5

THE ECONOMIC AND SOCIAL RESEARCH INSTITUTE

AN INTERDISCIPLINARY APPROACH TO THE MEASUREMENT OF UTILITY OR WELFARE

Fifth Geary Lecture, 1972

JAN TINBERGEN

THE ECONOMIC AND SOCIAL RESEARCH INSTITUTE

AN INTERDISCIPLINARY APPROACH TO THE MEASUREMENT OF UTILITY OR WELFARE

JAN TINBERGEN

Fifth Geary Lecture, 1972

Copies of this paper may be obtained from
The Economic and Social Research Institute, 4 Burlington Road, Dublin 4

Price £0.50
(*Special rate for students £0.25*)

Jan Tinbergen is Professor of Development Planning at The Netherlands School of Economics, Rotterdam, and a Director of the Netherlands Economic Institute.
This Paper has been accepted for publication by ESRI, which is not responsible for either the content or the views expressed therein.

Printed by Cahill & Co. Limited, Parkgate Street, Dublin 8.

*An Interdisciplinary Approach to the Measurement of Utility or Welfare**

1. *Nature of this Essay*

(*a*) The prevailing opinion of economists today on the utility function is one of *agnosticism*. Usually it is held that utility cannot be measured. This is comparable to the attitude of physicists *vis-à-vis* heat before the concepts of temperature and quantity of heat were introduced. The idea that utility cannot be measured is unsatisfactory. It is unsatisfactory because utility is the basic concept of economic theory and represents something that should enter into the set of aims of socio-economic policy. In particular, comparison of the utility experienced by various groups in a population is desirable once we admit that we must aim at some type of social optimality or of social justice. We must consider present-day practice concerning these matters as pre-scientific and try to deal with them in a scientific rather than an intuitive way. Today's practice is, at best, an attempt by politicians to compare intuitively the utilities experienced by various groups and attain as satisfactory a situation as possible, especially with regard to income distribution [5]. At worst that practice is only seen, by many critics of today's societies, as the outcome of a struggle for power between various classes or groups in society.

(*b*) In this essay I will understand by an individual's *utility* or welfare his happiness in so far as this is dependent on social elements,

*I want to express my sincere thanks to my long-time friend, Dr J. B. D. Derksen who, in personal discussions, contributed to what I have to say in this text. All errors are mine, of course.

leaving out such very personal elements of a religious, cultural or psychological character as art or love. Expressed positively, such elements, also psychological, as satisfaction from consumption (personal and collective), from work and from effort, both absolutely and in comparison to others, will be considered elements making for social happiness or welfare.

(*c*) Economists normally consider that the content to be given to such concepts as social welfare and justice is not part of their science. Often they say that statements on such matters imply *value judgements* or are the subject of ethics or morals. The way I see for tackling the problem of the definition and measurement contains not only such *ethical* aspects but also elements of a *methodological* nature. The latter may also be considered as a subject of the philosophy of science. I see my attempt as essentially scientific and in a sense, therefore, objective; it is at least a first step along this road. But, for the reasons just set out, it cannot be a chapter of economics only. Essentially it is *interdisciplinary*, in that it contains some elements of an ethical, and others of a methodological, nature. *Objectivity*, in this context, is reached if different scholars agree on the content to be given to the underlying concepts and if repeating the measurements involved produces essentially the same result. Objectivity is close to accuracy, if measurement is at stake; but no series of repeated measurements yields exactly the same results and we can only hope for agreement among a large *majority* of scholars. This majority is larger in, say, physics than in socio-metrics, because of the nature of the subjects; but it never amounts to unanimity.

(*d*) In the good econometric tradition set out in so masterly a way by R. C. Geary [1], I am presenting not only a theory in the form of a refutable hypothesis, but also a few *applications*. The latter will constitute by necessity some very simple first steps along the road indicated. From the preceding remarks, the reader will also understand that in this essay I assume the attitude, common to scientific workers, of mentioning explicitly all assumptions made, including, as does Gunnar Myrdal [3], the *explicit* formulation of the value judgements used, even though I present them as contributions to the realm of ethics. On many points the ideas presented here are close to those of Kolm [2].

2. *Elements Entering into the Utility (or Welfare) Concept*

Utility will be defined as a measurable function based on the contributions mentioned in Section 1. This function contains three types of elements to be called variables, parameters and coefficients.

(*a*) *Variables* represent phenomena that can assume different numerical values, sometimes by the individual's own choice, sometimes beyond his personal choice as when they reflect natural or social conditions. Examples are the figure or figures representing the nature of the job chosen and the consumable income going with it, as the consequence of salary and wage scales and tax rates. In a more detailed treatment of the subject, the quantities consumed of each of a number of items may also belong to this class of elements. Such details will not, however, be considered here.

(*b*) *Parameters* are figures characterising an individual's *nature;* among them appear his productive *abilities* as well as his *needs.* Examples of the former are his IQ, and other innate data such as the ability to guide other persons; examples of the latter are family size (a datum in the short run at least), and the individual's health or tastes. In the long run only *innate* characteristics should be among the parameters. Capabilities which can be attained by a learning process are among the variables; but the easiness by which this adaptation takes place again constitutes a parameter.

(*c*) *Coefficients* are numbers, dependent on the units in which utility is expressed as well as on those in which the variables or parameters are measured, indicating the effect of changes in variables or parameters on utility or welfare. Examples are the influence of a unit more of consumption on the individual's welfare, or the influence exerted on welfare by a given change in job. In some theories also the impact of an hour of additional effort will be indicated by a coefficient. Coefficients are constants; if a given variable exerts a curvilinear influence on an individual's utility, more than one constant or another mathematical form may be needed; for instance, a logarithmic form of the utility function may be sufficient.

3. *Methodological and Ethical Principles Proposed*

(*a*) In order to arrive at an objective theory of the measurement of utility the methodological and ethical principles to be introduced must appeal to a large majority of scientists and citizens. The methodological principles can best be submitted to scientists (including here not only natural, but also social and human scientists); the ethical principles must appeal to a large majority of citizens. Both requirements imply that the principles to be used must be of a rather general nature, so as to unite many; or to put it somewhat differently, they should express what unites people rather than what separates them. The implication of the approach is the assumption that there exist such almost universal principles. I purposely speak only of large majorities and not of principles accepted by all. Even in *physics*, large parts of which are considered to be objective, examples occur where unanimity does not exist. The accepted measurement of temperature by the degree of expansion of a number of substances constitutes an example. While almost all substances expand regularly upon the successive application of equal quantities of heat, in many temperature ranges exceptions to this behaviour occur. They occur wherever some substances change from the solid into the liquid phase (melting point) or from the liquid into the gas phase (point of evaporation) and they also occur on other occasions; thus, between 0^0 and 4^0 centigrade water contracts instead of expanding. In human matters we intuitively expect larger differences in behaviour and hence should not hesitate to call a method or statement objective even if the majority is less than in physics.

(*b*) The methodological principle I propose to apply is the one common to all natural sciences, including biology, namely that we select, from the theories conceived, the *simplest that does not contradict observation.* As soon as new observations are made which do not fit that theory, we choose, from the remaining theories, the simplest fitting the extended body of evidence. The process of testing a theory against observations requires that theories be formulated as "refutable hypotheses", an expression current in statistical testing [1]. Seemingly qualitative statements or propositions can often be transformed into quantitative form and so obtain the form of refutable hypotheses. Theories that do not permit such a

transformation will not, as a rule, be of much use to the aim of defining an optimal social order or social justice.

(*c*) The ethical principle I propose to introduce will be called the *principle of the fundamental equality of man.* The principle will be interpreted as meaning that we assume human beings to be equal except for those aspects where measurement has shown them to be unequal. This interpretation is eminently scientific in that it accepts the results of any generally accepted (or not disqualified) observations. It is scientific also in the sense that it introduces something which distinguishes man from other living beings. It is in the best of ethical traditions, where time and again, during the cultural history of man, some aspect of equality has been assumed to exist. In Christian theology the phrase "equality before God" is among the fundamental concepts concerning man. Among human rights, we find the legal maxim of "equality before the law", implying equal punishment for equally illegal acts; and also the rejection of many forms of discrimination. Equal voting rights are given to all citizens in a long list of countries. According to some critics in some social systems some people are more equal than others; and in some Swiss cantons women are only on their way to franchise. But, as I warned above, we cannot expect complete unanimity in human matters. The precise content to be given to the principle of equality of man unless inequality can be proven is that the *coefficients of the utility functions of all individuals are identical.* Every new observation of inequality gives rise to the introduction of a new parameter into the utility function.

(*d*) The two main principles proposed open up a road leading to the continual refinement of measured utility functions, once a start has been made. One possible start will be presented in this essay, fitting the main facts of observed income distribution in the Netherlands reasonably well. Programmes of further research, theoretical as well as statistical, will be proposed. Concrete, though very provisional, content will be given to the concepts of an optimal social order and of social justice on the basis of this start in utility measurement.

4. *Initial Hypotheses Made on the Utility Function*

(*a*) The results shown here may be seen as an elaboration of

previous work dealing with the theories of income distribution, positive and normative [4], and of the optimal social order [5] on the one hand and empirical work on income distribution and its determinants [6] on the other. The statistical material available for the Netherlands, not very different from that for some other countries, severely limits the degree of refinement of this first exercise, but still provides a first example of the feed-back from observation to theory. The only example of a parameter used in the first verification of the theory offered above is the *level of schooling* attained. While strictly speaking this level does not represent a parameter, it was taken as such as a first approximation. For a considerable number of professional, industrial and social classes figures are available on the frequency distribution of schooling level attained and of average incomes (male working population over 25 years old). The level attained was measured in units of three years' successful school attendance, ranging from 2 (primary school completed) to 6 (university master degree obtained), called v, and was supposed to constitute the parameter for each group, whereas the upper quartile s was supposed to represent simultaneously the *level required* and the *variable* characterising the *job chosen.* For some ten larger social groups comparable American figures are available and the actual American level of schooling v' (median) used as a test of the required level in the Netherlands. There was a reasonable correlation between s and v'. Since in some previous work quoted already the level of schooling was found to be able to explain the larger part of income differences, an attempt only to use this one parameter v together with the corresponding variable s was not considered hopeless beforehand.

(*b*) On the basis of the methodological principle of simplicity the first assumption made on the mathematical shape of the utility function for a person i of the active population was*

$$\omega_i = ln\{x_i - c_0 s_i - c_1 v_i - \tfrac{1}{2}c_2(s_i - v_i)^2\} \qquad (4.1)$$

where x_i represents consumable income (income after tax) in hfl 1000 per annum;

s_i required level of schooling (in units of 3 years);
v_i actual attained level of schooling, in the same units;

*Unlike the other work quoted, I neglected family size as a possible parameter in this essay.

and c_0, c_1, and c_2 represent coefficients. The expression { } may be referred to as consumable income corrected for some inconveniences, to be discussed below. The choice of the mathematical form of the (natural) logarithm expresses a well-known psychological assumption, namely that equal percentage increases in (corrected) income produce equal absolute increases in welfare. The choice also appears to simplify the problem of the optimal social order (cf. Section 7). The form of the impact of increases in s_i and ν_i was assumed to be quadratic, as a first approximation. The linear terms $c_0 s_i$ and $c_1 \nu_i$ are the simplest way of representing a monotonic impact. A positive value of c_0 indicates increasing sacrifices needed for jobs which require increased schooling; a negative sign is also possible, however, and indicates increasing pleasure in work with jobs requiring more schooling. It is uncertain whether a coefficient c_1 should be introduced at all, since nobody can experience a change in ν_i, representing here an innate property. We will return to this question later (Section 6). Considerable psychological realism underlies the introduction of the quadratic terms in the special form of the square of the difference $s_i - \nu_i$, to be indicated by the *tension* between the ability required from the individual and the ability possessed by him. The essence of the quadratic terms is that an individual feels happiest if required and possessed capabilities coincide, other things assumed equal, and that both a positive and a negative tension make the individual less happy—about equally so for equal positive and negative tensions. Since large deviations between s_i and ν_i are hardly possible, the form of the square seems to be a sufficient approximation.

(*c*) Whether all coefficients introduced can be measured depends further on the *structure of the model* used to describe the operation of the economy (or society) considered. The well-known phenomenon of identification has to be used in order to find out which coefficients of a model can and which cannot be estimated.

Moreover, any measurement, in order to be acceptable, requires a set of sufficiently high correlation coefficients obtained in the process of testing. A theory must be rejected if some or one of the correlation coefficients are below some assumed standard. The choice of such a standard is subjective. My personal preference is for multiple correlation coefficients around 0·85, implying that more than two-thirds of the observed variance in the variable to be

explained can be accounted for by the relation considered. The standard to be chosen also depends, however, on the form of the relation tested and may be set lower for relations explaining differences or small balances than for relations containing absolute values of the variables involved.

5. *The Need for Further Data*

(*a*) As already observed, the examples to be given in this essay (Sections 6 and 7) can be seen only as an initial attempt to illustrate the approach. A continuous series of improvements in both the model and the figures used is desirable and constitutes a natural follow-up which I hope will develop. Part of this series is being prepared already. For the time being our efforts are limited by the availability of data. Three stages in the use of better data can be suggested. Improvements can be undertaken by more extensive use of *existing data*. Existing data include data for other countries, for incomes and schooling as well as for a factor to be discussed later, namely the degree of independence of a job and the most suitable individuals to fill such a job. Another example of data already available relates years of practical experience, which can be deduced from age and years of schooling of the same person. Finally, it is probable that I did not use the material available on the factors included in the correct way.

(*b*) A second, more elaborate, follow-up can be obtained by the use of data which have not yet been put into an appropriate statistical form, but *can be collected* from such sources as job evaluation and career planning. Job evaluation may supply us with data on job requirements, whereas data on career planning, available in large enterprises and government services, may supply the corresponding figures on capabilities offered. Furthermore, studies examining which components of capability are innate, and which can be acquired by learning, will be helpful in refining the choice of parameters and variables in the utility function. Finally, studies on the correspondence between capabilities required, as described in job evaluation, and capabilities offered by the population, as available in career planning, will also be of great importance to our subject.

(*c*) Apart from the data which can be made available from existing sources, one can imagine the development of *future sources* as a consequence of scientific development in general. Not only can job evaluation be improved, but also the correspondence between job evaluation aspects and aspects of human capabilities. Moreover, new and better measures may be found to ascertain differences in needs or tastes. More direct and more objective ways of measurement may be developed, thus helping to improve the material described under (*a*) and (*b*).

6. *First Results of Testing*

(*a*) In the models used for the initial series of tests mentioned in Sections 3 and 4, the supply of labour of different quality was derived from utility functions as described in Section 4. For each individual the possible incomes to be obtained from alternative choices of jobs were considered to be given, i.e. free competition in the labour market was assumed. These (spendable) incomes are the difference between gross income and direct taxes, with the latter assumed to be given. Gross incomes from alternative jobs were conceived as marginal productivities derived from a *unique production function* for the economy as a whole, of a *generalised Cobb-Douglas character*. Information about the demand side of the labour market appears to be irrelevant, in this simplified model, for the estimation of the utility or welfare function. It becomes important, however, when the gross income scale, reflecting the relative scarcity of each type of labour considered, has to be explained or when the possibilities of changing it, in order to change income distribution, have to be considered. For the testing of our assumptions on the utility function the demand side is irrelevant, as stated; the essential characteristic of the model is its subdivision of the labour market into 21 compartments, later reduced to 19, each compartment being characterised by its s and ν as discussed in Section 4. The statistical material used is shown in Table 1 and the way it has been estimated described in the Appendix. Since ν is supposed to reflect the (only) parameter in the utility function and free competition is assumed to exist, a free choice of all individuals with a given ν_i between those compartments with differing values of s where that value of ν is admissible enables us to test the relation, derived from (4.1):

$$x_i - c_0 s_i - c_1 v_i - \tfrac{1}{2} c_2 (s_i - v_i)^2 = x_j - c_0 s_j - c_1 v_i - \tfrac{1}{2} c_2 (s_j - v_i)^2 \quad (6.1)$$

where some terms in v_i cancel. Because of our assumption of the fundamental equality of man, we can use as observation pairs any two pairs with equal v_i for all values of i.

The correlations found are unsatisfactory, however: cf. Table 2, Cases A–E. As a consequence, the theory in its initial form *had to be rejected.* The correlation coefficient between $x_i - x_j$ and $z_y = (s_i - v_i)^2 - (s_j - v_i)^2$ indicated by x and z in Table 2, is low (0·14) and the regression coefficients, even after the introduction of other explanatory variables, are insignificant. A new variable which seemed to have high explanatory power will now be discussed.

(*b*) Consideration of the differences in social groups suggested the relevance of an additional capability, namely the capacity to take independent decisions. The groups considered were divided into three groups, ranked according to the increasing importance of such decisions, represented by a parameter W. Persons having to take decisions of minor importance only and usually described as socially dependent were given the value $W = 0$; for persons heading small and medium-sized enterprises $W = 1$; and for heads of limited companies or persons in liberal professions $W = 2$. As can be seen from Cases A–E in Table 3, satisfactory correlations could only be obtained if W were included. This state of affairs constitutes a textbook example of the necessity of introducing an additional element into a theory in order to fit the facts in a more satisfactory way. The basic idea may also be expressed as the capacity to give leadership to other people, which is given a central place in the analysis of income distribution by Tuck [7].

(*c*) The *interpretation* of the role of W, however, poses some questions. Must we consider W as a parameter, as suggested, which implies that it indicates an innate personal capability, or must we consider it as a job variable? I opted for the first interpretation, which gives to W a role comparable to the one of v in the simplest model. This implies that we have to consider as separate groups not

TABLE 1: *Estimated Income Distribution of the Netherlands, ca 1965, According to Schooling (v), Required Schooling (s) and Degree of Independence (W) or (W')*

s ↓	v=	2	3	4	5	6	W W'	Description of Main Groups Included
6	f			0·1	0·3	0·9		Professions
	l			42·3	42·3	42·3	2	Directors of limited companies
	l'			19·0	19·0	19·0	6	
	x			14·0	14·0	14·0		
5	f		1·5	1·1	0·5	0·7		Independent industrialists, teaching staff of secondary and third level
	l		19·5	19·5	19·5	19·5	1	
	l'		14·2	14·2	14·2	14·2	5	
	x		11·3	11·3	11·3	11·3		
4	f	0·7	5·3	4·4	1·3	0·1		Civil Servants in general service; private employees
	l	10·2	10·2	10·2	10·2	10·2	0	
	l'	9·7	9·7	9·7	9·7	9·7	2	
	x	8·3	8·3	8·3	8·3	8·3		
3	f	0·1	4·0	1·0	0·1			Independents in trade and services
	l	15·0	16·2	17·4	18·6		1	
	l'	12·0	12·5	13·0	13·4		3	
	x	9·9	10·2	10·5	10·8			
3	f	10·7	6·6	0·1				Primary school teachers; Part of administrative personnel; Police; Mining workers; Farmers
	l	8·9	10·8	12·7			0	
	l'	8·5	10·2	11·3			2	
	x	7·4	9·0	9·4				
2	f	46·6	13·8					Workers, including retail trade employees, part of administrative personnel
	l	5·3	8·4				0	
	l'	4·9	8·1				2	
	x	4·4	7·1					

f: frequency in per cent of number of tax payers
l: income in thousands of guilders
l': labour income contained in *l*
x: labour income after tax

TABLE 2: *Results of Testing the Two Theories by Equating Utility between Cells with Same v (Theory I, Cases A through E) or between Cells with Same v and W (Cases F, G, H)*

Case No.	N	Regression Coefficients for x on:							Simple Corr. between x and				Const. Term e	Mult. Corr. Coeff.
		ζ	w	q	q'	q''	m	ν	ζ	w	q	ν		
A	16	0·09 (0·15)	1·64 (0·66)	−0·18 (0·22)	.	.	.	.	0·14	0·89	−0·83	.	+0·55 (0·29)	0·90
B	16	0·10 (0·16)	.	−0·05 (0·39)	.	.	0·47 (0·28)	.	0·14	0·89	−0·83	.	+0·46 (0·33)	0·87
C	16	0·09 (0·15)	1·60 (1·01)	−0·17 (0·37)	.	.	0·02 (0·39)	.	0·14	0·89	−0·83	.	+0·55 (0·33)	0·90
D	12	0·12 (0·17)	1·73 (0·56)	.	−0·06 (0·16)	0·07 (0·07)	.	.	0·14	0·89	−0·83	.	+0·42 (0·35)	0·90
E	16	0·08 (0·14)	2·11 (0·30)	.	.	.	.	.	0·14	0·89	−0·83	.	+0·54 (0·28)	0·89
F	6	0·44 (0·12)	.	.	.	.	.	1·10 (0·40)	0·63	.	.	−0·20	−3·6 (1·42)	0·91
G	6	0·28 (0·28)	.	.	.	.	.	−0·17 (0·44)	0·79	.	.	−0·70	+1·47 (3·34)	0·79
H	6	0·32 (0·06)	.	.	.	.	.	.	0·79	.	.	.	+0·9 (0·5)	0·79

N: number of observations (differences between two cells with same utility).
x: differences in labour income after tax between two cells or $x_1 - x_2$.
ζ: $s_1^2 - s_2^2 - 2\nu(s_1 - s_2)$; s schooling required, ν schooling available.
q: $W_1^2 - W_2^2 - \nu(W_1 - W_2)$; W degree of independence (as in Table 1).
q': $W'^2_1 - W'^2_2 - \nu(W'_1 - W'_2)$; W' degree of independence (same, but 3 instead of 2).
q'': $W'^2_1 - W'^2_2$; $w = W_1 - W_2$; $m = w\nu$.
In Case F lowest row of observations in Table 1 has been excluded.
In Cases G and H lowest row of observations in Table 1 have been combined.