How Should We Decide Which Ethical Preferences to Include in The Economic Analyses of The Health Sector

Paper presented to the Health Services Research Association of Australia and New Zealand Conference, 2-4 December 2001

Professor Jeff Richardson
Director, Health Economics Unit, Monash University

April, 2002
ISSN 1325 0663
ISBN 1 876662 49 2
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 of these disciplines within 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 Public Health at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

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 4434  
Facsimile + 61 3 9496 4424  
E-mail CHPE@BusEco.monash.edu.au  
Or by downloading from our website  
Web Address http://chpe.buseco.monash.edu.au
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

It is argued that orthodox economics has failed to acknowledge the extent to which ethical judgments permeate evaluation and policy advice. It has likewise failed to make explicit and debate the prior issue of how we should approach the task of evaluating ethical theories – that is, the ethical assumptions which underpin policy advice. Three such approaches are discussed. These are: (i) the approach implied in economic theory (‘Platonic orthodoxy’); (ii) the approval adopted by the ethics profession (‘Ethics’) and (iii) an approach labelled ‘Empirical Ethics’. It is concluded that the latter is the most secure and rewarding basis for theory and policy. The argument draws upon several empirical studies which are summarised.
How Should We Decide Which Ethical Preferences to Include in The Economic Analyses of The Health Sector

Paper presented to the Health Services Research Association of Australia and New Zealand Conference, 2–4 December 2001

1 Introduction

There are three separate types of questions which we might ask in relation to economic evaluation in the health sector or elsewhere. First, what interventions meet a particular criterion? (For example, benefits should exceed costs or net costs should be less than a given threshold.) The methodology to answer this (positive) question is (social) scientific: the answer depends upon empirical evidence. Secondly, we may ask which considerations we wish to have included in the evaluative criteria (utility or life years, discounting of the future; particular costs, equity defined in a particular way, etc). Answers to this second question require normative—ethical—analysis. Finally there is the question of how we decide what it is that we should include in our criteria. This issue is not directly ethical but about the derivation of ethical ‘principles’; that is, it is about the epistemology of ethics. The present paper is about these last two questions.

The paper proceeds as follows. First it is argued that ethical questions have been buried by economists and that, contrary to the impression commonly conveyed by text books, ethical issues permeate applied economics and, most obviously, health economics. Secondly, a number of examples are given of ethical issues that are unresolved and, generally, unrecognised as needing resolution or even discussion. In Section 3, three approaches to the epistemological question are considered. These are that we should accept the orthodox—‘proven’—assumptions about the nature of social values; that we should adopt the analytical framework of the discipline of ethics; and, finally, that we should adopt the approach labelled here ‘Empirical Ethics’. In Section 4 some of the likely criticisms of Empirical Ethics are considered and in Section 5 it is concluded that, while the present proposal for empirical ethics may be embryonic, alternative options for answering the epistemological question have not, to date, been suggested and that the current bases of the discipline’s epistemology—assumption and historical inertia—are unsatisfactory. Ethical issues have been ignored to the detriment of the discipline and the suggestions made in this paper represent a defensible foundation upon which to reconstruct this important and neglected area of economic analysis.

---

1 This paper draws together and extends themes from 5 recent studies in order of importance:

1 Richardson J. Empirical Ethics, or: The poverty of ethical analyses in economics and the unwarranted disregard of evidence in ethics. Working Paper 120, Centre for Health Program Evaluation, Monash University, Melbourne, 2001. Also in:

2 Richardson J. Why economic costs may not be of interest in a national health scheme, or: Costs, fairness and reverse order analysis. Working Paper 126, Centre for Health Program Evaluation, Monash University, Melbourne, 2001.

3 Richardson J. Age weighting and discounting: What are the ethical issues? Working Paper 109, Centre for Health Program Evaluation, Monash University, Melbourne. 2000. Also in:


2 Ethical Judgements Permeate Economics

The recommendation of a policy—the assertion that one state of the world is better than another—is always and unavoidably based upon an ethical judgement or belief. The common view amongst economists that issues of efficiency are different in kind as they represent a value free basis for economic policy is wrong. In all its various forms efficiency results in the achievement of the same outcome with the use of fewer resources. But this is only desirable if the surplus resources are used for some activity which is desirable: increased production of the same output; increased production of other useful commodities or increased leisure. There is, however, no objective criterion to determine what is or is not desirable. A judgement must be made and this judgement is unavoidably a normative—ethic—issue. What distinguishes efficiency from more overtly ethical questions is that it is relatively easy to define and that there is universal agreement about its desirability. Other cases are less clear cut. Two examples are given below.

Example 1: Welfarism or Extra Welfarism

Orthodox economic theory assumes that individuals seek to maximise utility and that this will occur if they are given choice and information about the consequences of their choices. The counter view, commonly expressed in health economics, is that the purpose of the health system is to generate health, not choice.

Evidence on the strength of the social preference for choice and for health was obtained by Olsen and Richardson [1] in parallel surveys in Australia and Norway. In these, respondents were asked to choose one of two health programs. The first—Program A—maximised lives saved by screening a segment of the population. The second—Program B—saved fewer live but was unambiguously preferred by the population. Three sets of reasons were given for this. First, the helicopter improved equity and fairness by increasing the access to health services by populations in remote areas. Secondly, it provided a better service in cases of an emergency. In the third scenario respondents were simply told that the helicopter was the preferred choice of the population. The questionnaire was designed to test the social preference for these services by a representative group of citizens who were aware of individual preferences. Results, summarised in Table 1, unambiguously indicate that, faced with a choice between health maximisation and the endorsement of population choice (which is most likely to maximise utility)\(^2\) survey respondents rejected the choice criterion and opted for the health maximisation. Respondents were relatively unconcerned about the satisfaction of people’s private preferences.

\(^2\) Maximising choice would, by definition, maximise utility if every person had the same intensity of preference. Utility would not be maximised if the policy selected was supported by those with a low intensity of preference and opposed by a smaller number but people with intense preferences.
Table 1  Results and program (percent)

<table>
<thead>
<tr>
<th></th>
<th>Choice %</th>
<th>Emergency (V1) %</th>
<th>Equity Fairness (V2) %</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aust</td>
<td>Norway</td>
<td>Aust</td>
<td>Norway</td>
</tr>
<tr>
<td>Program A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening: (Max Lives)</td>
<td>62</td>
<td>63</td>
<td>48</td>
<td>42</td>
</tr>
<tr>
<td>Program B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helicopter</td>
<td>28</td>
<td>24</td>
<td>42</td>
<td>43</td>
</tr>
<tr>
<td>Unsure</td>
<td>10</td>
<td>13</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>n</td>
<td>334</td>
<td>227</td>
<td>321</td>
<td>258</td>
</tr>
</tbody>
</table>

Numerous empirical studies of this sort have been carried out in the past decade. These are summarised in Menzel et al [2]; Nord et al [3]; Nord [4] and Ubel [5]. These find prima facie evidence for a range of social—ethical—preferences, including a social preference with respect to:

- severity of a health state per se
- some personal characteristics, including age
- the perspective of the decision maker: individual or societal; ex ante or ex post
- certain contextual factors. The so called ‘Rule of Rescue’ has been cited to explain the social demand for highly costly, ineffective but dramatic and public interventions
- communitarian values; that is, the maintenance of ‘solidarity’ or ‘community capital’
- the maintenance of hope
- achievement of a ‘fair innings’.

**Example 2: Positive Net Benefits**

‘Discrimination’ on the basis of treatment costs is the most fundamental principle of economic evaluation but it may also be perceived as unfair. This is illustrated in Table 2 below. In this, the social benefit of treating each of five diseases A-E is 40. However the cost of the health services needed to cure them rises from 10 (Disease A) to 41 (Disease E). Following economic orthodoxy benefits exceed costs for the first four health states and, consequently, these should be covered by health insurance and included in a national health scheme.
Table 2  Costs, benefits and fairness: 5 diseases

<table>
<thead>
<tr>
<th>Disease</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment benefits</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Treatment costs</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>41</td>
</tr>
<tr>
<td>Optimal rebate = subsidy</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Patient cost</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>41</td>
</tr>
</tbody>
</table>

From the patient’s perspective an increasingly costly illness results in increasing social generosity until disease E is reached at which point social generosity does not taper off but terminates abruptly. For the patient who has no interest in global utility maximisation such an abrupt response to such a marginal change in service cost (from 40 to 41) will almost certainly appear to be unfair (even if the reason is explained). For many, a ‘fairer’ outcome might be to provide benefits to all categories of patients but with the benefit a function of various considerations only one of which is the disease cost. The ‘fair’ functional relationship would almost certainly not have an abrupt termination of benefits. This, in turn, implies that there is no special significance in the point at which marginal costs equal marginal benefits. This point identifies maximum social net benefit. But if this ceases to be the principal social objective then there is little reason to bestow special status upon this point. In Richardson [6] I suggest that ‘reverse order analysis’ should be undertaken which commences with fairness and then considers a trade-off between equity and fairness. It is suggested that the principle social concern may relate to financial flows between groups, not ‘costs’ as envisaged in economic theory.

3   How Do We Answer Ethical Questions

This latter conclusion raises the epistemological question of how we are to determine social objectives. Three options are considered below. These are described as (i) Platonic Orthodoxy (PO). This is were we make the assumptions of the orthodox welfare model. A near cousin to this orthodoxy is to assume ‘healthism’: that the objective of the health system is to maximise health; (ii) orthodox ethical analysis. This is defined by the type of analysis conducted within the discipline of ethics; (iii) empirical ethics, which is the central suggestion of this paper [7].

Platonic Orthodoxy: The assumptions which define the welfare model and economic orthodoxy are relabelled here to highlight the similarity between the epistemology of economic orthodoxy and the epistemology of, inter alia, the Platonic theory of Forms, or , more generally, the epistemology of philosophical rationalism [8]. The defining characteristic of this is that a set of fundamental assumptions or axioms are adopted which have been derived from intellectual intuition or (in earlier rationalism) divine revelation. It is evidence free. Subsequent ‘truths’ are derived by the application of logic or mathematics. The methodology is adopted in the model of welfare economics which translates various axioms into recommendations for the maximisation of social welfare.

There are two broad sets of problems associated with this methodology. The first includes the fact that some of the assumptions may be wrong. Orthodox economists, for example, considers one state to be superior to another if ‘winners’ in this state could compensate losers (the Kaldor–
Hicks potential compensation principle). In the health sector, however, compensation cannot be paid, even in principle, to those who have lost their life. It is likewise (politically) impossible to tax the beneficiaries in a national health scheme and compensate the losers (tax payers) without losing the purpose of the health scheme. More generally, distribution and net benefits cannot be separated in the health sector; and there is no strong evidence to suggest that the objectives of the population correspond with those postulated in Platonic orthodoxy.

The second set of problems are associated with the methodology of Rationalism [8,9]. Assumptions or axioms are always potentially fallible and, more specifically, their truth or otherwise is likely to be context specific. The ultimate authority for assumptions/axioms in Rationalism is their innate plausibility: but the intuition of the analyst is a highly fallible basis for a social science. The analyst may insist that their intuition is fashioned by empirical evidence but this simply suggests the need for empirical evidence and not intuition.

**Option 2: Ethics:** There is considerable appeal in the proposition that ethical issues should be in the domain of the ethics discipline. But ethics, per se, as practised by formal ethicists, cannot demonstrate the superiority of particular ethical assumptions. To do this pre-supposes the existence of an evaluative criterion, A, which can be used to demonstrate the superiority of an assumption. But there is a prior question whether or not this evaluative criterion is acceptable and to determine this there is a need for a further criterion to evaluate A, viz, criterion B. However to justify B requires criterion C which requires criterion D... Justificationism of this form encounters an infinite regress. That is, logic alone cannot demonstrate the superiority of an ethical system.

Ethical analysis can identify inconsistencies and weaknesses in arguments and can suggest (not prove) the desirability of extending the principle from one context to another. It cannot, however, provide a secure ethical basis for evaluation and policy and, for similar reasons, it cannot quantify the strengths and weaknesses of different ethical theories.

**Option 3: Empirical Ethics:** In the last ten years there have been numerous empirical enquiries into social preferences; that is, into the stated preference for an alternative distribution of benefits and costs or for an alternative procedure for evaluating the desirability of health related interventions. The suggestion here is that the procedures adopted in these studies be described as ‘Empirical Ethics’ and that a set of procedures be formalised in such a way that the adoption of the procedures be recognised as an acceptable methodology for determining ethical preferences and for resolving ethical issues. The suggestion here for such a set of procedures is follows:

a) **An iterative elicitation of values**

   (i) Hypothesis generation and clarification should be carried out using qualitative survey techniques. Alternatively any other methodology for generating a hypothesis should be regarded as acceptable. This includes the relatively large number of investigator generated hypotheses underlying the empirical inquiries reported in the literature. While qualitative surveys are one approach to hypothesis generating, the process should be of less interest than the subsequent testing and validation of the hypothesis.

   (ii) Deliberation and clarification of responses. As part of the process of elicitation, a variety of techniques should be used to achieve deliberation and a precise statement and specification of values. In this and the next step below a variety of techniques
should be used to encourage deliberation. These include interview-reinterview; different forms of information feedback (including Delphi techniques), the triangulation of issues, ie eliciting issues using different techniques and a comparison of results; focus group and other discussion techniques.

b) Quantification of social (value) preferences. The chief technique used in the literature to date has been the person trade-off. Other decision analytic trade-off procedures are possible including Stated Preference Discrete Choice Modelling (SPDCM) and willingness to pay procedures. (The latter should not—necessarily—be treated as having particular ethical significance. It may be used simply as an instrument for eliciting social preferences which may or may not then translate into a corresponding number of dollars expenditure in a health scheme.) As noted above, it is important that, to the maximum extent possible, responses are elicited using validated deliberative procedures.

c) Critique and testing. At all points, but particularly following initial quantification, results should be subject to ethical critique. That is, the logical consequences of results should be derived and provided to subjects to assist their deliberation. For example, a utility weight of 0.75, (as this is usually interpreted), implies that returning four people from such a health state to full health is equivalent to saving the life of a single person and returning him or her to full health. Respondents providing a utility weight of 0.75 might be unwilling to accept this inescapable consequence and opt to revise their initial estimate of the utility score. Similarly, an ethical critique will highlight apparent incongruities in results when they are transferred to another context. This may, again, suggest the need for revision. Alternatively, it may suggest that the contextual factor has ethical significance. Unsuccessful testing should lead to a further iteration of the earlier steps. Successful testing and retesting should qualify the hypothesis for tentative acceptance.

In sum, the acceptance criteria advocated is that there is majority support for an ethical issue or that the social value assessed on a suitable instrument reach a pre-specified threshold. Results should be capable of withstanding a robust but reasonable ethical critique. It would be desirable to formalise a convention (albeit stated in qualitative terms) which, once satisfied, would justify the ‘tentative adoption’ of an ethical position. This does not indicate ‘truth’ or ‘certainty’; rather the convention would be analogous to the 5 percent convention used in statistics for the tentative rejection of the null hypothesis. In this discipline it has been accepted that certainty cannot be achieved and that it is necessary to replace this beguiling but unobtainable goal with an explicit and pragmatic criterion.

It should be emphasised that empirical ethics, like statistics, does not—cannot—lead to certainty and it would be unhelpful to subject it to the criticism that certainty cannot be demonstrated. Arguments of the form that ‘it is conceivable that …’ should be considered illegitimate. Utilitarianism, in principle, permits slavery, sadism, etc. But this is uninteresting in a context where such options do not exist for easily defensible ethical reasons. In contrast, arguments of the form ‘in normal circumstance this theory implies…’ are legitimate. For example, the simple interpretation of utility scores implies lower priority for disabled persons who can never achieve full health. This implies either lower priority care for the disabled—a potentially fatal flaw in utility measurement—or the existence of an ethically relevant contextual factor, possibly arising from a conflicting ethical principle. In this particular example—double jeopardy—the latter explanation appears to be correct.
4 Evaluation of Empirical Ethics

The most fundamental objection to Empirical Ethics which is likely to be raised by ethicists may be summarised in the following way.

‘It would be foolish to believe that ethical issues can be resolved simply by voting. Even with sophisticated procedures designed to elicit deliberative responses, people might select immoral options. This is recognised by the need for the ‘laundering’ of preferences; but laundering is simply a recognition of the fact that populations may be wrong in their ethical perceptions. They may vote for slavery, racist or other immoral policies. Ethics is immensely complex and it is no more sensible to elicit opinions from even well informed and intelligent amateurs than to vote on the laws of physics.’

There are two implicit assumptions in this argument and both are unacceptable. First is the belief that ethical arguments must be fully consistent and applicable in all possible contexts. The ‘immense complexity’ of ethics often arises only when ethicists seek the implications of an ethical theory in an unusual or abhorrent context. Thus, and as noted above, simple utilitarianism may be criticised as it appears to permit torture if this generates more utility than it destroys. However, utilitarianism may have broad application in uncontroversial areas. It is almost certainly a correct and defensible description of population values in markets less contentious than health care. More generally, and as also noted earlier, there is ethical content in all normative advice and the ethical basis for much of this is simple and uncontroversial. There is no great complexity in the ethical principles which lead to the recommendation of technical efficiency. There is no ethical complexity in accepting that severity per se may be of independent importance in prioritising health services or that we may wish to give priority care to individuals who have previously been disadvantaged, have not yet had their ‘fair innings’ (life span), have been discriminated against, etc, etc. A significant part of Empirical Ethics has been, and should continue to be, the identification of such issues and the demonstration of their importance relative to more generally accepted goals.

The second implicit assumption in the passage is that an argument or procedure is flawed unless it achieves certainty and universality: that a procedure or ethical theory must be rejected if it might conceivably produce a ‘wrong’ conclusion—one that is repugnant to our moral intuitions. However, ethical theory and normative economic analyses do not and cannot provide answers to ethical questions which are unambiguously true or immune to criticism. But this conclusion also applies to positive analyses. It is always possible that the current laws of physics, for example, may be falsified by new observations. A physical law is never unambiguously ‘true’, but represents the best current theory or hypothesis which may subsequently be corrected, generalised or rejected altogether.

It may, of course, be argued that the ideas that are ultimately successful will be those that have been most skilfully marketed and this does not represent gold standard methodology: these ideas may be ‘wrong’ in some context and the procedure suggested here is, therefore, flawed. However this rejoinder once again assumes that there must be certainty—that the theory must be universal and true in every context, no matter how contrived.

A softer ‘interpretation’ of the ‘successful marketing objection’ is that, even if it is recognised that certainty is ephemeral, it is unacceptable to have the rhetorical strength of an argument as the deciding criterion between the competing theories. While this version of the critique has some
rhetorical power it is not persuasive unless an alternative, superior, methodology is proposed which avoids this objection. Acceptance and rejection of a theory is never permanent. Consequently, a superficially attractive rhetorical argument may subsequently be rejected following more dispassionate argument and evidence concerning the consequences of the theory. It is for this reason that a free market in ideas is essential. If criticism never occurs in such a market then it is difficult to distinguish ‘rhetorical’ from ‘legitimate’ ethical argument.

An alternative formulation of the ‘ethics is too complex’ argument is that ‘most people are intellectually and educationally incapable of grappling with the type of issue raised here’. This is a view stated forcefully by Reinhardt [10] in a somewhat different context:

‘I, for one, believe that, if this nation is ever to have … a truly humane safety net all around, an elite espousing these goals would have to impose that state of affairs on generally confused plebs that have quite unstable, often logically inconsistent and utterly malleable preferences.’

Reinhardt, of course, was not the originator of this view. Plato thought of it first and it has been one of the most influential arguments in favour of totalitarianism since its first publication in ‘The Republic’. Setting aside this not-so-subtle argument ad hominem it is not self evidently true that the Reinhardt/Plato view of the individual is correct. At best it remains an unproven hypothesis. The successful testing of ethical issues and the derivation of apparently sensible conclusions in the large number of studies cited earlier suggest that the hypothesis is false. At worst it may be asserted that the conclusions from these studies are unreliable or invalid. But this hypothesis also needs to be tested. If, indeed, it is found that results are unacceptable then the appropriate response would be to improve the techniques for conveying simple ideas to the population. There has been remarkably little inquiry into these methodologies. (For example, a world survey of the literature documenting techniques for eliciting deliberative responses found virtually nothing. Cognitive psychologists we contacted confirmed the desolation of the field.) If subsequent developmental work still encounters a fundamental cognitive barrier to some groups of the population then it may be legitimate to eliminate such groups from empirical ethics surveys. The direction of bias introduced by data censoring is clearly capable of investigation. In sum, the elitist view of intellectual discourse may (or may not) accurately describe the past. It cannot demonstrate the inevitability of its assertions in the future as the methodology suggested here has not been developed or seriously tested.

5 Conclusions

A final criticism of empirical evidence may be that it is ‘arbitrary’. Ultimately it has no solid foundation. However, like Platonic Orthodoxy and the conventional discipline of ethics there is no ultimate source of authority and, sooner or later, an assumption must be made to justify the transition from positive to normative analysis.

An appropriate response to this ‘criticism’ might be ‘exactly so’! There is no ultimate authority and this should be made very explicit in social research: the implications of this disturbing fact should be clearly articulated and discussed. But this has not occurred. The discipline of economics has buried this pivotal fact and treated ethical issues as being technical matters where economists have particular authority. In doing this orthodox economics has diminished the intellectual integrity of the discipline and impoverished its foundations.
While it is true that, like Platonic Orthodoxy, Empirical Ethics has no ‘solid’ (error free) foundation, this is not equivalent to being arbitrary: nor does it make different approaches to ethical analyses equivalent. Empirical Ethics is based upon the hypothesis that normative propositions in economics will be more acceptable to the population and to policy makers when their ethical foundation is based upon stable population values elicited with deliberative procedures and subjected to ethical criticism.

References

6 Richardson J. Why economic costs may not be of interest in a national health scheme, or: costs, fairness and reverse order analysis. Working Paper 126, Centre for Health Program Evaluation, Monash University, Melbourne 2001.