CENTRE FOR HEALTH
PROGRAM EVALUATION

WORKING PAPER 80

The Role of Willingness-to-Pay in in Resource Allocation in a National Health Scheme

Professor Jeff Richardson
Director, Health Economics Unit, Centre for Health Program Evaluation

August, 1999
ISSN 1325 0663
ISBN 1 875677 92 5

CENTRE PROFILE

The Centre for Health Program Evaluation (CHPE) is a research and teaching organisation established in 1990 to:

- undertake academic and applied research into health programs, health systems and current policy issues;
- develop appropriate evaluation methodologies; and
- promote the teaching of health economics and health program evaluation, in order to increase the supply of trained specialists and to improve the level of understanding in the health community.

The Centre comprises two independent research units, the Health Economics Unit (HEU) which is part of the Faculty of Business and Economics at Monash University, and the Program Evaluation Unit (PEU) which is part of the Department of General Practice and Public Health at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

PUBLICATIONS

The views expressed in Centre publications are those of the author(s) and do not necessarily reflect the views of the Centre or its sponsors. Readers of publications are encouraged to contact the author(s) with comments, criticisms and suggestions.

A list of the Centre's papers is provided inside the back cover. Further information and copies of the papers may be obtained by contacting:

The Co-ordinator
Centre for Health Program Evaluation
PO Box 477
West Heidelberg Vic 3081, Australia
Telephone + 61 3 9496 4433/4434 **Facsimile** + 61 3 9496 4424
E-mail CHPE@BusEco.monash.edu.au

ACKNOWLEDGMENTS

The Health Economics Unit of the CHPE is supported by Monash University.

The Program Evaluation Unit of the CHPE is supported by The University of Melbourne.

Both units obtain supplementary funding through national competitive grants and contract research.

The research described in this paper is made possible through the support of these bodies.

ABSTRACT

Argument and evidence are presented to support the claim that individual based willingness to pay is an inappropriate criterion for decision making in the health sector. The following propositions are supported:

1 that to achieve allocative efficiency we must eventually place a dollar value on life;

2 that proposition 1 does not imply conventional willingness to pay.

3 that willingness to pay for one's own life is an unhelpful concept in the context of a National Health Scheme.

4 that the population would reject the willingness to pay criterion as the basis for measuring benefits in a national health scheme.

5 that people accept paternalism; they are prepared to override other people's preferences and endorse the importance of extra welfarist objectives and, more specifically endorse the maximisation of health per se, not utility.

6 that the argument that social decision making is imperfect (Arrow's Impossibility Theorem) and that we must therefore have individual based WTP is wrong;

Despite the general rejection of individual based WTP, it must be accepted in some contexts. Consequently, total health resources must be allocated according to different criteria in different contexts and the boundaries between these contexts may shift over time and differ between countries It is concluded that social willingness to pay is a function of outcome, process and ***context***. Consequently, viewed from a contextless perspective decision criteria will appear incoherent and inconsistent.

TABLE OF CONTENTS

Abstract

1 Introduction 1

2 Theoretical Orthodoxy and Social Objectives 3

3 The Need for the Dollar Valuation of Life 6

4 Social Versus Individual Willingness to Pay 7

5 Qualifications 14

6 Conclusion 16

References 17

List of Tables

Table 1 Choice of Criterion: WTP vs Health Outcome 10

Table 2 Willingness to Pay for Health Services 11

Table 3 Paternalism vs Respect for Public Preferences 12

The Role of Willingness-to-Pay in Resource Allocation in a National Health Scheme

1 Introduction

Health economists have normally included two dimensions of outcome in the economic evaluation of health programs, namely, the quality and length of life. Because of the conceptual difficulty in placing a dollar value on lives or life years, as required by cost benefit analysis (CBA) this has often been replaced by cost effectiveness or cost utility analyses which appear to side step this conceptual problem and to provide enough information for the ranking of health programs[1]. It is further argued that this information is sufficient for any health system or institution with a fixed budget as health output will be maximised by selecting programs in the order of their cost effectiveness until the budget has been exhausted.

Recent literature has reflected a renewed interest in CBA and in the use of willingness to pay (WTP) techniques to place a dollar value on the quality of life and, more fundamentally, the value of life years (Donaldson et al 1996, 1997; O'Brien & Gafni 1996; Olsen & Donaldson 1998). This resurgence has been associated with the development of the techniques of contingent valuation in transport and environmental economics and their transference to health economics. Especially after the publication of the report of the NOAA Commission[2] in 1993 there has been increasing acceptance of the view that correctly conducted stated preference techniques may accurately predict revealed preferences and thereby increase confidence in the validity of stated preferences with respect to the quality of life.

1 WTP techniques have been used by environmental and transport economists to measure intangible benefits including both the 'option value' (of, for example, a wilderness) and the value of human life. It is the validity and even the meaning of the latter that is contentious.

2 Following the oil spill from the super tanker Exxon-Valdez in Alaska in 1989 stated preference techniques were employed to estimate the public's valuation of the environmental damage. The estimate was the basis of a law suit which fuelled debate over the validity of the techniques employed. To resolve the issue the US National and Oceanic and Atmospheric Administration (NOAA) established a panel chaired by two Noble Laurents in Economics, Kenneth Arrow and Robert Solo to investigate the issue. Their report (NOAA 1993) concluded that stated preferences were capable of reflecting true preferences if, but only if, a series of 'correct' technical measures were adopted to ensure the reliability and validity of the results.

There is also a common acceptance of the view that the value of life per sé may be inferred from individual willingness to pay for a reduction in the risk of death.

While the methodologies used to estimate the WTP for quality improvement and for life per sé have been quite different, the two concepts are now linked. The techniques for estimating quality adjusted life years, if correctly applied, convert quality improvement into an equivalent number of years of healthy life. Consequently, in a simple 'seamless' world—where contextual factors are unimportant—the monetising of either quality improvement or life years immediately implies a monetary value for the other.

2 Theoretical Orthodoxy and Social Objectives

Underlying the growing enthusiasm for willingness to pay is the often unstated assumption that WTP is the 'theoretically correct' metric for quantifying the value of a health program; that it is the 'approach suggested by economic theory' and, following Mishan's dictum, that 'it is better to measure the theoretically correct concept imperfectly than the wrong concept (Mishan 1971). This type of argument may at first appear to be compelling and to argue that 'the theoretically correct' approach is wrong appears almost self-contradictory. These assertions are, however, misleading and their force arises from, at best, the omission of important information and, at worst, from the misuse of terminology. CBA and the use of the WTP as an index of utility is based upon the theory of (new) 'welfare economics' which evolved after Robbins' (1938) seminal work and subsequently described in the classic works of Little (1965) and Graaf (1967). This might legitimately be described as 'orthodox' or 'orthodox neo classical theory'. But it cannot be described as 'economic theory' as if this were the only available theory or the only theory accepted by economists in all possible contexts. Other theories and other analytical frameworks exist. The decision making approach of Sugden and Williams (1978) for example is not based upon Welfare Theory. This and similar approaches do not assume the same social objectives or, at least, the same maximand in decision making as assumed in the neo classical orthodoxy. In sum, appeal to the authority of 'economic theory' is rhetorical and bordering upon the illegitimate. There is no 'economic theory' which equates with 'established truth'.

For this reason the most that can be legitimately claimed for the use of the WTP technique is that it is the measure suggested by 'one economic theory' or the measure that is consistent with economic orthodoxy. But with this restatement the apparent force of the initial claim is significantly reduced and attention is correctly focused upon the acceptability of the alternative theories.

Determining the relative merits of theories is itself problematic and when this issue is pursued the relationship between CEA and CUA on the one hand and CBA and the WTP on the other hand quickly becomes ambiguous. It is, for example, argued that the WTP is the most flexible technique for measuring the value of both a life year and a variety of additional factors that are important to the individual such as risk, hope, fear and other 'ex ante' emotions, prognosis, social relationships and other contextual factors. This may or may not be true. The extent to which QALYs can be extended to incorporate additional variables has not been properly explored[3]. However, a prior question is whether or not these factors should be included in economic evaluation and more fundamentally still, what is it that should be measured and included in the relevant outcome of economic evaluation.

The answer to this question in orthodox economics—in welfare theory—is unambiguous. We should measure 'utility' and, consequently, anything which affects an individuals preferences should be included in the analysis. Revealed preference under conditions of adequate information is, by definition the gold standard and, consequently, the stated WTP must be the preferred form of stated preference techniques subject to the caveat that measurement must be

[3] For an example of the inclusion of risk and prognosis in TTO measurement see Richardson, Hall and Salkeld (1996). For discussion of the need to extend QALY measurement see Menzel et al (1999) and Nord et al (1999).

'valid' (truly measured what it purports to measure). In this theoretical scheme CEA and CUA are simply pragmatic devices for avoiding the practical difficulties of using the WTP to evaluate life per sé and social objectives are some function of individual utilities.

Following Sen, these orthodox assumptions regarding the objectives of economic analysis have been labeled 'welfarism' (Sen 1982). By contrast with this, 'extra welfarism' – to use Culyer's (1991) term – is the theory that social objectives may be broader and include a variety of objectives which are not directly related to individual preferences; objectives such as basic capacities or, in the present context, life and a minimum quality of life. Culyer (1991) implies that extra welfarism is a broader theory than welfarism in the sense that it permits both utility and non-utility objectives to influence social welfare. Since there are a variety of candidates for the non-utility objectives, 'extra welfarism' is more properly viewed as a class of theories rather than a simple, coherent theory, and there is no reason why this class should only include theories that add additional objectives to welfarism. Coherent theories may entail the impotence of welfarism, or, at the opposite extreme, completely replace utility as the centre piece of social welfare in a particular context. Thus, for example, while individuals may, if allowed, reveal a preference for health programs that reduce the dis-utility of risk, that permit hope or minimise regret, there may simultaneously be agreement that these components of utility are not the appropriate subject for consideration in a tax financed National Health Scheme.

Viewed from this perspective, the use of CEA and CUA may reflect a fundamentally different theory of social objectives than the WTP. Rather than being a conceptually second best, pragmatic devices for avoiding the measurement problems of CBA they may come closer to the measurement of social values than the WTP. The theory that our society would like to maximise life, or health, rather than the monetary value that people with variable income and wealth would place upon life and health has a prima facie appeal which cannot be dismissed without evidence or argument. Rather than representing a rather dubious measure of utility, the QALY may be seen as the embodiment of a theory in which a particular combination of the quantity and quality of life is proposed as more closely approximating social objectives than previous measures; but a combination which can be progressively improved as we learn more about social objectives. From this perspective, the whole class of QALY-like measures may be tested as possibly best representing social values. This class includes the original one year focused QALY, the HYE, the DALY, the various unnamed combinations of perspective (social, individual, ex ante, expost); different time horizons (1 year, duration of a single health state; multiple health states) different measurement techniques (TTO, PTO, SG, WTP); and various combinations of value weights (age, SES, social characteristics). This plethora of measures may each be proposed and tested as best representing social objectives or as representing the best compromise measure when the practical problems of applied research with a finite research budget is recognised. Self interested WTP represents only one such measure – albeit a measure of special interest as it is closest to the welfarist ideal—but not thereby assured of ascendancy over other contenders for the position of best measure of relevant social value. This must be demonstrated by an analysis of social attitudes and an analysis which, ultimately, draws upon empirical evidence of what social values are held by the public and by patients after deliberation and ethical debate.

In sum, the theoretically appropriate metric in economic evaluation depends upon which maximand is implied by the theory which correctly describes our objectives and behaviour and this cannot be determined without recourse, at some point, to empirical evidence. There are

neither logical nor methodological reasons why objectives may not be context specific and the appeal of universal, context free, culture free, social objectives has probably more to do with the intellectual aesthetic tastes of analysts than with empirical observation. The applicability or otherwise of orthodox theory in other sectors where market forces have been allowed to determine the allocation of resources does not imply that the objectives of this theory also apply in the health sector where market forces have been specifically rejected as the basis for resource allocation. This suggests that the applicability of the WTP technique in the health sector need to be carefully re-evaluated.

The present article defends the following nine propositions which relate to these issues:

1 To achieve allocative efficiency we must eventually place a dollar value on life; sooner is better than later.

2 Proposition 1(a) does not imply conventional willingness to pay.

3 A willingness to pay for one's own life is an unhelpful concept in the context of a National Health Scheme.

4 (a) The hypothesis that people want a national health scheme only to provide the benefits that would be selected by private individuals using the willingness to pay criterion is empirically incorrect.
(b) If consulted, people would reject the personal willingness to pay criterion.
(c) People accept paternalism; they are prepared to override other people's preferences and endorse the importance of extra welfarist objectives.

5 (a) The argument that social decision making is imperfect (Arrow's Impossibility theorem) and that we must therefore have individual based WTP is wrong.
(b) Social decision making is (conceptually) easy.

6 Despite the general rejection of individual based WTP, it must be accepted in some contexts. Consequently, total health resources must be allocated according to different criteria in different contexts and the boundaries between these contexts may shift over time and differ between countries.

7 Social willingness to pay is a function of outcome, process and *context*. Consequently, viewed from a contextless perspective decision criteria will appear incoherent and inconsistent.

8 In practice the application of social willingness to pay techniques will probably be limited to the calculation of the value of a quality adjusted life year/DALY with this latter metric being used to evaluate the benefits from health program interventions.

9 Despite the previous propositions there is a case for the measurement of individual WTP but its role is comparatively modest.

3 The Need for the Dollar Valuation of Life

Proposition 1: To achieve allocative efficiency we must eventually place a dollar value on life. Proposition 1b: Sooner is better than later.

The truth of the first proposition is fairly self-evident. As long as we exclude services from a national health scheme we are, implicitly, placing a dollar valuation on life. The value of the lives that could be saved by the excluded services is implicitly less than the cost of these services. Whatever the size of the health budget there must be a boundary at which the value of lives is implicitly or explicitly equal to the opportunity cost. If the value of life rises/falls then the size of the budget should rise/fall. In the absence of an explicit valuation of life, as at present, there is no mechanism for achieving efficiency between the health and other sectors and it is highly likely that, at present, services are either wrongly included or excluded because of the lack of this information.

The common rationale for postponing a decision about the dollar value of life is that, using CEA and CUA, it is possible to rank services and achieve efficiency within the health sector. It is correctly argued that the available evidence suggests widespread inefficiency and it is concluded from this that the first priority should be internal allocative efficiency: if efficiency requires the expansion of some services then this should be funded by the contraction of others.

This latter conclusion is not compelling. Consider a scenario in which H_1 and H_2 represent two health programs. Economic evaluation establishes that for H_1, benefits > costs; and that for H_2, benefits < cost. The previous argument would suggest that H_1 should be expanded using resources from the contraction of H_2. This is certainly the optimal strategy in a first best world. However it is possible that, for political reasons (the power of providers or beneficiaries) it is impossible to contract H_2. In this case the options are either to fund H_1 from outside the health sector or to do nothing. If it is known that benefits exceed costs for H_1 then the former strategy is clearly better. The program should be funded and this conclusion is independent of the source of funds. However this conclusion cannot be reached on the basis of CEA and CUA. It can only be reached if benefits can be monetised[4].

[4] This conclusion may not always be true. If, for example, H_1 and H_2 were both hospital based procedures then the only long term strategy which will successfully contract H_2 may be an overall budgetary constraint upon the hospital thereby coercing the supporters and beneficiaries of H_1 to bring pressure on the supporters and beneficiaries of H_2.

4 Social Versus Individual Willingness to Pay

Proposition 2: The need to monetise the value of life does not imply individual WTP.

The first and most common objection to the use of WTP is that those with the greatest income have the greatest ability to spend and, consequently, their preferences would receive disproportionate and socially unacceptable importance. The usual rejoinder to this argument is that the willingness to pay may be adjusted to take into account the distribution of income.

This is only superficially a solution as may be easily seen with a simple example. Consider a society of two individuals A and B who receive respectively incomes of $180,000 and $20,000. If each was prepared to pay 0.05 and 0.1 percent of their income for the personal receipt of a service then the total, self interested, WTP would equal $90 + $20 = $110. Since this sum is dominated by the WTP of A the WTP could be recalculated after adjusting for the differences in individual incomes. For example, the percent of income that A and B are willing to pay for the service could be applied to the average income of $100,000. This would result in a total WTP of $50 + $100 = $150.

While this procedure overcomes the ethical objection to the unfair weighting of high incomes it also breaks the nexus between individual utility and revealed WTP. The $150 does not correspond with what the society composed of A and B would be willing to pay as this amount is $110. The adjustment of income has created an index of the strength of social preferences and this index could be compared with other indices. However, unlike the indices produced by the time trade off or person trade-off the precise meaning of this index is rather obscure. If the relationship between unadjusted WTP and income was linear then it would be possible to interpret the $150 as being an index of the summation of individual preferences where each individual's preferences are of equal importance. However with a non linear relationship the $150 represents the summation of weighted preferences where the weights are determined by the non linearity of the relationship. For example, if Individual A actually did have an income of $50,000 rather than $180,000 then she may have been willing to spend, not 0.05 percent, but 0.2 percent of her income. This would mean that her true preference (0.2) received a weighting of 0.25 (0.05/0.2) and by comparison with unadjusted WTP the procedure would heavily discriminate against the wealthy Individual A[5].

In sum, the individual based WTP may either adjust or not adjust for income differences. In the latter case the WTP may be unacceptable on equity grounds. In the former case the WTP no longer represents the dollars which society is prepared to pay and, additionally, is likely to discriminate against some groups in the community. Under these circumstances there is little reason to give special status to individual WTP. It does not capture the requirements of welfarism. Consequently, there is little reason to ask individuals a question which, from the perspective of orthodox theory, is irrelevant, and, as an index of wellbeing, is ambiguous.
A further objection to either form of WTP based measures of personal preference is that the payment made by each individual is largely unrelated to the benefits that they will obtain from a

[5] It is, of course, possible to calculate the non linear relationship from a large enough sample of respondents. This may overcome the problem of unequal weights but not the problem that the WTP from adjusted income no longer represents the actual WTP.

National health scheme. Individual payments will be primarily for the health benefits of others. This is also true for any form of social decision making. However it is particularly damaging for the WTP as it, uniquely, derives its authority from the ethical proposition that people should receive benefits which, on the margin, are equal in value to their willingness to pay for the benefits. Other forms of social averaging are not based upon this normative requirement. To comply with the normative requirement the relevant WTP question becomes how much each individual is willing to pay for other's health services but with the important caveat that the individual will be eligible to receive the services that others receive. As in the famous case described by Rawls, the individual should be behind a 'veil of ignorance': (s)he will not know whether or not (s)he will need the benefit and (s)he may not even know the likelihood of needing such a benefit.

The personal WTP for a service is likely to diverge from the social WTP; that is, from the amount the individual, acting as a citizen, will be prepared to pay, through taxation, for the addition of a service to the national health scheme. The entirely self interested individual would be prepared to pay a maximum amount based upon their subjective estimate of the likelihood of using the service and the dollar estimate of the benefits they would receive. Acting as a citizen the individual is likely to adopt a criterion which is different in kind. While influenced by self interest the citizen is also likely to be influenced by a set of social considerations; their empathy for others and their views on the nature of the society in which they wish to live. There is clear empirical evidence that issues of equity and social justice are of importance to the population in the context of the collective financing of health services. The communitarian values underlying these social values were clearly articulated in Canada by Evans and Law (1995) when they observed that 'Canada's system of universal public insurance for health care is … widely regarded as an important symbol of community, a concrete representation of mutual support and concern … as David Peterson, the Premier of Ontario, pointed out … 'There is no social program that we have that more defines Canadianism or is more important to the people of our country'.

The overwhelming support for Australia's Medicare program reported in the Health Insurance Commission's Annual Reports and elsewhere indicates that a similar statement could probably be made in Australia. When the willingness to pay for social programs reflects such a value system it is likely to diverge significantly from the private willingness to pay.

Proposition 3: The willingness to pay for your life is an unhelpful concept in the context of a National Health Scheme.

This and the next proposition reinforce the theme that the individual willingness to pay for one's own life is an inappropriate concept for the valuation of health services. Proposition 3 reiterates the unoriginal theme that while the WTP for the risk free benefits of breakfast cereal, theatre, travel, etc is a useful measure of the benefits obtained from these products there is no corresponding useful concept of the willingness to pay for the certainty of saving one's own life. The value of this to an individual would normally be their entire income and wealth less the minimum required for biological survival with, perhaps, some consideration of the inheritance to be received by heirs. In a purely private market individuals pay insurance premiums which give access to a range of services which may or may not save their life at some later date. That is, there is a WTP for a reduction in the risk of death and not for the certainty of life. In this respect,

once again, WTP based CBA differs fundamentally from CEA and CUA where the unit of benefit is not calculated with a knowledge of the risk of death and is concerned with realised (ex post) health outcome and not with the subjective benefits of (ex ante) risk reduction[6].

The two notions of ex ante risk reduction and the value of ex post health outcome have often been linked using the Von Neumann Morgenstern (NM) axioms and it is argued that the value of saving a life can be inferred from the value of reducing the risk of death. For example, it is argued that a willingness to pay of $2,000 for a reduction in the probability of death of one in a thousand implies a willingness to pay of $2,000 x 1,000 = $2 million for a reduction of a likelihood of death of 100 percent. The calculation depends upon the truth of the NM axiom that choice under risk is unchanged by linear transformation and that, consequently, multiplying two sides of an equation by a factor of 1,000 does not change the choice. While perhaps appealing, the axiom is flawed by the fact that it is empirically false and self evidently so when the implied willingness to pay exceeds a person's resources. More generally, people do not maintain choice under linear transformation and – as explicitly recognised by Von Newman and Morgenstern – risk behaviour is heavily influenced by the specific disutility of risk (hope anticipation, anxiety, the dislike of risk per se etc) and this is not taken into account in the NM axioms (Richardson 1994).

An alternative argument is that, as the individuals who pay for private or national health insurance face only the risk of ill health and death, then it is appropriate that the WTP should be for risk reduction and not for the certainty of outcome. Once it is acknowledged that these alternative end points are not equivalent, then this argument amounts to an acceptance of the fact that WTP based CBA is a conceptually different form of analysis from CEA and CUA where the risk free (quality adjusted) life year is the unit of output. The question of which is the appropriate basis for decision making is not, however, an issue of economic theory but an ethical issue to be resolved by social decision making. Economic theory per se cannot be of assistance – there is no immutable law of society which asserts that utility, as defined by economists, is the only possible social objective - and the only help which may be offered by economists with respect to this issue is to empirically investigate what it is that decision makers and the community wish to have maximised. The question cannot at present be unambiguously answered because, to date, economists have not asked the question or carried out the necessary research. Prima facie, however, the evidence from public policy strongly favours the output units used in CEA and CUA. Even a superficial review of policy statements and the object of national statistical collections would indicate that social concern is with the incidence of disease and the level of mortality and not with minimising the disutility arising from people's risk aversion. Of course it might be argued that the best way to minimise risk aversion is to minimise morbidity and mortality. Once again, it is an empirical issue whether or not this is the motivation behind health programs. However, evidence cited in relation to Proposition 4 suggests that the population is not particularly concerned with individual preferences or with risk aversion, but are concerned with health per se.

6 Some have argued that it is for this reason that utilities should be calculated using the standard gamble. The logic of this argument is, however, incorrect. The standard gamble employs a probability of death to derive an index of the quality of life of a health state. This is not, except by coincidence, be equivalent to the risk of death associated with a health state. It does not give any indication of the absolute value of life but only the relative value of life in two different health states.

Proposition 4a: The hypothesis that people want a national health service to provide benefits that would be selected by private individuals for themselves is wrong.

Proposition 4b: If consulted, people would reject the personal willingness to pay as an appropriate criterion.

Proposition 4c: People are prepared to override other people's preferences for services in order to achieve the objectives that they believe appropriate for a national health service.

Two small and one large empirical study have been carried out by the author to gain an indication of the objectives of the general population in this context. In the first, a mail questionnaire was sent to a random sample of the Australian public selected from the telephone directory. Three questions were asked. In each, respondents were requested to indicate which of two programs, A or B, should be included in Medicare. Respondents were told that the outcomes from the two programs were the same but that program A was preferred by the majority of the population who were willing to pay the increased cost of the program. The reason for this varied. In version 1, program A provided greater personal contact and reassurance. In version 2, program A was presented in such a way that they could experience regret and, to avoid this, they were prepared to pay the extra cost of program A. In the third case, program A involved risk and people were prepared to pay the extra for program A to avoid this risk.

In each of the three cases the arguments for A and B were summarised and, importantly this included the statement that respondents were told that *'A should be preferred because this is what people with the disease want and would be prepared to pay for. Whenever possible Medicare should give people the services that they prefer*'.

In sum, the three options involved the choice between a program (B) where the only benefit was the ex post improvement in health and a more expensive program (A) which people preferred and were willing to pay for because there was an additional ex ante benefit associated with the process of the program which patients were prepared to pay for. As shown in Table 1 the 85 respondents unambiguously rejected program A, thereby rejecting the argument that people should be given the services for which they are willing to pay. By implication, the implied criterion for the inclusion of services in the National Health Scheme was the cost and ex post impact upon health status.

Table 1 Choice of Criterion: WTP vs Health Outcome

'People are willing to pay extra for Program A because of: '				
Program	(i) Reassurance	(ii) Regret	(iii) Risk	(iv) Total
A WTP Criterion	31%	28%	37%	32%
B Health Benefit Criterion	69%	72%	63%	68%

In the second study Proposition 4(b) was tested using a convenience sample of 74 MPH students enrolled at The University of Melbourne. The willingness to pay criteria was explained to the group and its use for the allocation of resources in the general economy was similarly explained. The students were then told that 'it has been argued that health services should be valued in the same way as other market products; that is, that the value of each treatment for each person should be set equal to the person's willingness to pay'. Students were then asked whether they would support the government in allocating resources in the health sector and including them in Medicare on the basis of this criterion of value.

As shown in Table 2, in response to two related questions, 84 percent indicated that this proposal was 'ethically objectionable' and 97 percent (all except 2 students) indicated they would not support the government in this proposal.

These results are unsurprising as national health schemes in every western country (including the national Medicaid and Medicare schemes of the USA) were established to remove this criterion as the basis for allocating resources.

Table 2 Willingness to Pay for Health Services

Is WTP for health services?		Health	
(i)	Ethically objectionable	Yes	84%
(ii)	Should resources be allocated this way?	No	97%

Proposition 4c has been tested more systematically in a collaborative survey in Australia and Norway (Olsen et al 1998). In both countries a stratified population sample was sent a mail questionnaire which, inter alia, asked respondents to choose between two programs; namely, a screening program which would save the larger number of lives and a helicopter ambulance service which was preferred by the public. Importantly, respondents were asked to adopt the perspective of a decision maker and to vote as they would wish their decision makers to vote. It was emphasised that their constituency preferred the ambulance service. In three versions of the questionnaire the reason for the public's choice of the helicopter service was changed so that it emphasised the importance of public preferences (version 1); equity (version 3); and the 'rule of rescue', viz that in an emergency everything possible must be done. Results in Table 3 are striking. In none of the three versions and in neither country was the helicopter service preferred. That is, respondents rejected the principle that public preferences should be respected even when this was reinforced by issues of social justice. Without these reinforcing factors only one quarter of respondents accepted the sovereignty of the public/consumer and even with the inclusion of the social justice issues the proportion only rose to a little over one third.

Table 3 Paternalism vs Respect for Public Preferences

	Percent Selecting Different Options			
	Questionnaire Emphasised			
	Public Preference %	Rule of Rescue %	Equity %	Total %
Public Preference (Helicopter)				
Australia	28.1	41.4	32.2	33.8
Norway	24.2	42.6	45.0	37.6
Maximum Lives (Screening)				
Australia	61.7	48.3	54.3	54.8
Norway	63.0	42.2	42.0	48.7

Source: Olsen J, Richardson J & Mortimer D (1998).

The final evidence relevant to this issue arises from the enquiry into population objectives undertaken by Dolan and Cookson (1998). In this, a series of two hour focus groups were conducted with ten groups of six people on each of two occasions. After initial explanation of the purpose of the focus groups each group was asked how they would set priorities in the UK National Health Scheme; that is, what principles they would apply. The outcome of this exercise was a list of principles and the frequency with which these were mentioned by individuals and by groups. It was possible to classify the results under a number of general principles; viz, that priority should be given to (i) procedures providing greater health gain; (ii) patients with life threatening and urgent conditions; (iii) procedures favouring the young; (iv) procedures where there was not a self inflicted illness; (v) procedures which met some threshold requirement relating to either health improvement or final health state; (vi) the length of time queuing. Conspicuous in this list is the complete absence of any principle which could be interpreted as relating to the economist's notion of utility. The criteria were entirely couched in terms of different dimensions of health or health related personal characteristics.

In combination with the results from the three surveys above this empirical evidence strongly suggests that the ethical basis of the willingness to pay criterion, viz that individual preferences per sé should be respected is, in fact, firmly rejected in the context of a national health scheme.

Proposition 5a: The argument that social decision making is imperfect and that we must therefore accept individual preferences and the WTP as the criterion of value is wrong. Proposition 5b: Social decision making is (conceptually) easy.

These propositions have been included here in anticipation of one possible rejoinder to the present line of argument, viz that Arrow's Impossibility Theorem demonstrates that rational social decision making is not possible[7].

[7] Arrow considered a situation in which three people had the preference ratings such as those below:
Person 1: A > B > C Person 2: B > C > A Person 3: C > A > B
With simple voting we would find a majority (two persons) would favour the following:
Vote A versus B → B; Vote B versus C → C. This implies that A should be preferred to C. However, Vote A versus C → C.

There are two objections to this argument. The first, and more compelling objection, is that the logic is wrong. Even if it was agreed that simple voting was a flawed method for collective decision making it would not follow that the libertarian alternative of rejecting all collective decision making should be the default option. Depending upon the criterion adopted in a second best world and the issue over which decision making was necessary, a flawed collective system may still prove to be the second best option.

Secondly, and more fundamentally, Arrow's Impossibility Theorem does not flaw representative government and its capacity to decide issues such as the criterion for program evaluation in the health sector. At best, the theorem demonstrated that it is, in theory, possible for voting to be inconsistent. To the author's knowledge there are no examples in any country where such inconsistency has, in practice, impaired decision making. At worst the impossibility theorem is an intellectual game in which the rules are: (i) that decision making must be 'rational'; (ii) that rationality is defined as consistency; and (iii) that decision making is by majority vote with sequential voting between options. The rules of this game do not correspond with the rules of government. First, the limited concept of rationality embodied in welfarism does not correspond with the notion of rationality more generally. Alternative criteria of rationality would allow a variety of devices to avoid problems associated with the paradox (eg including the strength of preference; simultaneous voting on all 3 options etc). Secondly, decision making on particular issues is seldom if ever made by a series of sequential referenda. Thirdly, in the case of inconsistency, a dispute resolution mechanism exists, viz the referral of the issue to parliament.
Proposition 5b is the trivial observation that social decision making is (conceptually) easy. It may be observed daily in every social collective: at the local state, and national level by local state and national governments.

The popularity of the Impossibility Theorem does not arise from its demonstration that decision making is impossible but that rational decision making is impossible in the abstract model described above. The probable rejoinder to the apparently casual dismissal above, of a widely acclaimed contribution to economic theory, is likely to be that the theorem demonstrates that simple voting procedures cannot maximise social welfare in a social welfare function. Depending upon the definition of the social welfare function this may or may not be true. The relevant issue in the present context, however, is that it does not demonstrate the impossibility of either establishing the criteria for decision making or of maximising a proximate objective, viz life years or QALYs.

5 Qualifications

Proposition 6: Despite the argument for the adoption of the social willingness to pay criterion, individual willingness to pay for health related services will almost certainly be accepted to a greater or lesser extent.

This will occur because any national health scheme must, inevitably, draw boundaries between services that are included and excluded. In liberal democratic countries it is unlikely that government will prohibit the private purchase of the services that have been excluded. There are two major categories of these. The first are services which, after (casual or careful) evaluation are omitted because their benefits or the probability of favourable benefits are too small. Secondly, there are services which are excluded because their total cost is too great. Of course it is possible that these two criteria could be simultaneously applied.

The result of this qualification is that, from a national perspective, resources must be allocated according to different rules in different contexts, viz social WTP for services in the public health system and individual WTP for services outside it. As the boundary between these sets of services is likely to shift this implies the possibly intellectually unsatisfying,conclusion that, from a contextless perspective, the social value of services could fluctuate. The conclusion is unsurprising if it is accepted that 'value' implies a value system and that values are mutable. In principle this is no more surprising than the value of warm clothes varying with the climate.

Proposition 7: Social WTP is a function of outcome, process and context.

Propositions 1 to 5 imply that the social WTP will be a function of outcome only: mortality, morbidity and HRQoL. While these outcome variables are undoubtedly of importance and may remain the primary benefit incorporated in CEA and CUA it is likely that the optimal criteria for resource allocation will include process and contextual factors. There is clear evidence that process issues relating to equity, access and the systemic determinants of social justice are a major concern to the population. While perhaps not the primary criterion, individual's preferences are clearly not irrelevant even in the context of an NHS and, particularly if, as in the case of Australia and the UK, the NHS permits individuals to purchase 'rights' to priority or supplementary services through private health insurance.

Additionally, there is preliminary evidence that the importance of services and of the social WTP depends upon the context in which a decision is be made. When budgets are being dispassionately allocated on the basis of outcome, the opportunity cost of services and health outcome may be the primary concern. Dispassionate decision making is, however, rejected in the context of an emergency where it is expected that all available resources should be used irrespective of cost (the so called 'rule of rescue'). Similarly, 'rational' resource allocation may be rejected at the 'bedside' where the provider of care is expected to have only the patient's interest in mind.

The extent to which process and contextual factors impinge upon CEA and CUA is largely unknown because it is only recently that their importance has been recognised and the empirical

research required to answer these questions has not been conducted. If these variables do prove to be important then, once again, from the contextless perspective of consequentialist economic orthodoxy, optimal context specific decision criteria may appear to be incoherent and inconsistent.

Proposition 8: In practice the application of social willingness to pay techniques will probably be limited to the calculation of the value of a quality adjusted life year/DALY with this latter metric being used to evaluate the benefits from health program interventions.

The methods for eliciting a social willingness to pay have not been well developed and, given the unfamiliarity of the form of question and the need for deliberative processes the task is likely to be problematical. Additionally, and following from the empirical evidence cited here, people are likely to resist the dollar quantification of health states. Consequently it is likely that results from the use of QALY/DALY procedures will be more reliable, valid and cheaply elicited and that, therefore, these techniques should continue to be the primary metric for health service evaluation.

Proposition 9: Despite the previous propositions there is a case for the measurement of individual WTP.

There are three reasons for this final qualification. First, in the context of a national health scheme, economists might be expected to adopt the criteria nominated by government decision makers. (This is a position forcefully argued by Sugden and Williams [1978]). Irrespective of public opinion, governments may nominate welfarist criteria. Even if this is not presently done, future governments may do so (and possibly because of the advocacy of economists!). Secondly, many of the decisive arguments in this debate are empirical. It has been suggested here that the evidence does not support welfarism and that the public would prefer the adoption of a social WTP criterion based largely upon non welfarist objectives. However, this empirical evidence is incomplete and may be wrong.

Thirdly, and perhaps most basically, measurement of all but the simplest outcomes is imprecise and even conceptually problematic. With their tradition of full information, rationality and revealed preferences, economists generally assume that when questions are asked and information sought, questions are perfectly understood, stable attitudes for preferences exist and the answers provided mean precisely what the grammar of the reply implies. In the psychometric tradition almost the opposite occurs and it is assumed that replies received may have little relationship to their literal meaning or to the precise question and that their meaning must be learned by a series of 'validation' studies, ie there must be a series of studies to determine what stated preferences can be used to predict. From this perspective individual WTP could be viewed as simply another metric which may be used to increase confidence if it, as well as other metrics, indicate the same conclusion from an economic evaluation.

In sum, individual WTP could be viewed as a form of sensitivity analysis, firstly, testing the validity of other psychometric scales of population preferences and, secondly, increasingly confidence that between the different metrics we have measured the imperfectly understood social objectives which economic evaluation should seek to satisfy.

6 Conclusion

The justification for using the willingness to pay criterion of value in economic appraisals of health programs has received only the most superficial attention. A series of conventional assumptions have been made about social objectives which do not stand up to criticism. Economic evaluation should not impose a set of conventional and analytically convenient values upon the community. Rather, it should reflect the social values of the community. These may not be fully coherent and consistent and they may be different in different countries. These are questions for empirical research which has, to date, been largely neglected because there has been little recognition of the complex interplay of social values and social context and the implications of this for economic evaluations. These issues should be of priority concern to economists or their analysis runs the risk of either misleading the community or of being ignored.

References

Culyer AJ. 1991, 'The normative economics of health care finance and provision', in *Providing Health Care: The Economics of Alternative Systems of Finance and Delivery*, (eds) A McGuire, P Fenn & O Mayhew, Oxford Economic Press.

Dolan P & Cookson R 1998, Fairness in health care: What the public thinks, Paper presented to the HESG Meeting, Galway, July.

Donaldson C, Map T, Ryan M & Curtin K 1996, 'Estimating the economic benefits of avoiding food borne risk: Is willingness-to-pay feasible?', *Epidemiology & Infection*, vol 116, pp 285-294.

Donaldson C, Map T et al 1997, 'Assessing community values in health care: Is the willingness-to-pay method feasible', *Health Care Analysis*, vol 5, pp 7-29.

Evans RG & Law NM 1995, 'The Canadian Health Care System: Where are we and how did we get here?' in *An International Assessment of Health Care Financing, (eds)* D Dunlop & J Martens, The World Bank, Washington DC.

Graaf J, de V 1967, *Theoretical Welfare Economics*, Cambridge University Press, Cambridge.

Johansson PO 1995, *Evaluating Health Risk and Economic Approach*, Cambridge University Press, Cambridge.

Little IMD 1965, *A Critique of Welfare Economics*, Oxford University Press, London.

Menzel P, Gold M, Nord E, Pinto Prades J, Richardson J & Ubel P 1999, 'Towards a broader view of values in cost effectiveness analysis of health care', Hastings Centre Report, May/June, pp 2-10.

Mishan E 1971, 'Evaluation of life and limb: A theoretical approach', *Journal of Political Economy*, vol 79, no 4, pp 687-705.

National Oceanic and Atmospheric Administration, 1993, Report of the NOAA panel on contingent valuation, Federal Register, vol 58, pp 4601-4614.

Nord E, Pinto Prades J, Richardson J, Menzel P & Ubel P 1999, 'Incorporating societal concerns for fairness in numerical valuations of health programs', *Health Economics*, vol 8, pp 25-39.

O'Brien B, Stoddart G, Donaldson C, Torrance G & Drummond M, 1998, Willingness to pay for what? Alternative definitions of health care program benefits for contingent valuation studies, Working Paper 89/07, Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario, Canada..

O'Brien B, Goere, ER, et al 1998, 'Assessing the value of a new pharmaceutical: A feasibility study of contingent valuation in managed care', *Medical Care*, vol 36, pp 370-384.

O'Brien B & Gafni A 1996, 'When do the 'dollars' make sense: Towards a conceptual framework for contingent valuation studies in health care', *Medical Decision Making*, vol 16, pp 288-299.

Olsen JA, Richardson J & Mortimer D 1998, Priority setting in the Public Health Service: Results of an Australian survey, Technical Paper 9, Centre for Health Program Evaluation, Melbourne, Australia.

Olsen JA, & Donaldson C 1998, 'Helicopters, hearts and hips: Using willingness-to-pay to set priorities for public sector health care programs', *Social Science and Medicine*, vol 48, pp 1-12.

Richardson J 1994, 'Cost utility analysis: What should be measured?', *Social Science and Medicine*, vol 39, no 1, pp 7-21.

Richardson J, Hall J & Salkeld G 1996, 'The measurement of utility in multi-phase health states', *International Journal of Technology Assessment*, vol 12, no 1, pp 151-162.

Robbins L 1938, 'Interpersonal comparisons of utility', *Economic Journal*, vol 48, pp 635-641.

Sen A 1982, *Choice Welfare and Measurement*, Harvard University Press, USA.

Sugden R & Williams A 1978, *The Principles and Practice of Cost Benefit Analysis*, Oxford University Press.