Choice behaviour and verbal behaviour: a critical assessment of their relevance for practical policy

presented at the Tinbergen Conference,

Empirical models and policy making

May 1997

B.M.S. van Praag[1]

P. Frijters[2]

---

[1]Director of the Foundation for Economic Research,
professor at the University of Amsterdam.

[2] University of Amsterdam, Tinbergen Institute research associate.

CONTENTS

1. Introduction

Ten years after the Tinbergen Institute was founded as a joint initiative of the Erasmus University Rotterdam, The University of Amsterdam and the Free University of Amsterdam, we are together to celebrate this occasion. First, we would like to congratulate the universities and the board of the Tinbergen Institute for their tenacious activities during those ten years and for the remarkable success of this enterprise. Clearly, there is a lot left to be desired but, given the difficulties, we believe that the Tinbergen Institute represents a significant contribution to the promotion of economic science in the Netherlands.

There is no doubt that Tinbergen himself should be considered as the father of the modelling approach. With his background in physics and his political penchant towards democratic socialism and his belief in the feasible society, he endeavoured to describe the economy by a set of a few equations, to specify these equations functionally and to estimate the parameter values in the equations. The next step was to identify problems and objectives, to design policy instruments and to calculate policies in terms of instrument variables. Finally he advised the authorities in the Netherlands and abroad on economic policy and how to implement those policies. Although Tinbergen's approach may be considered in hindsight as too optimistic and too simplifying, there is no doubt that his approach signified a watershed between the old-fashioned approach based on intuition and not much else and the modern approach of economic policy built on models and empirical data. He founded the Dutch Central planning Bureau in 1945, which nowadays has a counterpart in nearly every developed country.

Nevertheless, there have been blind spots in his approach. Tinbergen himself recognised this and devoted some chapters in *Economic Policy: Principle and Design* to so called "qualitative policy".

Despite this attempt, the restriction in Tinbergen's approach has remained in the following decades. It implies that all variables that are considered unmeasurable have been left out of the model and as a consequence that all models used in practice are partial in a sense. Hence they do not incorporate *behavioural reactions* of private individuals to e.g. new airports, such as when individuals move to another place. Second, they lack a social welfare function or, less ambitiously, an idea about how the citizens evaluate their individual state of affairs and the state of the nation as a whole. Objectives, if any, are cast in terms of aggregate employment, growth and price stability.

Fundamental choices by individuals and the corresponding effects on individual welfare and/or well-being, linked to changes in infrastructure, environment, the taxation system and social security, usually remain outside the scope of our observation.

There are two aspects of importance here:

a. The **Positive question**. How will individuals change their behaviour as a reaction to changes in policy?
b. The **Normative question**. How would individuals evaluate a change in policy?

For instance, let us assume that a government considers a change where automobiles are not taxed with respect to the value of the automobiles but with respect to the number of kilometres driven. As a behavioural change we may expect frequent

drivers to reduce the number of kilometres they drive and we may expect the number of car-owners to increase as the threshold to buy a car decreases.

Secondly, some individuals will feel that their well-being has improved while others will feel their position has worsened. Obviously, the change in policy is not a Pareto-improvement, but does this imply that we may ignore the welfare distribution shift when asked to advise for or against the proposed policy?

Another example is the cleaning of a lake or polluted industrial plots. A third example is the question how Amsterdam Airport should be extended. A fourth example is the public discussion on the railroad use from Rotterdam harbour to the German hinterland. A fifth example is the cost and benefit of health policy measures. All these problems, to which we can add many others, call for a quantitative empirical assessment of the preferences of citizens. In fact, nearly all policy proposals are good for some and bad for others, making virtually all policies dependent on an implicit weighting of different interests. Although this is not well-integrated in the Tinbergen models, there are a couple of isolated approaches to solve such problems. Isolated from what? These approaches are mostly initiated by psychologists, marketeers, environmental or regional economists. They are not embodied in mainstream economics.

Let an individual i have income y_i and let his well-being be influenced by his level of income y_i and other variables z_i and the price vector $\mathbf{p}$ of goods. We may think of the vector $\mathbf{z}$ as including the level infrastructure, level of environment, the family situation, and the health of the individual. Well-being may now be described by an indirect utility function

$$U=U_i(y_i,\mathbf{p}_i,\mathbf{z}_i)$$

The question is now how much individual i increases or decreases in welfare if $\mathbf{p}_i$ or $\mathbf{z}_i$ changes. Actually we are interested in the money amount needed to compensate a change in $\mathbf{p}_i$ or $\mathbf{z}_i$. Hence we look for the empirical solution of the equation

$$U_i(y_i+\Delta y,\mathbf{p}_i+\Delta p,\mathbf{z}_i+\Delta z)=U_i(y_i,\mathbf{p}_i,\mathbf{z}_i)$$

in terms of Δy. In the context of the paper we assume p constant and focus on changes in z.

In this paper we shall briefly outline and evaluate five approaches which have been applied in recent years. The first approach is based on the observation of behaviour and stays within mainstream economics. We call it the equilibrium approach, exemplified by Blomquist et al. (1988), as it is based on the assumption of a behavioural equilibrium. The other four approaches are based on the observation of response behaviour to more or less complicated question modules.

The method of contingent valuation approach (Hanemann (1994)) is rather straightforward in asking how much one is *willing to pay* for more or less of something. Conjoint measurement tries to estimate a preference structure on a multi-dimensional set of alternatives by observing choices for hypothetical products. The Leyden method uses evaluations of incomes and estimates indirectly the effect of other variables on *welfare*. The Cantril question estimates *well-being* and again we may indirectly assess the effect of other variables.

2. The equilibrium approach (revealed preference)

Mainstream economists are wary of trying to estimate utility functions. Actually, they argue that it is impossible to estimate utility levels. However, it may be possible to estimate indifference curves from observed behaviour.

This is the method which is used in an article by Blomquist et al. (1988). They consider differences in *quality of life* (QOL) between 253 American counties. Let county k have a better QOL than county j, than this implies that an individual needs a monetary compensation to move from k to j. The monetary compensation is defined as the QOL-difference between j and k. The county k is characterised by a 16-dimensional amenity vector a_k whose amenities include climate variables like precipitation, humidity, etc., infra-structural variables like "violent crime" and the teacher-pupil ratio and environmental variables like "visibility", "disposal sites", etc. Let (indirect) household utility be described by

$$v=v(w_k, r_k, a_k)$$

where w_k stands for the wage rate, r_k for the rental price of land and a_k for the amenities in county k, then the implicit price of a_k is found by total differentiation. We have:

$$v_w \, dw/d\, a_k - v_w \, qdr/d\, a_k + v_a = 0$$

and hence

$$v_{ak} / v_w = qdr/da_k - dw/da_k$$

where all the derivatives of v are assumed to be positive and q stands for the Hicksian demand for land.

Hence if people live in different counties which differ with respect to one amenity, say a, by an amount Δa, then we need to increase money income by

$$(qdr/da - dw/da)\Delta a$$

to ensure equal utility in both places. Obviously if amenities differ for all 16 components the money difference will be approximately

$$QOL = \sum^{16}_{i=1} (qdr/da_i - dw/da_i) \, \Delta \, a_i$$

If we assume that all individuals in the different counties enjoy equal utility, because the individual would move otherwise until a spatial equilibrium is reached, then our empirical situation is *utility equality*. This is Blomquist's equilibrium assumption. If we then estimate the function r(a) and w(a), i.e., housing rent[3] and wage rate level as a function of the amenity vector a, then we can numerically assess QOL, as defined above. Blomquist et al. assess the amenity difference vector **a** for 253 counties and they estimate the functions r(a) and w(a) for these 253 observations.

The study as such is highly ingenious and the results are very interesting. The approach is also methodologically attractive as it totally circumvents the problem of how to evaluate welfare explicitly. Nevertheless, we have some major points of critique. The first point is the basic assumption that there is a spatial equilibrium, such that every individual is at the same indirect utility level. Secondary assumptions are that all prices other than housing are the same in each region and transport costs are negligible. Clearly these assumptions are blatantly unrealistic. The two latter assumptions are testable and the method may be refined in order to cover the differences in price levels and the existence of transportation costs. However, the first assumption of utility equalisation over individuals cannot be tested if one does not accept a welfare measurement method to investigate the hypothesis. Another big problem is that the utility function and the conditions will vary with age, birth place and family circumstances. The availability of good schools and kindergartens will for instance not be relevant to families without children. It seems also very likely that many individuals will be reluctant to leave their place of birth, even if it is objectively speaking unattractive, just because it is their native region.

We may therefore qualify the equilibrium approach as courageous and ingenious, but its strict assumptions on possible tastes and choices mean that it fails to yield

[3]Blomquist et al. approximate the price of land by standardised housing rents.

politically relevant results. It is the only method based on actual observed choice behaviour we will investigate in this paper. There are of course many more examples of similar hedonic price analyses (e.g. Roback (1982, 1988), Vijayendra (1993)) which go back at least to Rosen's (1974) model of implicit markets. The analyses all have in common that they assume perfect knowledge and equality of utility functions across individuals.

3. The Contingent evaluation approach

Contingent valuation is now a very popular method in the USA to assess environmental values. Literature surveys are found in Hausman (1993) and Johansson (1993). A critical discussion of the methodology is found in Hanemann (1994) and Diamond and Hausman (1994).

The type of problem for which the method is used is sketched by an example. Let us assume there is a polluted lake in your neighbourhood. In order to assess the willingness of citizens to clean up the lake you may ask them:

"How much would you be willing to pay for cleaning up the lake?"

The resulting amount is called the *willingness to pay (WTP)* of the individual. Similarly one can imagine the situation of a lake being threatened by pollution. Then the symmetric question would be:

"How much would the authorities have to pay you in order that you would accept the pollution of the lake?". The resulting amount of money will be called the *willingness to accept* (WTA) of the individual.

The questions are obviously attempts to solve our initial problem.

Let the individual's welfare position be described by two dimensions, viz., his income y and the situation of the lake measured by some pollution index z where the lake becomes less polluted if z increases. The welfare of the individual in the reference situation is then U(y,z). In the case of the WTP measurement, the WTP-question boils down to finding the number WTP to solve

$$U(y,z) = U(y-WTP,z+\Delta z)$$

where $\Delta z>0$

Symmetrically, WTA may be seen as the solution to the equation

$$U(y,z) = U(y+WTA,z-\Delta z)$$

Notice that the question does *not* presuppose equal preferences over individuals, nor any knowledge of a cardinal utility function, or indeed even of the existence of a cardinal utility function.

The aggregate WTP over the population may be interpreted as *the* value the community would want to raise to clean up the lake pollution. For instance, if we assess the damage of the Exon Valdez oil spill, aggregate WTA (or WTP) would be an assessment of the damage.

The methodology has an intuitive appeal. If there is no market mechanism available for individuals to reveal their preferences, it might be the best approximation to the market mechanism.

On the one hand the method of contingent valuation (CV), as it is formally called, is frequently depicted as the only sensible way to assess the money value of public goods, especially in environmental affairs and has been recommended as the legal base for assessing environmental damages. On the other hand there has been much critique on the methodology, notably by economists (see Hausman (1993)).

The critique is the following:

A. *People are unable to answer such questions in a reliable way*

Many experiments have been done which show that individuals are sensitive to the precise wording of the question, have difficulty in understanding the question asked, that they answer differently to repeated questions, that WTP and WTA values are not equal, that respondents answer inconsistently in the sense that if you ask for five lakes first and *then* for one lake only, you get a one lake-value which is below the value one gets when *only* answering for the one lake, etc. Such embedding effects are decribed in Kahneman and Knetsch (1992).

We would like to raise some additional questions:

B. *Strategic answering will be present at a large scale*

It is obvious that if the willingness to pay is not followed by the actual requirement to pay the promised amount that people will be inclined, especially when they know the objective of the survey, to exaggerate their position. This will be especially the case if there is a public opinion campaign or/and when social desirability issues are involved. To be fair to the method, the design of the questionnaires has been altered to take account of the strategic problem. The contingent valuation questions are now often worded in terms of whether individuals would accept a proposed compensation for pollution or a proposed bill for cleaning up pollution. By restricting individuals to a yes\no answer in a referendum type-situation there is less incentive for an individual to hide his true preferences and no possibility to exaggerate his willingness to pay or his unwillingness to accept to absurd levels.

C. *Dependence on the existing income distribution*

In public goods decisions it is a good democratic principle to equate the influence of citizens, independent of their wealth, income and status.

In this method it is obvious that answers will depend on the income of the respondent. Hence rich people will have a higher WTP than poor people. Hence an equal WTP amount does not reflect equal marginal welfare additions. It follows that if the mean WTP would be the basis for assessing public values and for deciding on public policy via cost-benefit analysis where the benefits are defined via WTP measurement, then the rich will have a disproportionate influence. Hence it may be a plutocratic mechanism without correction. However, it is hard to see, and in our opinion impossible, to correct for those effects without knowledge of cardinal welfare differences. Even if one takes the median reported value as the WTP to be used in political analysis, thereby reducing the influence of the rich, there remains the problem that the actual welfare losses of those with a lower than median WTP are not necessarily smaller or greater than the welfare gains of those with a higher than median WTP.

D. *Neglect of the finiteness of the public purse.*

When we evaluate public goods (either potential losses or gains) we have to be aware that the effort to clean a lake or a polluted plot, etc., will mean that less money will be available elsewhere, e.g. for education or health care. In the case of moving Amsterdam Airport (very improbable, but just for the sake of argument) environmental advantages are counterbalanced by a loss of wealth and jobs for the region. Hence the correct but much more difficult question would be

> "How much would you like to pay for a removal of Amsterdam airport, being aware of the risk that you or some of your fellow compatriots might lose your job?"

It is illegitimate to assess such large scale works without reckoning with and making the respondent aware of the important side-effects as well. Though contingent

valuation studies may want to give the respondent as much information about the wider context as possible, there is an obvious limit to how much information one can give in an interview.

E. *Welfare assessments depend on the situation.*

It is well-known that welfare functions depend on the situation. Van Praag (1971), and Van Praag and Kapteyn (1973) found that the evaluation of income levels y by individual i, say $U_i(y)$ depends on the current income of i (see section 5). People partially adapt to their income and use it as their reference position to evaluate other incomes. Van Praag called it preference drift.

The same phenomenon is also known as the "hedonic treadmill" (Brinkman and Cambell (1971)). That a similar adaptation mechanism is probable for environmental changes is more than likely. If we call A and P the *ex ante* and *ex post* evaluation of (y,z) then we have two possible equations describing the willingness to pay for cleaning a polluted lake before it is cleaned, say WTP_A, and the willingness to pay to clean up a lake after it is cleaned (WTP_p):

$$U_A(y_A,z_A) = U_A(y_A\text{-}WTP_A,\ z_A + \Delta z)$$

and

$$U_P(y_A,z_A) = U_P(y_A\text{-}WTP_p,\ z_A + \Delta z)$$

It follows that it is well possible to find two different values of WTP, where the last value, say WTP_P, will probably be lower than WTP_A due to the adaptation to the new situation. The utility U_A function and U_P will differ. Hence, people assess a lower value to cleaning up the lake after it has been cleaned than they are willing to pay for cleaning it when it is polluted.

It is open to discussion which WTP we have to use.

Our conclusion is that the CV-method, although intuitively attractive, cannot be used as a serious basis for public decisions on the provision of public goods, amenities or environmental changes.

4. The Conjoint measurement approach.

A more sophisticated way of using respondent's knowledge of their own preferences than contingent valuation is the method of conjoint measurement. It originated in marketing research as a way to get insight in the market potential of new goods (see e.g. Luce, R.D., Tukey, J.W. (1964)). Consider prototypes of new varieties of a product, say peanut butter. The prototypes may vary in taste, fat contents, colour, shape of the package and price. In short, the various prototypes may be described in terms of attributes or characteristics $Z_1,..,Z_k$ (c.f. Lancaster (1966), Van Praag (1968)) which assume attribute levels $z_1,..,z_k$. A prototype may be described by a k-vector **Z**. When offering a choice of m prototypes $\mathbf{Z_1,..,Z_m}$ to N prospective consumers, we may observe how the various prototypes are ordered according to the preferences of the respondent i. Consider two prototypes say $\mathbf{Z_1}$ and $\mathbf{Z_2}$

We find that individual i prefers $\mathbf{Z_1}$ to $\mathbf{Z_2}$ while j has inverse preferences. The idea is then that we attain a utility value U(Z) to a prototype Z and that the observed orderings roughly take place according to the ordering of utility values. It is obvious that if we have a lot of respondents and a sequence of prototypes with enough variation in attribute values, then it becomes possible to estimate U(.), i.e., its functional specification and numerically specified parameters.

The experiment may also be enriched by adding to rank ordering (RO) the question which prototypes would be acceptable buys (AB).

If respondents are expected to have different utility functions depending on their age, gender, income, education, etc., in short a vector of individual attributes Y, then the utility function may be *individualised* as U(Z;Y) and again U(Z;Y) may be estimated from observing RO- and AB-behaviour.

The way in which the utility function should be estimated is a matter of debate. A simple way is to attach rank ordering numbers and to interpret them as utility values. These utility values are then explained by linear or non-linear regression whereby the attributes and individual variables Y are the explanatory variables. A more sophisticated approach is by rank-ordered logit models (see Beggs, Cardell, Hausman (1981)).

In this approach we are not observing real choice situations as the prospective consumers are observed in an artificial situation and are aware of it. Nevertheless, it is a good way to get insight into preferences for as yet non-existing goods. For many commodities we do not have actual prototypes, and we have to rely on symbolic presentations of those prototypes. This can be done with so-called *vignets*, verbal descriptions of the choices, perhaps supported by photographic or computer presentations. The vignet method paves the way for getting insight into preferences when the alternatives are intangible, e.g. services, or where it is too costly to produce sufficient variety (e.g. car prototypes). Here we may also think of variables like infrastructure, climate, etcetera.

The problem with extending the descriptive technique is that the choices become more fictitious and that consequently, the risk is there that the observed choice processes are unrelated to reality. The weak spot is the formulation of vignets. They have to look realistic and they have to be unambiguous in interpretation. It follows, and that is certainly a restriction, that respondents have to be knowledgeable on the subject or that they have to be quickly introduced to the subject. For example, it does

not make sense to enquire on the preferences for classical music amongst a sample who is not familiar with that music.

There is no reason why the method is only applicable to preferences relating to commodities or services. E.g., Van Beek, Koopmans and Van Praag (1993, 1997) presented vignets of low-skilled workers to recruitment officers in firms, where the attributes were, amongst others, age, education, health, and ethnic origin. As an example of a vignet from Van Beek et al. (1997):

Table 1. **Example of a vignet; a profile of a prospective worker**

gender	Male
age	23
health	Healthy
origin	Netherlands
education	Junior technical
command of language	limited
travelling time	one and a half hours
current status	unemployed for one year
availability	part-time, work in shift or irregular hours no problem
family situation	married/living together, two young children, breadwinner
working experience	none
wage costs	Individual wage cost subsidy of FL 600,- a month

In a similar way we might subject preferences on environmental, infrastructural or societal issues to the conjoint measurement approach. We might think of variation in tax structure, public tariffs or inequality, or the appropriateness of specific penalties or fines for specific crimes.

Let us now look at the advantages and disadvantages of this approach. The advantages of the conjoint measurement approach are:

1. It gives the possibility to include non-existing but conceivable options. Hence it is an ideal method to study changes in reality which are conceivable but not yet realised.

2. It is an ideal method for hedonistic analyses where the effect of multi-dimensional and non-separable attributes may be studied.

3. It is a rather cheap method since we may supply many choice situations, say p, to one respondent, thereby obtaining a large number of total observed choices from a relatively small sample of respondents. In that case we are creating p observations per respondent. Evidently these p observations are not mutually independent. First, the respondent-specific part Y remains the same over the p observations; second it may be that random disturbances in the choice patterns may be correlated over the p observations. Hence we cannot assume that the p observations per respondent are mutually independent. When we have n respondents, it follows that the trick of asking p choice situations does *not* increase information p times. However, that information is increased considerably at low marginal costs may be taken for granted.

4. A fourth aspect which is statistically rather useful is that we avoid selection biases, which are unavoidable when we consider only the best and realised choice for each respondent. Again the example in Van Beek et al. is very instructive. There we do not only get an idea of what is considered the best worker who is employed, but we also find out why the other workers are not employed.

The disadvantages of the method are also obvious. The design of the questionnaire, including the description of the vignets is much more difficult and subjective then asking only for objectively controllable facts. Although it is unnecessary that vignets are "orthogonal", as frequently stated, it certainly helps when vignets are not too

similar and that attributes are not nearly linearly dependent across vignets. We estimate that designing a conjoint measurement survey is more tricky and more expensive than a traditional questionnaire.

How far are the results from conjoint measurement reliable?

There are two ways in which reliability may be assessed. First we may look at standard deviations, t-test statistics, F tests, Mac Faddens R^2, etc. They have the same interpretation as in other cases. The second point is how far the estimated model gives good predictions. We do not know of studies where predictions have been compared to realisations but we do not see beforehand why the predictive value would be less than for other extrapolations based on model estimation data. Actually the mere fact that thousands of these studies have been ordered and performed on behalf of profit maximizing firms (Catting and Willing (1989)) in a sense is the best proof of its predictive value. Nevertheless, we recommend some wise reluctance. It would be good practice to test and report the sensitivity of the estimates to changes in the questionnaire.

Conjoint analysis is an approach in which the use of many alternative choices helps in avoiding socially desirable answering effects and can be used to determine the relative importance of many different attributes. If prices and/or income have been included in the attributes, it is also possible to assess WTP. In that sense it may be seen as an improvement on CV (see also Kamp and Maxwell (1994).

In conclusion we see conjoint measurement, if wisely and prudently implemented, as a valuable extension to traditional data collection. For a survey, see Willing and Cattin (1989), and for a discussion of recent refinements, Wedel et al. (1995).

5. The Leyden approach

A totally different version of verbal behaviour is used in the so-called “Leyden approach”[4]. The Leyden approach focuses on estimating the so-called welfare function of income (WFI). The empirical basis is the Income Evaluation Question (IEQ), which runs as follows:

> "Whilst keeping prices constant, what after-tax total monthly income would you consider for your family as:
>
> very bad,.....
>
> bad, ...
>
> insufficient,.....
>
> sufficient,.....
>
> good, ...
>
> very good, ... "

Let the six answers of individual i be denoted by c_{ij}. The verbal label bad, .., good describe satisfaction levels and the respondent i is asked which household income level c_{ij} he associates with each verbal label j.

If verbal labels have the same emotional significance for respondents, the answers become comparable across individuals. Obviously the answer c_{ij} will depend on personal circumstances, say Y_i, which includes income. Let two respondents s and

[4] The Leyden approach was developed by a group of researchers at Leyden University (see Van Praag (1971)).

t identify c_{s4} and c_{t4} as the household income they would call “sufficient”, and let us assume that a sufficient income is translated into a utility value $U_{j=4}$ then there holds

$$U_{j=4}(c_{s4},Y_s) = U_{j=4}(c_{t4},Y_t)$$

knowing c_{ij} and Y_i for many respondents we are actually observing the functional relationship

$$c_{ij} = c_{ij}(Y_i)$$

The best approximation found thus far includes at least family size (fs) and household income y (see Van Praag and Van der Sar (1988)). More precisely, the equation

$$\ln(c_{ij}) = \beta_{0j} + \beta_{1j} \ln(fs_i) + \beta_{2j} \ln(y_i)$$

has been found, where typical values are β_1=0.1 and β_2=0.6.

The fact that income is one of the determinants is called *preference drift*. This is an example of how norms and values shift with the individual’s reference situation (Kapteyn (1977)). We find that somebody will be at satisfaction level j if

$$\ln(c_{ij}) = \ln(y_i)$$

which is the case if

$$\ln(y_i) = \ln(c_{ij}) = (\beta_{0j} + \beta_{1j} \ln(fs_i))/(1- \beta_{2j})$$

The interesting point is here that $\beta_{1j}/(1- \beta_{2j})$ represents the shadow price elasticity of extra household members. Similarly, by including any other variable Z we may wish to investigate into Y_i, we may find the shadow price of Z. In Van Praag (1988) and Frijters and Van Praag (1996) we used this method to obtain shadow prices for climate conditions in Europe and Russia, thus yielding climate equivalence scales. The method may even yield more information than shadow prices and equivalence scales, it is also possible to use it for cardinal welfare computations. It was found in Van Praag (1991) that satisfaction levels are numerically translated on a finite scale by an equal partition of that scale: if there are six verbal labels for a respondent to

answer to and the scale goes from zero to one, then $u_1 = (2-1)/12$, and in general if there are k verbal labels there approximately holds

$$U_j = (2j-1)/2k$$

See also Van Praag (1971) and from a psychological angle Parducci (1995).

This entails the cardinalisation of the answers to the IEQ on a zero-one scale. If we look at the answers to the IEQ and set them against the values of U_j , the best empirical fit found is

$$U(c;y)=\Lambda(c;\mu,\sigma)$$

where Λ is the lognormal distribution function and μ and σ are, respectively, the mean and the variance of the logarithm of the c_{ij}'s.

We may interpret this function as a cardinal utility function, although any monotonic transform will provide the same information.

An example of the Leyden method may be found at the end of the next section, where we show a table on the estimated effect of climate variables on welfare and well-being in Russia.

6. The Cantril-method

This method uses self-reported satisfaction as an estimate of the well-being level of individuals. By answering the question

> "How satisfied are you with your life as a whole at the present time on a scale from 0 to 10?"

one gets a number denoting the level of satisfaction of that individual[5]. As with the revealed preference method and the Leyden method, the monetary value of amenities may now directly be found by comparing individuals with different incomes and different amenities who are at the same well-being level. Well-being here does not only consist of material welfare, as captured by the Leyden method, but also includes non-monetary circumstances which may influence the degree to which one is satisfied with one's life.

This question has been posed in many varieties, also with symbolised answers, like weeping and laughing faces, as in the Michigan tradition. A legitimate question is whether all respondents use the same frame of reference. A second question is whether the same answered *value* implies the same feeling for each respondent.

The frame of reference is certainly important: are we dealing with satisfaction at this moment only, satisfaction with work, income, health, or an amalgam of these life aspects. People have been found to adapt to their own situation. Individuals frequently consider only those variables which are less than perfect, causing dissatisfaction. Nevertheless, the volatility with respect to response behaviour does not seem to be very high. The rather great influence of present mood and recent events on satisfaction answers, be they general or on specific life areas, merely requires a large amount of observations to find the structural parameters. The high volatility *per se* does not bias the underlying relationships. With respect to the valuation issue, the comparability between individuals, we refer to Van Praag (1991). There it was found that the translation of verbal labels into points on a scale was rather stable. The opposite also holds. If we therefore assume that verbal lables like

[5]There are many variations to this question. In fact the original Cantril question asked individuals to position themselves on a vertical ladder.

"good", "bad", etc. may be identified with specific levels of satisfaction the translation issue looks rather harmless.

If we do not accept the validity of the satisfaction question we arrive at a scientific *solipsism*. It is then not possible any more to attach meaning to the concept of satisfaction or to the values given.

Theoretically the use of the Cantril well-being function, say

$$W_i = W_i(\mathbf{z_i}; y_i)$$

might be a solution to the problem of formulating a social welfare function of the type

$$\hat{W} = \sum W_i$$

which might be used for evaluating policy changes and its impact on the well-being of specific social subgroups, like workers, entrepreneurs, old-aged, and the unemployed.

The advantages of the Cantril question are:

1. An all-encompassing measure of utility is used which is based on the evaluation of the individual himself, instead of an evaluation by an outside observer.
2. The effect of all policy aspects on well-being may be examined in theory (e.g. Frijters and Van Praag (1996), Plug, Van Praag and Hartog (1996), see also the job-satisfaction literature: e.g. Clark and Oswald (1994)).
3. Informational or behavioural restrictions do not influence the outcomes of the method.

An example of both IEQ and Cantril results is given in the next table. We find a strong preference drift and effects from family size and age in the IEQ. We notice that the income effect in the Cantril function (=0.25) cannot be directly compared with the preference drift effect, but rather with $1-\beta_2 \approx 0.38$. It appears not

unexpectedly that income has a weaker effect on well-being than on welfare. In this analysis by Frijters and Van Praag (1996) the objective was to analyse the effect of climate variables and to calculate welfare or well-being equating compensations to neutralise climate differences. Here we have tried to measure the effect of climate variables in Russia (data is from 1993 and 1994) both on levels of welfare and on levels of well-being. A reduction in μ implies a reduction in financial need and thus a higher evaluation of one's own income. As we may see, welfare is mainly influenced negatively by harsh winters with strong January winds, whereas well-being is influenced by such factors as the amount of sunshine. Included as climate variables are the average temperature in January and July (JANTEMP and JULTEMP), the average annual temperature (TEMPAV), the difference between maximum and minimum temperature in one calendar year (TEMPDIF), the average level of annual precipitation (PREC), the average amount of precipitation in the summer and winter (SUMMERPREC and WINTERPREC), the number of raindays a year (RAINDAYS), the number of sunhours a year (SUNHOURS), average wind speed a year (WIND), average wind speed in January (JANWIND) and the height of the region above sea level (HEIGHT).

We refer to Frijters and Van Praag (1996) to for a thorough analysis. Here it is only shown as an example of what can be done with the IEQ and the Cantril Question.

Table 3: results of climatic model-estimation*: the effect of climate on the amount of income needed to reach a fixed standard of welfare (μ) and the effect of climate on overall well-being (Cantril).

	Least squares:				Ordered probit:	
	μ		Cantril		Cantril†	
Intercept	-4.35	(2.0)	17.72	(5.3)	-	
ln(y)	0.62	(56.6)	0.25	(16.9)	0.14	(6.6)
ln(fs)	0.17	(8.7)	-		-	
ln(age)	2.14	(5.1)	-3.92	(6.6)	-2.69	(3.2)
ln^2(age)	-0.30	(5.4)	0.53	(6.6)	0.36	(3.2)
Dummies:						
rural	-0.08	(4.3)	2.17	(4.7)	1.82	(2.8)
Volga and South Russia	0.13	(4.3)	0.08	(1.8)	0.08	(1.3)
wave2	0.36	(17.8)	-0.22	(7.9)	-0.17	(4.3)
Climate variables:						
ln(JANTEMP)	-1.30	(3.3)	-		-	
ln(JULTEMP)	3.84	(5.4)	-		-	
ln(TEMPDIF)	-2.31	(6.9)	-		-	
ln(JANWIND)	4.07	(4.9)	-5.67	(6.5)	-6.75	(5.6)
ln(HEIGHT)	0.11	(5.5)	-0.11	(3.4)	-0.08	(1.9)
ln(RAINDAYS)	-		0.86	(6.1)	0.36	(1.9)
ln(SUNHOURS)	-		0.84	(4.5)	0.37	(1.4)
ln(WINTERPREC)	-		-0.50	(4.3)	-0.16	(1.0)
ln(SUMMERPREC)	-		-0.35	(4.6)	-0.22	(2.1)
Interaction Terms:						
ln(JANTEMP)*ln(JANWIND)	-1.12	(4.8)	1.59	(6.7)	1.85	(5.7)
ln(PREC)*ln(TEMPAV)	0.06	(5.5)	0.21	(4.3)	0.09	(1.3)
ln(HUMIDITY)*ln(TEMPAV)	-		-1.45	(7.0)	-1.29	(4.4)
Rural*ln(PREC)	-		-0.34	(4.6)	-0.28	(2.7)
R^2	0.731		0.094			
McFaddens pseudo-R^2					0.007	
% of observations correctly predicted					55.8%	
N	4412		4412		4412	

*absolute t-values in parentheses. y denotes current household income not corrected for climate

Source: Frijters and Van Praag (1996).

7. Discussion and conclusion

In this paper we argued that public decisions frequently deal with problems for which the Tinbergen macro-economic model framework is insufficient. This is especially imminent in problems relating to the environment, infrastructure and the choice between options from different welfare perspectives, like *welfare*, *well-being*, differences in social subgroups, etc. The first problem is that we are unable to model behavioural effects for changes which have not been realised yet. Secondly, we have no idea as to the evaluation of the changes by individuals involved and subsequently the effect on social welfare, however defined.

In this paper we shed light on some relatively new techniques developed outside macro-economics, even partly in other sciences which may be helpful to incorporate among the main instruments in the art of modelling. All these methods aim or can be utilised to evaluate changes in non-market variables in terms of monetary countervalues. We summarize the main methods:

	EA	CV	CM	Leyden	Cantril
flexibility	-	+	+	-	+
Assumption credibility	-	±	±	+	+
multi-dimensional	+	-	+	+	+
experience with method	+	+	+	+	moderate
Sample size needed	n.a.	large	small	large	large
simplicity questionaire	+	+	-	+	+
predictive quality	±	-	±	+	+
relatively cheap		-	±	±	+ +
availability of data	??	??	-	mod	yes
turnover time	-	+	+	+	+

* + stands for favourable, - for negative and ± for indeterminate

We have tried to evaluate the five methods on various aspects.

The first aspect is *flexibility*. Is it possible to mould the method ad hoc in order to analyse different phenomenon? In this respect EA, Leyden and Cantril score low.

The second aspect deals with the *assumptions needed*. Are the basic assumptions underlying the method credible, where we do not mean the statistical or econometric assumptions on error terms or functional specifications. In that repect we have serious problems with EA. To such an extent, that we doubt the validity of the method.

The third aspect is the number of different *dimensions* which may be considered simultaneously. There all methods score high with the exception of the CV method. We notice that Leyden focuses on the value of income, say *welfare*, and Cantril on a wider concept, viz. *well-being*. The second concept is more comprehensive. Hence the Cantril question may be preferred in some contexts. It covers income as one of the explanatory variables and consequently we can assess trade-offs between non-monetary variables and income.

The fourth one is the availibility of *knowledge and experience* with the method. In that respect CV and CM have the best credentials. They have been applied many times. Especially CM is a routine in marketing (Cattin and Willing (1989)).

The fifth is the *number of observations* needed, which may be rather low in CM, but must be large for all other methods. For CV and CM oral interviewing is frequently preferred, while EA does not even always require primary data and Leyden and Cantril are routinely used in mail interviews.

The sixth is the difficulty of *setting up the research*, where Leyden and Cantril are the most simple to devise and implement the questionnaire for.

Seventh, the *predictive quality* may be partially tested by splitting samples and using the estimated model to predict results for the other half of the sample. Another possible test are the stability of the effects over various samples in time. In that respect CM, Leyden and Cantril score well. Nevertheless it will always be difficult to predict situations which are relatively new.

Eighth, *cost*s of using the method. For EA we may utilitise existing data sets, although it may be difficult in practice to link various data sets. CV calls for relatively large ad hoc data sets which may be derived from mail questionnaire although it is frequently based on expensive oral interviews. The same holds for CM analysis. Here however the number of interviewed may be kept relatively small.

Leyden data are relatively rare but Cantril questions have been put to respondents all over the world. We do not believe that any method is typically cheap. It depends on (nine) the availability of data which can be used for secondary analysis.

Finally, the *turnover time* for the analysis, i.e., the time needed to set up and collect the data, can be relevant. We estimate that any data set except EA, because it needs a lot of administrative data from various sources, may be quickly set up and collected, such that the total turnover time will be in the order of six months.

As it appears we do believe that it is possible to enrich present macro- and micro models by adding information derived from CM, Leyden and Cantril-type data collection. We are more pessimistic about EA and CV, although they may be very useful tols in some limited circumstances.

In the next decade, where the micro-foundations of macro-economics will become still more prominent, we expect and hope for a significant contribution from those techniques, te get a better grip on reality in macro-economic models.

Literature

Brickman, P., Campbell, D.T. (1971), "Hedonic relativism and planning the good society", in Apley, M.H. (ed.), *Adaptation-level theory: a symposium* . New York: Academic Press.

Cantril, H. (1965), *The pattern of human concerns* , New Brunswick, Rutgers University Press, New Jersey.

Clark, A.E., Oswald, A.J. (1994), “Unhappiness and unemployment” *The Economic Journal* , 104, pp. 648-659

Diamond, P.A., Hausman, J.A. (1994), "Contingent valuation: is some number better than no number?", *Journal of Economic Perspectives* , vol 8(4), pp. 45-64

Frijters, P. Van Praag, B.M.S. (1996), “Climate equivalence scales and the effects of climate change on Russian welfare and well-being”, Tinbergen Discussion Paper 96-91/3, University of Amsterdam

Hagenaars, A.J.M. (1986), *The perception of poverty*, North Holland, Amsterdam.

Hanemann, W.M. (1994), "Valuing the environment through contingent valuation", *Journal of Economic Perspectives*, vol 8(4), pp. 19-43

Hausman, J.A. (1993, ed.), *Contingent valuation. A critical assessment*, New York: North Holland.

Kahneman, Knetsch, (1992), “Valuing public goods: the purchase of moral satisfaction”, *Journal of environmental Economics and Management*, 22, pp. 57-70.

Kamp, M.A., Maxwell, C. (1994), “Exploring a budget context for contingent valuation estimation”, in Hausman (ed.), *Contingent Valuation*, North Holland Publishing Group, Amsterdam.

Kapteyn, A. (1977), *A theory of preference formation*. PhD. Leyden University, Leyden

Lancaster, K.L. (1966), “A new approach to consumer theory”, *Journal of Political Economy*, 74, pp. 132-157

Luce, R.D., Tukey, J.W. (1964), “Simultaneous conjoint measurement: a new type of fundamental measurement”, *Journal of Mathematical Psychology*, 1, pp. 1-27

Parducci, A. (1995), *Happiness, pleasure and judgement, the contextual theory and its applications*, Elbaum Associates, Mahwah, New York.

Plug, E.J.S., Van Praag, B.M.S., Hartog, J. (1996), “If we knew ability, how would we tax individuals?”, *Tinbergen Discussion Paper* 1-96-154, University of Amsterdam.

Roback, J. (1982), "Wages, rents and the quality of life", *Journal of Political Economy*, 90(6), pp. 1257-78

Roback, J. (1988), "Wages, rents and amenities: differences among workers and regions", *Economic Enquiry*, 26(1), pp. 23-41.

Rosen, S., (1974), “Hedonic prices and implicit markets: product differentiation in pure competition”, *Journal of Political Economy*, 82, pp. 34-55.

Tinbergen, J., (1956), *Economic policy: principle and design*, North-Holland publishing.

Van Beek, K.W.H., Koopmans, C.C., Van Praag, B.M.S. (1997), “Shopping at the labour market: a real tale of fiction”, *European Economic Review*, 41, pp. 295-317.

Van Praag, B.M.S. (1971), "The welfare function of income in Belgium: an empirical investigation", *European Economic Review*, Spring, pp. 337-369.

Van Praag, B.M.S. (1988), "Climate equivalence scales, an application of a general method", *European Economic Review*, no. 32, pp. 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 (1/2), pp. 69-89

Van Praag, B.M.S., Kapteyn, A., (1973), "Further evidence on the individual welfare function of income: an empirical investigation in the Netherlands", *European Economic Review* , nr. 4, pp. 33-62.

Van Praag, B.M.S., Van der Sar, N.L. (1988), “Household cost functions and equivalence scales”, *Journal of human Resources* , 23, pp. 193-219.

Van Praag, B.M.S., Plug, E. (1995), "New developments in the measurement of welfare and well-being" , presented at *The Ragnar Frisch Centennial 1995* , Oslo.

Vijayendra, R. (1993b), "The rising price of husbands: a hedonic analysis of dowry increases in rural India", *Journal of Political Economics* , pp. 666-677.