Review of Economics and Statistics, 67, 179–187.

Van Praag, B.M.S., 1968, Individual welfare functions and consumer behavior. A theory of rational irrationality (North-Holland, Amsterdam).

Van Praag, B.M.S., 1971, The welfare function of income in Belgium, European Economic Review, 2, 337–369.

Van Praag, B.M.S., 1975, Utility, welfare and probability, in: D. Wendt and C. Vlek, eds., Utility, probability and human decision making (Reidl, Dordrecht) 279–295.

Van Praag, B.M.S., 1988, Climate equivalence scales, European Economic Review, 32, 1019–1024.

Van Praag, B.M.S., 1991, Ordinal and cardinal utility: An integration of the two dimensions of the welfare concept, Journal of Econometrics 50, nos. 1/2, Annals, 69–89.

Van Praag, B.M.S., 1994a, The relativity of the welfare concept, in: M. Nussbaum and A. Sen, eds., The quality of life (The Clarendon Press, Oxford) 362–385.

Van Praag, B.M.S., 1994b, Ordinal and cardinal utility, in: R. Blundell, I. Preston and I. Walker, eds., The measurement of household welfare (Cambridge University Press, Cambridge).

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.