

Elicitation of Health System Benefits : economic rationality and the multidimensional utility analysis(*)).

Joan Costa Font and Joan Rovira Forns

Departament de Teoria Econòmica

Universitat de Barcelona

Diagonal 690, 08034 Barcelona

E-Mail : jcosta@eco.ub.es and jrovira@soikos.com

Paper for the UK HESG Meeting ,University of Birmingham , 6th -8th January, 1999.

(*) Paper in process, any comments will be welcome

Acknowledgements

The preliminary research has been supported by the EC Program called “Measuring and Monitoring the performance of health care systems in developing countries”. We are particularly grateful to Claudi Estva and Berta Alcaniz (Department of Social Anthropology, University of Barcelona), to Peter Sandiford (Liverpool School of Public Health and Tropical Medicine) and Emilia Sanchez (Servei Català de la Salut) for their comments.

1. Introduction

This paper explores a methodology to value health systems benefit (HSB) as macrodimensional welfare change introducing some aspects concerned with social rationality when eliciting preferences and values for health system benefits. It has been developed as an analysis of a former paper on valuing health systems benefits (Costa and Rovira, 1998). Reforming health systems is nowadays a current debating issue, especially if we relate to the introduction of new programs, social services related to health systems and other health care technologies that can affect directly or indirectly health related welfare of individuals. However, a high number of important decisions concerning to the size of the public coverage are being discussed without a relevant methodology to evaluate its economic consequences. The methodology purposed is based on the identification and valuation of the health care system outcomes as perceived by society, in order to guide health care policy and especially health care system reforms. The ultimate question that tries to answer this paper: in a health care system reform scenario, how can decision makers select which benefits should be improved?

The main results of the research are the development of a new instrument for valuing health system benefits, introducing a multidimensional concept of HSB and finally discusses the results in terms of rationality. More concretely, results suggest that despite individuals are coherence in logical terms, the economic rationality measured by the sensibility to changes in the cost is low. A multidimensional utility model has been displayed in order to guide empirical analysis such as the understanding of the methodology implemented. This is the initial outcome of an extended research on vaulting health systems benefits, which its final aim in to develop a social evaluation algorithm to guide health system reform policies.

The remainder of the paper is as follows. Section 1, introduces the conceptual background, discusses and justifies the further methodology such as some recent theoretical developments that have been taken into account. Section 2, focuses on the economic rationality applicable to this methodology. Section 3 departs from a multidimensional utility model trying to capture the different dimension's of benefits included in health care systems. In section 4, the basic results are presented and discussed.

(ii) Conceptual background

In an ideal (optimal) democratically context, decisions affecting remarkably society should be adopted by consensus within a decision process. Therefore, the decision maker should rely on individual social preferences to estimate some kind of empirical social evaluation function, defined as the generalization of individual social revealed preferences and values. That is, a social elicitation inducing enough evidence that an alternative "A is socially preferred to an alternative B¹".

¹ That is a social process of evaluation by which A has been declared preferable than B

The present methodology departs from a realistic context where policy makers and politicians are not removed from citizens and users. By the contrary, they try to identify and value health system benefits in terms of welfare gains. However, in most cases policy makers do not dispose of enough information about outcomes, even though sometimes there's no consensus on its determination and prioritization. Thereby, a rift appears between health decisions and social health expectations (Sandiford,1997). It seems reasonable to develop a methodology and clear instruments in order to study the health care systems outcomes and their impact in social welfare.

Individual's perceive health outcomes only as one dimension of health care system outcomes (Mooney,1985), that is, there are non health benefits - other dimensions of outcomes - derived from health care systems that affect individual welfare. So we define health system benefits as any *characteristic of health care systems that affect positively or negatively individual's welfare*. However, HCB can be classified in terms of health benefits and non health benefits, despite in this paper a 5-dimensional classification has been considered (Table 2).

Individuals are not conceived as just consumers or users of health services, but as citizens. That is as members of society, that perceive welfare gains provided from the existence and functioning of a health care system, regardless of being recipients of them (Calnan,1997). Thus, the effects of health care systems over social cohesion, or the increase in equity, can be considered as valued as much as than the capacity of services to prevent illness or to restore health. According to Williams (1996) :

"The counter argument is that priorities are set by the whole society and that, if what would give an individual satisfaction clashes with what society requires of the system, then its the social norms which should prevail".

The "citizens approach" focuses on a global vision from which benefits are socially preferred. Given that the purpose of the methodology is a social global value - despite the pilot nature of the experiment - we assume that society can be divided into a finite number of social homogeneous groups, composed of a reduce number of participants representative of a social strata, where every group is qualitatively different than other's, while within the groups there are no significant differences.

We conceive a collective decision process for simplicity in three steps. . A first step where participants identify which are the outcomes -positive or negative - that they perceive from the health care system. A second step based on ordering and valuing different "alternative health care outcomes" in terms of a health care programs. Finally, the third step consists on the aggregation of the individuals preferences and values into a collective choice. However, the present paper analyzes the first two steps of the collective decision process as a previous tool to aggregate individual - in terms of social groups - preferences.

This approach seems consistent with the new welfare economic theory. According to Debreu (1964) social welfare evaluations can be understand as separable in two steps, a first step where the group welfare is considered separately and a second step where welfare is aggregated for the whole groups of society. Some kind of social rule should

be inferred from individual social preferences. In this context the experimental approach seem to be operative and consistent with theoretical requirements. In our case, the aggregation of social preferences lies in a simple proposition " if we consider the population segmented in social groups (decentralized choice) with similar characteristics and approximately equal income levels within a group , then individual information can be aggregated without a significant lack of information" (Brekke, 1996).

One of the major limitation in a social collective choice formation is the representation of preferences and its intensities. One of the ways usually accepted to avoid this problem is to suppose that participants maximize a determined social utility function , but this provides us a clearly restricted information if we do not know which are the benefits perceived from the health care system. According to Brekke(1996), we assume that given a sufficient information, individuals are capable to compare social welfare from different social health states and rank them according to their benefits. However, we should define what we understand by utility in this context. The concept of utility in this paper captures the discussion of the agency aspect (Sen ,1986), defining utility as : "*the extent to which the individual 's collective personal interest is fulfilled*".

(iii) *Justifying the Methodology*

The measurement of health outcomes in health economic literature is being mainly carried out by the use of the Quality Adjusted Life Years (QALYs) and more recently the Healthy Years Equivalents (HYE), that assumes a priori that health maximization is the unique objective of health care ². Nevertheless, - and without dismissing the advantages of using a priori defined measure as QALY's are³ - the use of this measure when valuing HSB implies rejecting the multidimensional nature of the individual utility functions and the heterogeneous preferences over health systems (Olsen,1997)⁴. QALYs are a unidimensional measure of health outcomes appropriate well when valuing "micro welfare changes" , where their effects are exclusively reduced to health gains. However, despite recent advances in terms including equity , QALYs do nor account differences in socioeconomic characteristics over individuals and the process utility is excluded from valuation. Therefore, when there is a macro -welfare change , such as a health system reform (involving a large number of programs) other methods should be developed

According to the former arguments, and in order to consider the fact "that there is more than just health in the individual's health related preferences " , the present methodology do not assume a priori any dimension of health system benenfits. By the contrary, individuals are to identify the underlying outcomes, and thereby define implicitly the set of possible dimensions. Once the outcome dimension has been defined, the outcomes are raked and individually valued , and finally aggregated by groups according to a definite rule.

² This is explicit in the cost-utility analysis's techniques and cost-effectiveness when health is defined in terms of QALYs.

³ For an extended discussion see Williams (1996).

⁴ QALYs are unfair values of HCB, experimentalists show that the general public do not share the underlying values .

A set of famous priority setting experiments have been developed in health care with similar objectives as this research, though with different assumptions and methodology (the Oregon (US) experiment, the Somerset (UK) experience and the Norwegian and Swedish experiment). These experiments were implemented without clear foundations since they display close practical aims. However, the realism of the experiments that always introduce more complexity, permit to consider them as a clear paradigm of the elicitation of preference and values in health care regarding an open and accountable prioritization of public resources.

Furthermore, other similar methodologies have been developed. Cookson (1997) has been a reference guide the experimental part of this paper, actually he introduces a new approach to placing values on health and safety by using a qualitative technique (focus group) and the willingness to pay. In this paper we've adapted some of the improvements of the Cookson research in to a more general experiment. At the same time the quantitative technique has been modified according to the dimension of benefits expecting to cover.

(iv) *A New Approach for Revealing Health System Outcomes : describing the methodology*

What we call the new approach is based on combining the use of the focus group (qualitative technique) instead of the standard questionnaire approach with a monetary valuation instrument (quantitative technique). This methodology, though has been design for evaluating health policies, can be applied to any area of public policy for engaging priority - setting⁵. The focus group technique consists on a semi-structured group of individuals –commonly between 6 and 15-, coordinated by a manager who tries to capture information about personal experiences and feelings concerning to a determined issue (health care systems) by a discussion of some delimited questions. The discussion group should be active, and participants can answer a list of questions requested from the coordinator, or even ask for other questions related to the general topic research. This technique sacrifices the large sample of alternative techniques, but gains a great depth of information, although maintaining of survey design, rigor and so on (Cookson,1997). Its main point is its high efficacy for obtaining rapid information specially interesting to interpret the qualitative motives of quantitative responses, improving therefore individual expressed valuations.

The basic aim of this approach is to ask very precise but general questions to a reduced sample of a population representative of a social group. Thus, the basis of the sample is a set of individuals classified according to socioeconomic reasons. Moreover, participants should be given an information-processing task that they are capable to deal with according with its cultural and educational level-. Therefore, a correct understanding is essential in order to eliminate possible bias in the final results. Probably the main advantage of this method is the possibility to discuss the answers and to lead to a consensus that permits further group aggregation results. The standard value elicitation survey do not permit to improve reasoning and solve apparent contradictions of answers, checking after a discussion the former answers. Thus,

⁵ Actually, some paper have been identified using the focus groups in other fields as agricultural economics.

preferences and values obtained would be by using this method are expected to express more precisely real decision makings in any field. Can be defined as true preferences, since they are expected to display higher consistency levels⁶.

The quantitative method used for incorporate in an homogenous measure the benefit values has been an open ended "willingness to assign" an increase in 4000 million pesetas in the health public sector budget, as if they where decision makers. This method essentially asks individuals the maximum amount of money they would be willing to assign for a single benefit improvement subject to a maximum amount of financial resources available. The reason for modifying the willingness to pay, lies on the intention of presenting a realistic experiment, where scarcity of resources is captured. In fact, there was another possibility, the use of willingness to pay an taxes, however, because of technical inconsistencies we've rejected this option Therefore, participants adopt a broad perspective on the hypothetical health policy situation. Also, the sensibility of answers respect to the cost information is considered as a measure of a personal economic rationality, such as a measure of cost-preference sensibility.

2. Economic Rationality and Reasoned Preferences

When alternative programs are presented to be valued by individuals, the final valuation is expected to be consistent when answers are guide by some logical sequence that reflects a logical adequacy. However, rationality is something more than a pure logical adequacy, rationality then eliciting health care programs should encounter the question: are individuals able to maximize "something" according to some former goals defined in terms of health system benefits?. Preferences and values assigned with this approach can be viewed as "reasoned preferences" due it combines the qualitative preliminary approach with a quantitative approach that permits participants capture the overall effects of every programs, and the relation between every program and the underlying benefits.

The basic assumption underlying the preference elicitation is that individuals display rational preferences. Preferences can be represented as a numerical value in terms of "utility theory". The basic statement is somehow tautological, "if a person prefers and alternative x to an alternative y, the value of x $v(x)$ should be higher than the value of y $v(y)$ "⁷. However, we should distinguish between rational preferences related to individual behavior where the final outcomes depend on the interaction with other individuals from the approach that we are issuing in this paper. Here, individual know the possible set of alternatives and they value every program without any incentive to

⁶ By contrast, the traditional standard questionnaire approach lies on the design of precise and concrete questions to a large sample of population. The precision level is fixed and usually is high representative, however there's a great loss of information due to the method do not permit to improve the depth level fixed in the amplitude of the questionnaire design, despite the simplicity for further quantification of final results.

⁷ This is stated under the assumptions of completeness (for all x and y, $x \succ y$ or $y \succ x$), transitivity (for all x,y,z if $x \succ y$ and $y \succ z$ then $x \succ z$) and monoticity (if $v(x) \geq v(y)$ and $v(x) \neq v(y)$, then $x \succ y$). The same could be defined in weak terms if the indifference is included or not.

co-operate or to adopt strategically decisions due to they have no information of the aggregation procedure. However, individuals are asked to act as citizens , where a clear component of altruism is relying, therefore rationality in this case may differ from the pure hedonistic approach, here egoism can be viewed as inconsistent.

In order to analyze the responses we need some instruments that can distinguish the rational responses from other. In this study we have obtained ordinal and cardinal information about individual social preferences .Therefore a first test is the *reliability test*, related to the variation in true values within the method of measurement. In this paper this has been measured by as a MLE estimation of the willingness to assign (WTAS) as a endogenous variable in monetary units and the ordinal ranking (PR) in ratings from 1 to 10, as the exogenous variable $WTAS = \alpha + \beta(PR)$. The null hypothesis is that $\alpha = 0$ and $\beta = 0$ at a 95% confidence level.

A second test implemented has been the *consistency test* . This test tries to capture if individuals are completing the priority setting exercise in a meaningful way. Consistency refers to the adequacy between the decision model specified and the empirical results, "an answer is consistent if its choose the theoretically preferred option" , participants should value higher the preferred benefits. Therefore, "if participants prefers A to B, then s/he should assign at least the same monetary units , according to the weak preference criteria. Furthermore, consistency can be defined in a dichotomic way as "consistent or not" or as a relative value, that is "results are consistent in a defined percentage". The later concept has been assumed in this paper. The indicator used is constructed as $D_c = \left| \sum_j PR - R_{WTAS} \right| / MD$ where PR refers to the ranking exercise , R_{WTAS} the resulting rankings in a 1 to 10 scale of the WTAS exercise, and the MD is the maximum possible distance between rankings. Finally, a third test refers to the economic rationality test, that is when cost is included as an attribute , then the economic rationality tries to capture if individuals re-scale its former PR. If cost is independent of individuals preferences, then people are economically rational. The scaling procedure is the same as the consistency test but it refers to the distance between the RP and the cost priority (CP).

3. The multi-dimensional utility analysis (MDUA) : an outline of the model

(i) The Social Evaluation Function

A Social Evaluation function is a numerical representation of an individual social preferences over some courses of action, such a health care programs presumably embodying the welfare value judgements of society over health system benefits .According to Harsanyi (1997) a social preference can be stated over a set of uncertain prospects , and under specific conditions that can take the form of a weighted sum of individual utilities . Considering a society consisting of homogeneous social groups $I = (1, \dots, n)$. Let $Y = (Y_1, \dots, Y_n)$ be an allocation set of possible health programs (Table 3) reflecting each one a particular dimension of health system benefits $X = (X_1, \dots, X_n) \in R^n$. Therefore, for all social states S a measurable subset of Y, $Y(X)$ is monotonic transformation of X. The assumption of the paper is that each

individual values each Y according to a social welfare judgement as if s/he was a decision maker that should care about the overall social welfare :

$$V_i(Y(X)) = W^j(v^j(X_1), \dots, v^j(X_n)) \quad (1)$$

where $v^j(X)$ is the value function, that is a numerical individual welfare judgement according to a sufficient information to be assessed to every programs. We do not assume that all individuals arrive to the same judgements among society whilst we assume that inside every group there would be a high consensus on which benefits are valued higher.

(i) *The multidimensional utility analysis*

Multi-dimensional utility analysis (MDUA) is concerned on eliciting preferences and values of health care programs outcomes, as reflecting predetermined HSB. HBS in this study were identified and classified according to a five dimension system shown in Table 2. Each dimension is subdivided in a number of concrete benefits such that all benefits revealed by the population can be classified into every dimension. A social evaluation function is implied over the different dimensions of health system benefits. For simplicity we adopt the additive utility function and therefore we assume additive independence. The multidimensional utility function is as follows :

$$v(X^1_i) = \sum_{j=1}^5 k_j v_j^1(X_j) \quad (2)$$

where k reflects the relative valuation of the utility of the benefit X_j in the total individual social value function of individual i of group 1, where :

$$\sum_{j=1}^5 k_j = 1 \quad (3)$$

in these exercise the total valuation would be the number of marks assigned to every program (benefit).

(ii) *The willingness to assign approach*

Let us consider a citizen acting as a decision maker that has to assess an amount of monetary units as a result of a budgetary increase to a set of health care programs $Y = (Y_1, \dots, Y_n)$ reflecting each one a particular dimension of HSB $X = (X_1, \dots, X_n) \in R^n$. The initial health care programs available within a system are known $Y^0 = (Y^0_1, \dots, Y^0_n)$ while their associated budget is $M^0 = (M^0_1, \dots, M^0_n)$. A

health system reform is viewed as an extension of the former programs $Y^* = (Y^*_1, \dots, Y^*_n)$ that has an associated budget of M^* that should be assessed according to its benefits. Then the willingness to assign approach tries to determine the amount of monetary units that independently of costs every participant would assign according to its social welfare gains :

$$WTAS = E(V^*, Y^*_j) - E(V^*, Y^0_j) \quad (4)$$

where V^* reflects the final welfare value . Finally, a measure of the social effects of every program can be reproduced as:

$$SNB_j = \frac{WTAS_j}{C_j} \quad (5)$$

where SNB refers to the social net benefit, and C refers to the cost of every program j .

4. Methods

(i) Design of the Experiment

The experiment scheme design required two months investigating the different alternative techniques and the possible bias that could appear. Finally , due to the experiment aims the design has been realized as follows. The pilot experiment has been developed in close consultation with two social researchers that were familiarized in the field of social investigation and anthropology. The experiment (or pre-tests) has been implemented to six groups in two different exercises separately in two different sessions, between December and June 1998. It consisted in a first focus group (in order to obtain health system benefits) followed, in a second session by a participated survey ⁸(in order to quantify the qualitative results). Sessions lasted about two hours and a dinner was paid to each group. Before every session , participants were requested to fill up a personal questionnaire.

The experiment Scheme involved three different parts :

- (i) Part one : identification of which health system benefits every group perceived and was able to understand from the health system , regardless of its real provision .

INTERPHASE 1: transcription of results, codification of HSB, design a set of health care programs that capture such benefits and determining by means of a focus group the connection between health care benefits and programs .

⁸ The participated survey in a social research instrument very close to focus group, where participants after an acceptable understanding of the objectives of the research, due to a pre test discussion, and the meaning of every rule and exercise of the experiment, have to fill up a written survey by reasoning its answer.

- (ii) *Part two* : prioritization and valuation of the former outcomes in terms of health care programs. An additional exercise was implemented in order to capture the cost sensibility of individuals .

INTERPHASE 2: selecting and classifying the results and testing the consistency and economic rationality .

- (iii) *Part three.* Analysis of the final results with statistic descriptive measures and further discussion of results. A net social benefit indicator has been computed as an additional indicator for the health policy aims. Moreover, a set of consistency tests were implemented , the rationality of the elicitation process was taken into account and the similarity of the responses was computed by using the Euclidean distance measure (table 9).

At the first part of the experiment was introduced the focus group approach as a mechanism to reveal health system benefits. All the socioeconomic information about ages, health care interventions , and professional active and residence was considered in order to segment population. The selection of groups was realized in consultation with social researchers . The time availability was considered such as the willingness to participate. Therefore, the interesting group profile should be one that accomplish with some active mobility conditions considering the interpersonal social relationship network and the sufficient extent number of people (6 to 15). As a result, the groups selected for the pilot experiment in Barcelona city were three (see table 1).

In order to prepare this part of the experiment, a reference guide was elaborated containing the basic aim and criteria adopted , the identification of the moderator and other collaborators , the kind of dynamic conversation predetermined, the practical and operative rules, the role of every participant, resources and additional material etc. At the same time, the basic questionnaire of the experiment was prepared and mailed to a contacting persona of each group.

The first exercise was successful, every participant contribute to the discussion with its opinion and arguments, and due to the discussion some people could deliberate and finally change its first argument. Since the intervention of the moderator was active, any participant contribute significantly more than others, thereby the results can be considered as the reflection of the group as a one member opinion . The first outcomes of this exercise where a two hours taped recorded conversation, that was finally transcript into a text . With the final text, the next phase was to identify the general homogenous health system benefits that where classified with the intervention of health system experts, in terms of Coverage, Accessibility, Quality, Public Health and Social Dimensions (table 2) .

The identification of health system benefits tries to reflect in a simplified term the basic dimensions of outcomes. The outcomes that every participant reveal can be positive, that is benefits for the health care system, or negative , that is shortages of health care systems. In order to consider this perception , health system benefits present a positive (+) remark and health system deficiencies a negative (-) remark. This is probably, one

of the advantages of the methodology, health care outcomes are not elicited by the researcher but is the society considered in terms of social groups who defines the basic dimensions of outcomes. With the identification of health system outcomes the first part of the experiment is finished (see table 2).

The second part of the experiment is the prioritization and valuation of health care outcomes. In order to design this part, it was considered the possibility to use public policy scenarios in order to encourage participants to come to an appropriate understanding of the policy situation in which their expressed valuations might be ultimately used (Cookson,1997). However, as our aim was not vaulting scenarios, but eliciting priorities to outcomes derived within the global health system, it has been implemented an indirectly outcome elicitation while the scenarios approach will be tested in future studies. That is, instead of eliciting HSB, HSB have been elicited in terms of health care programs assuming that every program relate to a concrete outcome and the magnitude of outcome that rises every program is the same and constant (see table 3)

The health care programs list was designed with the consultation of researchers of the Catalan Health Service (Catalan Health Public Administration). They assigned to every outcome considered, a "new" but applicable health care program that supposed to be an increase the set of services that till the date was publicly provided. Furthermore, they assigned to every program an approximated cost computed accordingly to the basic direct and indirect costs that could generate. Thus, the experiment tries to be a real public policy scenario in order to capture the attention and interest of participants. When selecting the number of programs, was taken into account the results from the focus group because this part of the experiment is viewed as a continuation of the former one. Additionally, in order to elicit relative valuations across outcomes the magnitude of public risk reduction was considered implicitly for participants, but not as an objective measure. Every program (health care system improvement) was briefly described to respondents on separate information cards with no preliminary order, one of which is reproduced in table 4.

Moreover, a response questionnaire was design and mailed with the program description to the participants before the elicitation experiment. A reference guide was prepared with similar contents as the one prepared for the former part of the experiment explaining the participating survey technique. This is similar to focus group in the sense that permits respondents to change its answers after a previous discussion, but differs in the answering method, because respondents should answer in a determined sheet.

With regard to the value elicitation methods (exercises) the three different types were as follows:

- **Priority Ranking (PR)**. This was a simple method to rank the different health care programs considered. The resulting valuations are a rank order of priorities, that is the ordinal utility of the different programs (and hence considering monotony, the outcomes). Participants after a description of the health care programs elicited a numerical priority to every program in terms of 1 (less preferred) to 10 (more preferred). This elicitation exercise tries to capture the ordinal

preferences of participants for every program . After the elicitation the results were published and every group arrived to a consensus about the priority setting.

- **Willingness to Assign (WTA).** This was an open ended question that considering an hypothetical budgetary increase, respondents were asked to assign the total monetary amount to every program. The hypothetical amount was calculated considering the addition of the whole programs costs, that is the total set of programs should not be totally covered by the budgetary increase.
- **Cost - Priority (CP).** This was a method for obtaining the sensibility of respondents to the cost of every program, that is after revealing the cost of every program the ranking exercise was repeated.

The first and third elicitation methods are well-known, whereas the WTA method is a new method especially designed for the purpose of this study. Participants in general terms, felt comfortable with these exercises, despite the large amount of programs led to some difficulties . Two individuals did not understand the “citizens view” , and they answered as “if they were users”. Also, the use of a budgetary exercise relative to WTA caused arithmetic difficulties, that were finally solved.

The format of this second part was as follows. It began with an explanation of the programs and a brief discussion moderated by a coordinator which lasted about 40 minutes. Following the explanation, subjects had about 20 minutes to answer the priorities, and each elicitation was published in a blackboard in order to discuss and reach a consensus over 10 minutes. Subsequently, the WTA exercise was answered in 20 minutes and with 10 minutes was discussed and a consensus was reached. Finally, the cost – priority exercise was answered without any discussion. The total time was about two hours, probably is not enough time so for further exercises should have to be extended

The third part of the experiment is described in the next section.

(ii) *Values and Preferences Estimations*

The results analysis tries to assess how well the methodology measures the values of the sample population considered . The final results have been estimated in arithmetic means according to Jones Lee- Lomes (1995) , however other central tendency statistics have been considered as the standard deviation and the median .

Before an extensive analysis of results an internal consistency test should be assessed. It seems reasonable that participants in the first preference exercise would not vary sensibly its results in the WTA exercise, that is if somebody prefers an outcome to another, then the monetary intensity of preference should be higher. Additionally, if participants do not reflect any variation in its results when cost is given, the result could be inconsistent.

The priority setting of the first exercise has been analyzed within every group by applying the Borda rule , due to all the participants have expressed its real preferences simultaneously. The Borda rule is probably the most simple method , every participant assigns its priority in terms of 10 the most preferred and 1 the less preferred and its

aggregated by summing up the results. The WTA has been estimated by summing up the total monetary amounts for every program, and the cost-priority as the priority exercise applying the Borda rule. While the first and third exercise denote a simple preference, the second one denotes a trade-off between every program, and logically the information contained in it is richer than the other results.

Not all the exercises were developed by the whole groups due to it was a pilot experiment. The exercise distribution was as follows:

	PR	WTAS	CP
G1	X		X
G2	X	X	X
G3		X	X
G4		X	X
G5		X	
G6	X	X	X

The WTA results were in general terms rounded number between 100 to 900 pesetas, and it should be noticed that respondents were not presented with monetary cues. All WTA were presented in tabular format in the same page. These results are absolute monetary values in arithmetic and median terms. The WTA has been aggregated by using this two central tendency indicator following the Jones-Lee and Loomes (1995) methodology. The median has been used in order to inform about the presence of extreme data.

Finally, it has been applied a multidimensional scaling (MDS) to the ordinal data results in order to state the similarity between programs. The aim of the MDS is to define the relationship between objects when the underlying dimensions are unknown. It is a mathematical tool that enables us to represent the similarities of objects spatially as in a map. MDS procedures represent objects judged experimentally similar to one another as points close to each other in a resultant spatial map. They do not require a priori knowledge of attributes whereas they provide a space that reveals attributes of stimuli to be scaled. The measure used has been the Euclidean distance and shows that there are two clear dimensions of HSB that can be interpreted as health system benefits and non health system benefits (Table 9)⁹.

5. Results and Discussion

In this section the main results from the pilot experiment are analyzed. Nevertheless these results do not explain the priorities and values of the population of Barcelona due to the reduced sample. The results are in general terms consistent according to the consistency tests used, only for a few individuals is not clear the rationality of the priority setting exercise. This could be explained because of the large number of programs. Table 5 shows that for the two groups that complete the PR and the WTAS, the 69,5% and 82% respectively were consistent, that is the results of the WTAS displayed similar rankings than the PR. The results from the reliability test denote that

⁹ See Schiffman et al (1981).

the sign is the expected and the explicative capacity is high at a 95% confidence level. Finally, individuals display a reduce level of economic rationality, since they do not change significantly their responses when cost is given. This could be explained by the fact that due to they do not actually pay the programs, the cost attribute is not really internalized by individuals.

Most mean rankings and WTAS responses are significantly different each other at a 95 % confidence level (Tale 6). Respondents believe that the cancer program, the waiting lists and the mammography revision should be the programs implemented in an hypothetical health system reform. That is, the aggregate results denote a high consensus in terms of the most valued program, the programs 1, 6 and 7 seem preferred, and refer the that health care outcomes (quality of inputs, coverage and accessibility) are the most valued. By contrast, the programs 4,5 and 10 seem to be the less valued reflecting the outcomes (quality in process and social dimensions). Finally, in terms of preference intensity, there are clear differences within the different groups, though there is a coincidence in the preference direction. Is expected that there would be differences between the PR and the WTAS since WTAS measures the intensity of preferences, exhibiting the possible trade-off between programs. However, those programs most preferred in the PR are the best preferred in the WTAS. The WTAS results (table 7) denote the value of each program for every group. Table 8, shows the estimated results of the social net benefit for every program in relative terms denoting the unit of benefit generated by a unit of program cost. From this results, and if this would be the results for the whole population, we could deduce that there are three programs that would not be clearly financed, the programs 3,9 and 10. The desirable programs would be number 1, 2 6, and 8. Comparing this results with the former ones we can find some inconsistencies in terms that are hoped to be solver in further research. However, in general terms, there are great similarities between the different exercise results.

6. Concluding Remarks

This paper tries to contribute to the discussion about how to value health system reforms. Its important to stress the pilot nature of this study . On the one hand, the methodology tries to test the population capacity to answer questions concerning to priorities and valuations of health and welfare. On the other hand displays a alternative methodology to identify health care outcomes that population perceive, and ranking them according to theoretically consistent indicators. Moreover, the basic advance of this methodology is the large definition of health system outcomes, that permit the integration of social dimension outcomes as equity, and thus trade-off between equity and efficiency could be taken into account, by measuring the relative values between those programs aiming to cope with equity and efficiency as a health system benefit respectively.

The results have been consistent despite the reduced sample extension despite the responses do not seem to be cost sensible . One of the main advances purposed for further research is the aggregation of the group results in to an empirical social welfare function. There are a large amount of alternatives, depending on the group weighing, nevertheless the utilitarian approach seems to be operative. The use of the focus group approach has eliminated the problems concerning on the misunderstanding of the

questions due to the possibility to discuss the advantages of every program. However, despite this approach avoids the problems of the standard questionnaire, there representative problems arise and results exhibit patterns of inconsistency.

Regard to the use of the WTAS, in spite of its difficulties for participants, tries to adequate the hypothetical answers to real elicitation's. Furthermore, it seems theoretically appropriate for this aim comparing with other alternative techniques. For instance, the use of willingness to pay taxes is sensible to the financing instrument, that is the type of taxes used and the aggregation of results would presents a large number of bias due to that fact. The use of willingness to pay, is difficult to implement when health public sector provides health services where there's not a clear price.

The experiment evaluation is clearly positive, in all session participants have succeed the objectives purposed, and the methodology has been testes positively. Participants have been apparently motivated despite the hypothetical pattern of the experiment. Moreover, the interest and seriousness of the responses has been guarantied in all moment, due to participant's have been able in every moment to justify its responses when it has been required. By the way, it can be assumed that in a real context , the motivation would be higher. In general terms, the exercises purposed where easier enough to de applied and participant were satisfied to collaborate in the research due to "ultimately they were helping themselves " as they "reveal" .

References

- Arrow, K.J (1951). *Social Choice and Individual values*. (Cowles Commission Monograph, nº12) New York: Wiley.
- .Brekke K,A et al (1996). *Allowing Disagreement in Evaluations of Social Welfare*. Journal of Economics vol 63, nº 3, pp 303-324.
- Calnan, M (1997). *Citizens, users and health care*. European Journal of Public Health, 7:1-2.
- Carey, A (1994) *The Group Effect in Focus Groups : planning, implementing and interpreting Focus Group Research*. En Critical Issues in Qualitative Research Method, edited by Janice. M.Morse
- Cookson,R (1997). *The Focus Group Approach to Valuing Health and Safety : A Trial Study*. Paper for the 25 Anniversary HESG Meeting, 2-4 july.
- Costa , J and Rovira J (1998) . *Valuing health system benefits*. Economic Science Association Meeting, Mannheim , 1998 (mimeo).
- Charny, MC (1989). *Choosing Who Shall Not Be Treated on the NHS* . Social Service and Medicine 28 (12): 1331-13338.
- Culyer, A, J (1991) *Health, health expenditures and equity*. Discussion Paper 83, Centre for Health Economics, University of York.
- Debreu, G (1964). *Continuity Properties of Paretian Utility*. *International Economic Review* 5:285-293.
- Diamond, PA and Hausman , JA (1994). *Contingent Valuation : Is Some Number Beter than No Number?*. Journal of Economic Perspectives ((4) : 45-64.
- Feldstein,A.P. (1966). *Research on the demand of Heath Services*. Milbrank Memorial Fund Quaterly,44 nº 4 , 2ª parte pag 143.
- Gafni,A (1991). *Willigness to Pay as a Measure of Benefits*. Medical Care, December, vol 29, No 12.
- Gonzalez-BlocK, MA and Sandiford.P (1997) *Medición y monitoreo del desempeño de las reformas de los sistemas de salud*. Proyecto Eurpean Cimission, Monitoring and Performance in Developing Coutries.
- Harsanyi (1997). *Utilities, preferences and substantive goods*. Social Choice and Welfare.
- Mooney G(1995) *Efficiency in health care : just health gains?*. Australian Journal of Public Health, 1995 Aug, 19(4): 330-5.
- O'Brien,B Gafni, A (1996) *When do "Dollars" Make Sense ?*. Medical Decision Making, 16 :288-299.

- Olsen, JA (1991).** *A Social Welfare Function for the Norwegian Health Service*. Paper presented at the 12th Nordic HESH meeting, 19-21. August 1991, Copenhagen.
- Olsen JA (1997)** . *Aiming to what? The objectives of the Norwegian health service*. In Alban, A and Christiansen T editors: *The Nordic Health Care System* , Odense University Press, Odense , 1994.
- Schiffman, S, Reynolds ML, Young FW (1981).***Introduction to multidimensional scaling*. Academic Press inc.
- Sen AK (1970)** *Collective Choice and Social Welfare*, Holden –Day INC, San Francisco.
- Sen (1986)**, *Social Choice Theory*, in KJ Arrow and M Intriligator eds, *Handbook of Mathematical Economics*, vol III, pgs 1073-1182.
- Williams , A (1993)** *Priorities and research strategy in health economics for the 1990s*, *Health Economics*, 2, 2-95-302.
- Williams(1996)** .*QALY's and ethics : a health economics perspective*. *Social Science and Medicine*, 43, 12,-1996 1795-1804.

Table 1.
Participants in the pilot experiment.

	Segmentation criteria	Educational level	Group ¹⁰ Extension
G1	Middle Income	Retired workers with lower studies	8
G 2	Low Income	Undergraduate Students	13
G 3	Middle High Income	Skilled workers	8
G4	Middle –Low Income	Industrial workers	15
G5	Middle Income	Rural workers	9
G6	Middle – High Income	Post graduate students	13

Table 2
Health System Benefits identified

- 1. Coverage (X_1)**
 - 1.1 Indemnity (+)
 - 1.2 Education/ health information (+,-)
 - 1.3 Periodic Revisions (-)
 - 1.4 Odontology (-)
 - 1.5 Medicaments (-)

- 2. Accessibility (X_2)**
 - 2.1 Waiting Lists (-)
 - 2.2 Access directly to specialist's (-)
 - 2.3 Knowledge about the administrative process (-)

- 3. Quality (X_3)**
 - 3.1 Inputs
 - 3.1.1 Quality and professional ethics(+)
 - 3.1.2 Technological Equipment level (+)
 - 3.1.3 Comfort and Appearance of Installations +)
 - 3.2 Process
 - 3.2.1 Clinic informational systems (+)
 - 3.2.2 Equip. working (+)
 - 3.2.3 Personal medical work (+,-)

¹⁰ Each individual provides 30 observations.

3.2.4 Patient information (-)

3.2.5 Personal Treatment (-)

3.3 Outcome

4. Public Health (X_4)

4.1 Health Education at School(+,-)

4.2 Control and food hygiene(+,-)

5. Social Dimensions (X_5)

5.1 Equity/ privilege (-)

5.2 Consumer Sovereignty/ User freedom (-)

Table 3.

Health care programs list for the population of Barcelona

Program	Expected Outcomes and Costs (millions)
1. Breast Cancer : (Biennial Mammography to all the women between 50 and 65 years old)	(Coverage 1.3) Mortality reduction in a 15% <u>Cost : 150</u>
2. Coordination between primer and specialized attention (programmed meetings between Primer Attention Centers and specialist's)	(Quality 3.2.2) Improvement of quality attention due to a major integration and coordination of patient attention . Cost :100
3. Program for Professional Immediate Access to Clinic Historical Information (Computational System to obtain clinic information of patients immediately all over the system)	(Quality 3.2.1) Improvement in the efficacy and quality of attention <u>Cost : 500</u>
4. Attention and User Information (diffusion champagne on the existence of user attention units)	(Accessibility 2.3) Improvement in the access , information and user's satisfaction. Cost: 100
5. User Treatment (Formation program with courses till 30 hours for the whole administrative personnel with public treat.)	(Quality Process 3.2.5) Improvement in patient satisfaction <u>Cost :300</u>
6. Medical Revision (Volunteer medical revision for the whole population one every three)	(Coverage 1.3) Precocious detection and health consultation about hypertension, alcoholism, gynecological revisions, ect <u>Cost:500</u>
7. Waiting Lists (Plan for reducing accumulated waiting lists in non elective surgery)	(Accessibility 2.1) Reduction of the waiting time of 240 to 120 days <u>Cost:700</u>
8. Life styles (informational program on habit's and healthy lifestyles formation addressed to scholar's parents	(Public Health 4.1) Promotion of healthy life styles, improve scholar's communication and collaborations between parents and child's. <u>Cost:100</u>
9. Odontology (Free provision till 12 years of odontology services, currently only extractions covered)	(Coverage 1.4) Improvement of dental health <u>Cost:2000</u>
10. Medicament Bonus (Possibility for acquisition an annual bonus per worker in order to avoid co-payment. This will be implement with a system related to declared income on the personal income tax.	(Social Dimensions 5.1) Improvement in equity . Nobody would pay more , but people with less income could be benefited. The actual situation is : people should pay the 40% of the medicament price while chronic illness a 10%. <u>Cost : 2000</u>

Table 4.
Example of Breast Cancer Program.

<p>Program 1 : BREAST CANCER</p> <p><u>Expected Outcome</u> : Mortality reduction in a 15%</p> <p><u>Description</u>: Extension to the risk population of a program to eradicate the breast cancer. It will be implemented by a mailing quotation in order to realize a biennial mammography to all the women between 50 and 65 years.</p> <p><u>Actual coverage</u> : Currently it covers the 15% of the population.</p>
--

Table 5
Consistency Tests

	Degree of Consistency	Degree of rationality	Economic Reliability Test : Regression Results
Group 1	-	13%	-
Group 2	69,5%	13,5%	31,39 (0,046) R ² =0,696
Group 6	82%	6%	36.45 (0,016) R ² =0,667

Table 6.
Preference Elicitation by groups according to the PR

	Mean Rating Difference G1		Mean Rating Difference G2		Mean Rating Difference G3	
	Value	Bilatera Signif	Value	Bilatera Signif	Value	Bilatera Signif
ATEN	4,7500	,001	2,5385	,001	4,0769	,000
BONU	2,1250	,002	4,8462	,000	3,6923	,000
CANC	7,8750	,000	8,6923	,000	7,0000	,000
COOR	3,6250	,003	5,2308	,000	7,3077	,000
LISTS	8,1250	,000	7,0769	,000	5,3077	,000
ODON	5,5000	,000	5,3846	,000	3,8462	,001
PPICHI	6,8750	,000	5,4615	,000	6,8462	,000
REVIS	8,3750	,000	9,0769	,000	6,2308	,000
STYLE	4,5000	,001	3,6154	,000	6,3846	,000
TREAT	3,2500	,005	3,0769	,000	4,0769	,000

*Significantly different from all other means in column at 95% confidence level. Ratings range from 1 (less preferred) to 10 (most preferred).

Table 7.

Aggregate Priority Ranking (median) , WTAS (arithmetic mean and Cost Contingent Ranking(median)

Prog/pa rticp	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
RESULTS PRIORITY RANKING										
G1	8	3	8	3	5	9	8	4,5	4,5	2
G2	9,5	4,5	5,5	3,5	2,5	8,5	7,5	4	7,5	2
G6	7	7	3	3	3	8	7	6	6	4
RESULTS WILLIGNESS TO ASSIGN										
G2	826,9	293,75	381,25	125	158,89	707,69	575	142,78	497,91	575
G3	516,67	260	291,67	341,67	350	635,71	1025	480	542,85	600
G4	579,33	295	462,67	268,57	344,28	617,14	669,33	349,33	465,33	405
G5	806,25	380	266,67	250	90	825	475	280	368,75	668,75
G6	787,5	509,58	328,33	230,9	270	766,67	528	336,81	582,91	293,75
RESULTS COST PRIORITY										
G1	9,5	4,5	5,5	3,5	2,5	8,5	7,5	4	7,5	2
G2	10	7	6	3	3	9	7	3	2,5	4
G3	8	4	6	3	2	8	8	4	6	3
G4	9	3,5	4,5	2,75	3,5	10	6	2,5	5	8
G6	7	7	3	3	4	8	7	7	6	3

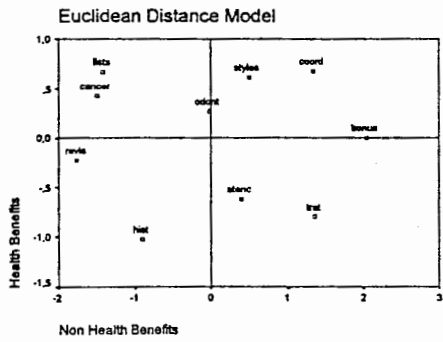
Table 8

Social Net Benefit Estimation

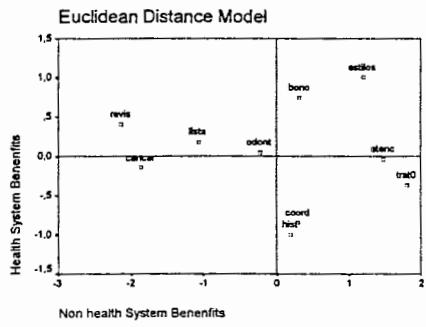
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
SNB G2	5,58	2,94	0,72	1,25	0,53	1,41	0,82	1,42	0,249	0,23
SNB G3	3,44	2,00	0,35	3,42	1,16	1,27	1,46	4,8	0,22	0,35
SNB G4	3,82	2,95	0,92	2,68	1,14	1,234	0,955	3,49	0,23	0,2
SNB G5	5,37	3,8	0,532	2,5	1,17	1,65	0,74	2,8	0,29	0,14
SNB G6	5,24	4,5	0,67	2,3	1,23	1,54	0,72	3,37	0,18	0,33

Table 9.
Results Multidimensional Scaling Model

Group 1 (S-Stress 0,102, RSQ=0,9377)



Group 2 (S.Stress= 0,49, RSQ=0, 9864)



Group 6 (S-Stress= 0,1045, RSQ =0,84)

