MONASH University
Business and Economics

Centre for Health Economics

Research Paper 2009 (34)

A Critique of Efficiency-Focused Economic Evaluation in the Context of a NHS: The Case for Empirical Ethics

(For publication in M Schlander (ed) *Economic Evaluation of Health Care Programs: Current Concepts, Controversies and International Experience.* Springer, New York)

Professor Jeff Richardson
Foundation Director, Centre for Health Economics,
Monash University

Dr John McKie
Research Fellow, Centre for Health Economics,
Monash University

Dr Kompal Sinha
Research Fellow, Centre for Health Economics,
Monash University

March 2009

Centre for Health Economics
ISSN 1883-1173
ISBN 1 921187 33 X

ACKNOWLEDGEMENTS

The research described in this paper is made possible through the National Health and Medical Research Council (NHMRC) Project Grant ID 491162 and NHMRC Professorial Fellowship ID 284363.

(Forthcoming in M Schlander (ed), Economic Evaluation of Health Care Programs: Current Concepts, Controversies and International Experience. Springer, New York)

Correspondence:

Centre for Health Economics
Faculty of Business and Economics
Building 75
Monash University
Clayton
Melbourne, Victoria 3800
Australia

Phone: +61 (0)3 9905 0754, Fax: +61 (0)3 9905 8344
Jeff.richardson@buseco.monash.edu.au

TABLE OF CONTENTS

1. Introduction 1
2. The failure of the orthodox neoclassical framework 2
3. Fairness and Evaluation Theory 3
4. Evidence on behaviour and motivation 7
5. Problems with costs 10
6. Empirical Ethics 15
7. Conclusion 15

References 17

List of Tables

Table 1. Results from the Monash Ethics Survey 9
Table 2. Spearmen Correlation Coefficient and Efficiency Index 13
Table 3. Probability of allocating resources to a patient 14
Table 4. Increase in life years which offsets the effect of the maximum shift in cost/LY years 14

List of Figures

Figure 1. Does Pareto efficiency increase social welfare 5
Figure 2. Revealed preferences and the resurrected primary of utility 7
Figure 3 Web based allocation exercise 11
Figure 4. Orthodox economic ordering 12
Figure 5. Extreme egalitarian ordering 12

ABSTRACT

We argue that the orthodox evaluation of health services has significant shortcomings with its measurement of benefits and a fundamental conceptual problem in its use of costs. These are important for the distribution of health resources in Australia and internationally, as the Pharmaceutical Benefits Advisory Committee (PBAC), Medical Services Advisory Committee (MSAC), and overseas governments broadly accept the authority of orthodox evaluation theory and follow its recommendations.

The Paper commences with a critique of welfare theory. We argue that the Pareto principle – one of the foundations of welfare theory – is neither necessary nor sufficient to maximise social welfare and that the common defence of utility as the maximand in orthodox studies rests upon a tautology, not evidence.

In contrast with current theory, empirical evidence suggests the importance of duty as a dominating motivational factor. While this evidence also suggests that fulfilling duty may be consistent with maximising happiness, it implies a focus upon relationships and community which is entirely absent from currently history.

When asked to allocate a limited budget between individuals, respondents – including a significant number of academic economists – almost without exception, shared the resources despite the cost of treatment varying significantly between the individuals; that is, they explicitly did not maximise health per dollar but selected a trade-off which heavily favoured sharing over cost-minimising but, consistent with the evidence cited above, focused upon relationships between people.

The implication of these results is profound for theory and practice. It suggests that theoretical framework should be primarily fairness, and not efficiency focused. In practice, there should not be a fixed threshold cost per QALY as a criterion for project selection, as the entitlement to a share of the budget is only weakly related to treatment costs.

A critique of efficiency-focused economic evaluation in the context of a NHS: The case for Empirical Ethics

1. Introduction

The theme of this chapter is that significant deficiencies exist in the measurement and use of both the costs and the benefits associated with health programmes and services.

On the benefit side, utility is too restrictive to serve as the sole objective for the purposes of economic evaluation in the context of an NHS, given the narrow interpretation of "utility" in neoclassical theory. The literature reveals other types of benefit that have been ignored: in the present context, in particular, objective criteria such as health per se. On the cost side, the practice of minimising costs per unit of benefit has also been rejected for reasons of fairness in population-based surveys. The underlying reason for these problems is the imposition upon evaluation methods of the straitjacket of orthodox neoclassical theory. While steps have been taken elsewhere to overcome the limitations of this approach, it continues to exert a strong influence in health economics.

This theme continues in Section 2 below in which we briefly review some of the problems embodied in the key assumptions of neoclassical theory. In particular, the assumption that there can be value-free comparisons between winners and losers of health program decisions is challenged. It is shown in Section 3 that both of the orthodox concepts used to circumvent this problem fail in the context of an NHS, and probably elsewhere.

Neoclassical theory might be described as an "efficiency-focused paradigm". Fairness tends to enter the theory as an afterthought and, in practice, is commonly neglected altogether. In Section 4 we summarise some earlier evidence to illustrate this claim in the broader context of health economic theory. In contrast, we argue that fairness is the raison d'être of an NHS, and that an overwhelming body of evidence suggests that it is seen by health consumers as central not marginal.

Commencing with Nord et al. (1995), there is a small literature indicating public rejection of the minimum cost per unit of benefit criterion. This literature is reviewed in Section 5. This is supplemented here with the results of a web-based survey that reveals a near consensus, even among economists, that resources should not be allocated strictly according to the minimum cost rule.

Since the applicability of neoclassical theory is severely compromised in the context of an NHS, there is the need for an alternative framework. We argue, in Section 6, that "empirical ethics" should be an important part of this new framework: that is, the measurement of population values, accompanied by ethical criticism of the findings, and subsequent re-measurement. The outcome of this process can then be used to modify evaluation methods or to provide relevant information concerning population values to social decision makers. We advocate this in preference to the search for a technically "correct" solution to social problems which carry only the authority of orthodox economic theory.

2. The failure of the orthodox neoclassical framework

According to neoclassical theory, individuals seek to maximise personal "utility" and social welfare is a function of these utilities (the theory of "welfarism"). At the level of the macro-economy, the theory demonstrates that a perfectly competitive economy will be efficient in the Pareto sense: no one's utility can be increased without another's utility falling (the first fundamental law) and, conversely, that any state that is Pareto-efficient may be achieved as the outcome of a perfectly competitive economy (the second fundamental law). Fairness has one role only: it enters the distribution of initial resources, and this determines which of the possible efficient outcomes will be achieved.

The assumptions that are of greatest relevance for economic evaluation concern the individual, and justify the assertion that each person's objective is the maximisation of their own utility. In its post-1930's incarnation, utility in economics usually refers to the strength of individual preferences as "revealed" in market choices. These are "complete", "independent" and "coherent", ie they cover all decision contexts as determined by the individual.

Concomitantly, individuals are assumed to be "rational". While this ambiguous term is sometimes defined formally and narrowly to mean "consistent" or "transitive" in the ranking of outcomes, the term is usually used more broadly to overcome Sen's objection that people who are foolish will be considered "rational" according to the narrow definition as long as they are consistently foolish (Sen 1976). To incorporate risk and uncertainty in the edifice, neoclassical theory posits that rational individual behaviour will be consistent with the axioms of Von Neumann and Morgenstern expected utility theory (EUT). The core of the rationality assumption, however, is that people are assumed to be sufficiently rational that when they are offered choice, and have sufficient information, they will select the outcome that maximises their utility.

When applied to health economic evaluation, these assumptions result in the orthodox prescription of either a willingness to pay or standard gamble metric, as these will best measure individual benefits. Efficiency will be achieved when benefits measured this way are equated with or ranked according to costs and the "efficiency" achieved in this way is almost universally deemed to be desirable.

In contrast, National Health Schemes (NHS) have been created in every developed country (including the USA) for reasons of fairness. The orthodox assumption is generally made (although seldom discussed) that fairness will be achieved through the distribution of initial resources – in this context, through the financing of the NHS – and not by modification of the measurement of benefits. Other grounds for government intervention are generally recognised in the health sector as a result of impediments to the market mechanism, which would otherwise achieve an efficient outcome as defined by the Pareto criterion. But these do not impact on the normative goals of individual utility maximisation or the enhancement of social welfare understood in terms of it.

While elegant, virtually every element of welfare theory is problematical, with the remoteness of the assumptions from the real world differing only in degree. For example, there is extensive evidence indicating that preferences are not complete. Poor information and complexity alone ensure this (Shiell, Hawe et al. 2000; Shiell, Seymour et al. 2000). Similarly, regarding coherence, large numbers of tests have demonstrated preference reversals and inconsistencies due to incoherence or cognitive limitations (Tversky, Slovic et al. 1990; Ariely, Loewenstein et al. 2003; Sugden 2004). The advertising industry has long believed that preferences are highly manipulable! Experimental evidence suggests that, as a result of complexity, effort and regret, choice does not always increase satisfaction (Bownds 2003; Schwartz 2004).

The meaning of "rationality" has been a subject of ongoing debate, but of chief relevance here are the well-known limitations on problem-solving in certain contexts. Experimental evidence once again suggests that this is particularly defective when numerical calculations or probabilities are involved (Kahneman, Slovic et al. 1982; Tversky and Kahneman 1986; Kahneman, Knetsch et al. 1991). Regarding the latter, the assumption that "uncertainty" can be satisfactorily represented by "risk" is an oversimplification, as risk presupposes a significant level of knowledge that individuals commonly do not have, especially in the health context. Furthermore, as argued in the chapter by Richardson and Pope (this volume), even with this information, EUT is invalid empirically and theoretically.

Finally, evidence is cited below that casts serious doubt on the two core normative assumptions of welfare theory: the assumption that individuals seek to maximise utility, and that social welfare is a function of utility. Indeed, it is questionable whether the individual or society maximises anything, or whether this notion is just an analytical convenience with little connection to the real world.

Theoretically, for every interpretation of "utility" there is a form of welfarism, since the latter is a function of the former. We have outlined above some of the key features of "utility" as it figures in neoclassical theory: e.g. preference satisfaction, where preferences are complete, independent and coherent, and revealed by market choices. It is the form of welfare characterised by "utility" so defined that is the target of the critique that follows. This allows that there are more sophisticated forms of welfarism incorporating wider interpretations of utility - in particular, those that take account of "participatory utility" (which can accommodate some communitarian values), "process utility" (which can accommodate fair procedures), and so on. It remains the case, however, that fairness has been largely neglected in the context of an NHS, and the analytical concepts found in orthodox neoclassical theory have been an important contributory reason for this. We briefly return to this issue in the second part of the next section.

3. Fairness and evaluation theory

Welfare theory states that marginal benefits should equal marginal costs. However, in an NHS – a collective insurance scheme – payment is a collective decision undertaken on behalf of the society by government. As pointed out by Richardson and McKie(Richardson and McKie 2007a), the marginal benefits that are dispensed via an NHS depend upon social generosity, which will depend upon a number of factors, only two of which are costs and the individual's own evaluation of benefits. The distribution of these benefits and costs will also affect collective generosity implying, inter alia, the likely relevance of some transfer payments. In this framework fairness and social values play a central role.

The elements of welfare theory that result in the side lining of fairness and other social values were noted above. "Macro level" fairness is achieved by a redistribution of resources. Economists may then focus upon the task of achieving or replicating the outcome of a competitive economy, as this will achieve Pareto efficiency, given the desired redistribution of resources. The fairness of this outcome depends upon the belief that Pareto improvement is also fair. While the Pareto criterion per se only provides a definition of efficiency, increasing one person's utility without reducing another's utility has been accepted as desirable, given a fair initial redistribution of resources.

When it is applied to economic evaluation, welfare theory appeals to the potential compensation or "Kaldor-Hicks" criterion. This is in recognition of the fact that economic evaluation typically creates winners and losers, and a direct comparison of the magnitudes of the gains and losses is difficult. The criterion, or principle, states that if benefits could be redistributed, post hoc, to

achieve Pareto efficiency, then the state should be deemed desirable. Translated into the context of economic evaluation, provision of a drug, for example, would be considered desirable if it led to net benefits that were large enough to leave recipients better off while compensating those who were worse off, even if resources were not redistributed in this way.

In sum, welfare theory provides theoretical justification for a focus on efficiency. Redistribution at either the "macro level" or to compensate losers is the role of others, and economists can focus on the achievement of potential Pareto efficiency. The rationale for this focus, therefore, depends to an important extent upon the concepts of Pareto efficiency and potential compensation.

The first problem is that "potential" is not the same as "actual". No amount of persuasion is likely to convince a person whose house has been resumed that this outcome is desirable because the resulting surplus generated by the freeway through their former property is more than sufficient to compensate them for their loss, even though they have not been compensated. They may agree that the new situation is "efficient" according to the Kaldor-Hicks criterion (indeed, they could not disagree if the criterion is treated as a definition), but they would certainly object that it is unfair. The Kaldor-Hicks criterion does not therefore permit economists to recommend policy unless accompanied by the caveat that the effect upon fairness – in this case the distributional consequences – must be independently judged or compensation actually paid. In practice, this second step is seldom if ever undertaken. Ongoing compensation for permanent loss would almost certainly be opposed by economists as excessive government interference and because of the adverse incentive effects of welfare payments and taxation.

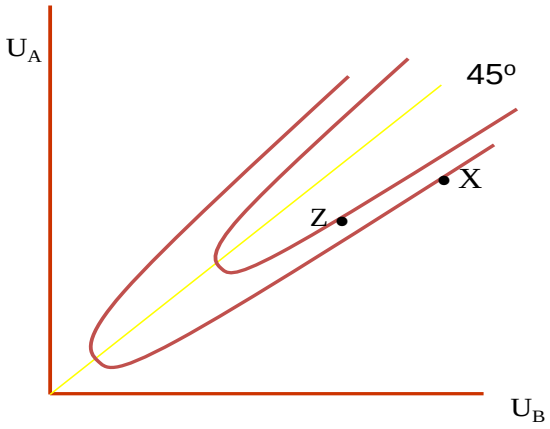
A second problem which arises in the health sector is that when patients are left in a state of extreme disability or allowed to die, compensation is impossible, even in principle, and the Kaldor-Hicks criterion is irrelevant[1].

Third, in the context of a National Health Scheme, the argument for an efficiency focus is not simply potentially inapplicable, but it conflicts with the purpose of most or all NHS's. This is to redistribute the cost of illness from those who are sick to the taxpayer, ie in broad terms, to carry out a redistribution from the healthy wealthy to the less wealthy sick (by comparison with a scheme where people pay a risk-related premium). In an NHS with this objective, the healthy wealthy taxpayer is the ultimate loser financially. Compensation would involve a reversal of this redistribution. The unhealthy less-wealthy who benefitted from the NHS would be taxed in order to compensate the wealthy-healthy. But this is not possible for social-fairness reasons: compensation of this form is not an option and Pareto efficiency cannot be achieved.

Of course, taxing the sick would probably anger a large percentage of the population and reduce their utility. In other words, if the effects upon the population at large are taken into account, and those effects are broadened to incorporate the damaging effects of unfairness, 'compensation' is unlikely to be achieved. In the neoclassical framework, however, this is seldom taken into account.

[1] A caveat is that, in principle, some people may be prepared to die if compensation to their heirs was sufficient. We are unaware of any empirical evidence of the number of people (or their heirs) who would accept this option or the amount of the compensation necessary. However, it is likely that the numbers and expenditures required make this possibility a theoretical escape clause for the compensation principle rather than a serious basis for a particular form of economic evaluation.

Figure 1. Does Pareto efficiency increase social welfare



Finally, we suggest there is another reason why fairness plays such a minor role in orthodox neoclassical theory. This is the near consensus that increased Pareto efficiency leads to an improvement in social welfare. In fact, this is not necessarily true, as is illustrated in Figure 1. The utilities of two individuals are shown on the two axes. The social welfare function is represented by two "social indifference curves" that wrap around the 45 degree line, which is the line of perfect egalitarianism – ie where utilities are identical. The point X dominates Z in the Pareto sense: the utilities of both A and B are greater at X than Z. Nevertheless, X lies upon a lower indifference curve than Z.

In the case illustrated in Figure 1, the "anomaly" is explained by a second criterion of social welfare - viz. the importance of fair outcomes. Even given a fair initial redistribution of resources, and despite the fact that both individuals have higher utility at X than Z, they might disapprove of inequality to such an extent that they would select Z as a preferable social situation to X. The most that can be claimed, therefore, is that a move from Z to X would improve "efficiency", not that it is preferable all things considered.

More generally, the Pareto criterion may conflict with social welfare if this is not welfarist – that is, if it depends upon more than utilities and is also affected by some other ethical or political principle. Pareto efficiency requiring a loss of political liberty or reinforcing discrimination might reduce social welfare; likewise achieving Pareto efficiency at the expense of procedural justice, various forms of duty including religious or social rules, loss of community, destruction of large or small-scale ecosystems and unfair treatment of future generations. Sen (1970) provides a formal proof of the potential incompatibility of liberalism and Pareto optimality.

Defence by Elaborated Definition

In all of the cases above it is possible that individuals gain utility directly from knowing that policy decisions are the result of fair procedures, from satisfying "the apparently universal desire and willingness to share"(Lindsay 1969, p. 351), from observing a vibrant community, and so on. This raises the question whether there is anything about social welfare that could not be captured by "utility". However, there are a large number of actions falling under the heading "duty" that people may willingly undertake even though there is no obvious personal pay-off for them. This includes duties to children, family, friends and fellow citizens. Individuals may support actions that reduce global warming even though it is to their immediate detriment (for example, through the imposition of a carbon tax or the prohibition on coal exports). Principles of humanitarian behaviour, including

the rejection of exploitation, may trump obviously self-interested goals. Some might be prepared to reduce their living standard in order to increase foreign aid or to accommodate refugees in their country. Self-interest might be sacrificed to achieve goals associated with respect for others, gender equality or religious principles (Krebs 1982; Kahneman, Knetsch et al. 1986a; Andreoni 1995; Keser 1996; Bolton and Ockenfels 2000). Individuals may make personal sacrifices rather than have infringements of individual liberties.

Some of these actions might be explained by "sympathy" (where the welfare of one person increases one's own), "reciprocal altruism" (where there is the anticipation of a future pay-off), "participatory utility", "process utility", and so on. However, neoclassical theory often fails to consider these possibilities. Moreover, it remains possible that some of these behaviours might be the result of what Sen calls "commitment", which is strictly counter-preferential: "a person choosing an act that he believes will yield a lower level of personal welfare to him than an alternative that is also available" (Sen 1976). While it is possible that some prima facia sacrifices may be motivated by the expectation of personal utility gain, it cannot simply be assumed that all, or even the majority of such behaviours, are in this category[2].

It is possible to devise a definition of utility that resurrects its universal applicability and aligns Pareto efficiency (defined in these terms) and social welfare. Using Samuelson's revealed preference criterion, it might be argued that if any action, including voting, was observed because of any of the motives above, then this would indicate that the individuals involved were obtaining utility from that action - otherwise they would not have done it. Using this argument, the reason for self sacrifice is that it maximises utility. In the extreme case of one person giving their life for another, this must be construed as utility maximising. However, the logic here is vacuous. The answer to the question "why does an individual take a particular action?" is that "this action maximises utility". But the answer to the question "how do we know that this action maximises utility?" is "because they have taken the action" (see Figure 2)[3].

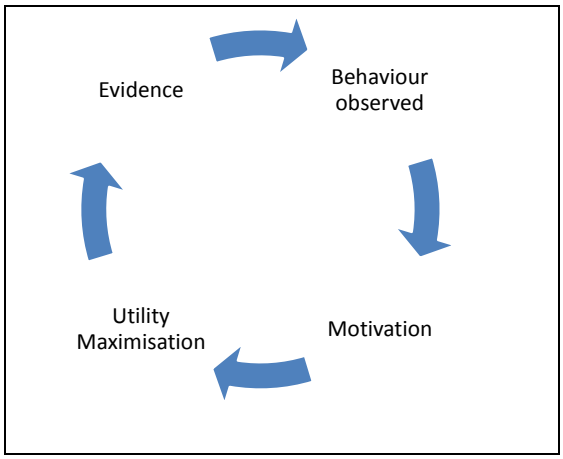
While the logic of this argument is vacuous, its consequences are potentially harmful. If "utility" is generated from the income received by individuals, then the analytical framework might appropriately focus upon the individual and the efficiency of production. However, if it is largely generated by, for example, fulfilling duties to others, the focus may need to shift to fairness and the relative treatment of people within communities. But the need for this is less obvious when there is no explicit distinction between utility from income and utility from duty, and such a distinction is seldom made in economics except as footnotes in some texts and, even then, in purely theoretical terms as a special type of "externality".

[2] Genuinely altruistic behaviour presents a well-known problem for evolutionary theory, but there has been a recent resurgence of interest in "group selection" in evolutionary theory which is compatible with individual "self-sacrifice" (Seeley 1996; Wilson 1997; Sober and Wilson 1998). The jury is still out regarding this complex issue.

[3] One of us (JR) proposes a test for economists who believe that "utility as it is usually used in economic analyses satisfactorily explains all human motivations such as the support for an NHS or obedience to the rules of a religion, etc. This uses the "punch on the nose criterion". When somebody asserts that they truly believe utility maximisation satisfactorily explains all behaviour, the test requires that the experimenter unexpectedly punch the person on the nose. (A punch on the head would suffice, but this is rather lower class.) If the response of the person is "why did you do that?", then they have failed the test. They should have known that the experimenter was simply maximising utility and that this is a fully satisfactory explanation.

The victim might, of course, wish to know the reason why his assailant maximised utility this way (social background, drunkenness, psychiatric illness and so forth). Similarly, we might investigate the reasons for happiness or the background to a particular view of duty. But orthodox economics pays little attention to alternative theories and subsumes numerous possible explanations under the one label, thereby impoverishing analysis and inviting the confusion of the (tautologically) universal concept of utility with the narrow concept of preference consequentialism which becomes the exclusive subject matter of an efficiency focused analysis.

Figure 2. **Revealed preferences and the resurrected primary of utility**



Initially, the term "utility" was used by Jeremy Bentham to denote pleasure and pain. This was not restricted to bodily pleasure and pain, but roughly corresponded to the physiologist's concept of "affect". As a motivation, it became closely associated with selfishness by Edgeworth, in the belief that this would explain the motivation for the narrow range of economic phenomena of interest in the late 19th century. During the 20th century, however, our understanding of individual motivations has increased, thereby requiring new words and concepts to make these distinctions, and the range of activities investigated by economists has increased. But rather than explain these activities by the newly understood motivations, orthodox economics has massaged and rationalised the definition of the all-purpose motivator "utility". The danger inherent in this practice was clearly articulated by George Orwell in an Appendix to "1984", his classic novel on tyranny in which the control of thought was achieved by the truncation of language.

It was intended that when Newspeak had been adopted once and for all ... a heretical thought ... should be literally unthinkable, at least so far as thought is dependent on words.... This was done ... chiefly by eliminating undesirable words ... countless other words such as honour, justice, morality, internationalism, democracy, science and religion had simply ceased to exist. A few blanket words covered them, and in covering them, abolished them. What was required in a party member was an outlook similar to that of the ancient Hebrew who knew, without knowing much else, that all nations other than his own worshipped "false Gods" ... he knew Jehovah and the commandments' of Jehovah; he knew, therefore, that all Gods with other names or other attributes where "false Gods" (George Orwell (1949), *Nineteen Eighty-Four. A novel*. New York: Harcourt, Brace & Co., pp 317-319).

4. Evidence on behaviour and motivation

The distinction between the selfish motivation of the individual, when this is interpreted narrowly, and the motivation of the individual as a citizen, has been noted by numerous writers including Aristotle and Adam Smith, the latter noting in The theory of Moral Sentiments that, however selfish a person may be, "there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it "(Smith 1982). As this quotation suggests, the important distinction for economics is not so much between selfish and selfless behaviour as between isolated self-interest and behaviour conditioned by social interaction. Beinhocker (2007) reports a

substantial body of evidence from controlled experiments and anthropology which indicates that human beings are conditional co-operators and altruistic punishers (Beinhocker 2006). This has also been described as strong reciprocity and defined as "a predisposition to cooperate with others and punish (even at personal cost if necessary) those who violate the norms of cooperation, even when it is implausible to expect these costs will be recovered at a later date" (Gintis, Bowles et al. 2005, p. 8). These behaviours are highly unsurprising in a race whose evolution and survival depended upon social cooperation. Again, this is not to deny that there may be a pay-off for the individual, since strong reciprocators "behave altruistically as long as others are doing so as well" (ibid), and may feel happy about doing so. Rather, it calls into question the naive interpretation of self-interest often associated with neoclassical theory.

Perhaps the most dramatic example of individual behaviour that violates utility maximisation (narrowly construed) is the so-called ultimatum game. In this, one individual is asked to divide a sum of money (e.g. $100) between themselves and a second person, who must accept or reject the offer. With rejection neither person receives any money. If the second person is utility maximising then they will accept any offer. Empirically, however, if the offer is too low (eg. $20) then the second person will reject the option in order to punish the first. Versions of this game have been played with real money many times in countries as varied as the USA, Mongolia and Zimbabwe. But with remarkable consistency, people reject offers that are less than about 30 per cent of the total (Beinhocker 2007 p.20).

Some further evidence from a recent Monash University survey is provided in Table 1. Question 1 represents a partial test of the Pareto principle. Strictly, the question is stated in terms of income whereas the criterion is stated in terms of utilities. But the result still requires an explanation, and several are possible. The two most obvious are, first, that people may wish to prevent the envy, and thus loss of utility, that accompanies discrepancies in wealth, and second, that they consider it unfair independently of any utility considerations.

Questions 2a and 2b test people's support for hedonic utility maximisation. Partly as a result of the influence of Kahneman and other hedonic psychologists (Kahneman, Diener et al. 1999), some researchers have recommended a return to hedonic utilitarianism, and the replacement of preference utilitarianism (Dolan and White 2007). Like Aristotle, some have argued that happiness is the only rational goal (Ng 2004). In apparent disagreement with this, only 14.3 per cent of Australian respondents to our Monash survey held that maximising happiness is more important than any other ethical principle.

In contrast, the motivation that received greatest support was one ignored in orthodox economic theory - namely, duty. Of respondents, 91.5 per cent agreed that they must fulfil some duties even if it makes them a little less happy (Question 3 a). Unreported results shed light on this outcome. Duty to children received support by 98.6 per cent of respondents and duty to spouse and parents 96.8 and 90.8 per cent respectively. People clearly perceive duty as playing an important part in defining their role in society. Ninety five per cent agreed that having duties is a natural part of being a member of society. Of respondents, 91.8 per cent agreed that duty connects people and makes them a community, and 77.8 per cent agreed that they fulfilled duties because it was their role.

Again, this is not conclusive. It is significant, for example, that 84.2 per cent of respondents also agreed with the statement: "I fulfil my duties to individuals and organisations (to family, country etc) because doing so will make me happier in the long run". While 77.8 per cent supported duty as a motivation and only 14.3 per cent supported happiness, this latter result suggests that motivations may not be clearly separated in people's minds. Further, it shows the importance of investigating why people fulfil their duties. It may be out of genuine self-sacrifice (which conflicts with the usual interpretation of welfarism) or it may be because it will enhance their own well-

being, or both. However, as noted earlier, even if duty is ultimately motivated by self-interest or utility, its importance suggests the need for a focus upon relationships and community, which is messing from the orthodox framework.

Table 1. **Results from the Monash Ethics Survey**

Questions/Statements	% (Strongly) Agree	% Neutral	% (Strongly) Disagree
1 Australia is better off if the wealthy receive even higher incomes so long as the income of the poor does not fall.	23.3	13.8	62.8
2a In any situation, the action that will produce the greatest happiness is always right.	22.8	22.8	57.4
2b Maximising happiness is more important than any other ethical principle.	14.3	19.8	65.9
3a I have some duties that I must fulfil even if doing so makes me a little less happy.	91.5	4.7	3.9
3b Having duties is a natural part of being a member of society.	95.0	4.8	0.0
3c I fulfil my duties to individuals and organisations (to family, country etc) not primarily because it will make me or others happy, but because it is my role (e.g. as a mother, father, employee etc).	77.8	5.6	16.7
3d People help others only because they gain something personally.	18.2	21.2	60.7
4a The entire population should receive some assistance for healthcare even when the majority of tax payers are opposed to this.	80.6	11.1	5.6
4b In the context of health it is okay to restrict people's behaviour against their will.	53.8	20.7	25.6

In a separate study, Richardson and Olsen, 2008, set out to test the commitment of two populations to the social principles of welfarism vs. extra welfarism. Individuals in Australia and Norway were asked to imagine that they were on a committee responsible for the allocation of a fixed volume of health resources. The first option was to implement a screening program for cancer that would save a relatively large number of lives, although this was not what the majority of the population supported. The second option was to finance a helicopter rescue service that would save relatively fewer lives, but was very popular with the public (in the question percentage support was cited). Survey respondents were asked to indicate whether they would vote for the service that the population chose or for the service that maximised lives but involved the paternalistic overruling of public preferences.

In three different versions of the questionnaire, the helicopter service was said to be supported (a) for no reason other than a strong public preference, (b) because the helicopter service would allow rapid service delivery in cases of disastrous misadventure (the Rule of Rescue), and (c) because the helicopter service would allow greater access to services for those in remote

areas. Of the 334 Australians and 227 Norwegian respondents, 68.7 and 72.2 per cent respectively supported the maximisation of life. Support for this option dropped to 53.8 and 49.0 per cent respectively in the context of the Rule of Rescue. When equity was the issue, 62.7 per cent of Australians selected life maximisation but, reflecting the greater commitment to equity in Norway, only 48.9 per cent of Norwegians selected life maximisation in this case. Overall the three scenarios indicated a strong commitment to the belief that, at least in the context of an NHS, the role of health services should be to increase health, not to achieve other goals.

This does not imply a rejection of all interpretations of welfarism. Rather, it implies a rejection by respondents of the idea that social welfare is maximised by respecting the self-interested choices of the public (subject to the law, and external effects) – ie by maximising "decision utility". This is essentially the Samuelson version of revealed preferences (which we earlier rejected as potentially tautological). Respondents may have reasoned that utility would be maximised by saving the greatest number of lives. However, the results do imply that, in the health sector, the interpretation of utility would have to correspond with "lives saved". This is defensible in a number of ways that do not require any notion of utility and, in particular, with a view of duty consistent with extra welfarism and the Australian results cited above.

5. Problems with costs

Until recently little attention has been given to the role of costs in economic evaluation. Irrespective of the unit of benefit and how it should be modified, it appears self-evident that this unit should be obtained at the lowest possible cost. However, in 1995 Nord and Richardson found that the public did not appear to agree with this conclusion (Nord, Richardson et al. 1995b). A random sample of 551 Australians were asked whether they agreed or disagreed with a number of statements which included the following:

1. Amongst patients who are equally ill, those who can be helped at low cost should have priority over those who can be helped at high cost because this will allow more people to be helped when money is limited.
2. It is unfair to discriminate against those who happen to have high-cost illness. Priority should therefore not depend on the cost of treatment (except in cases where costs are extremely high).

Only 19 per cent supported the first and 81 per cent supported the second argument. Because of the possibility that people were giving unreflective responses, a sub-group (119) were questioned intensively in the context of an interview. Initially, the two propositions were repeated but emphasising the limited budget available for total treatments. Secondly, interviewers conducted a structured argument which challenged the individual's logic and pointed out that less health would be achieved with the second option. Thirdly, a numerical example was given to illustrate this point and, finally, the individuals were asked to allocate a budget between several options that varied in cost-per-life saved and clearly showed the mortality consequences of the allocation.

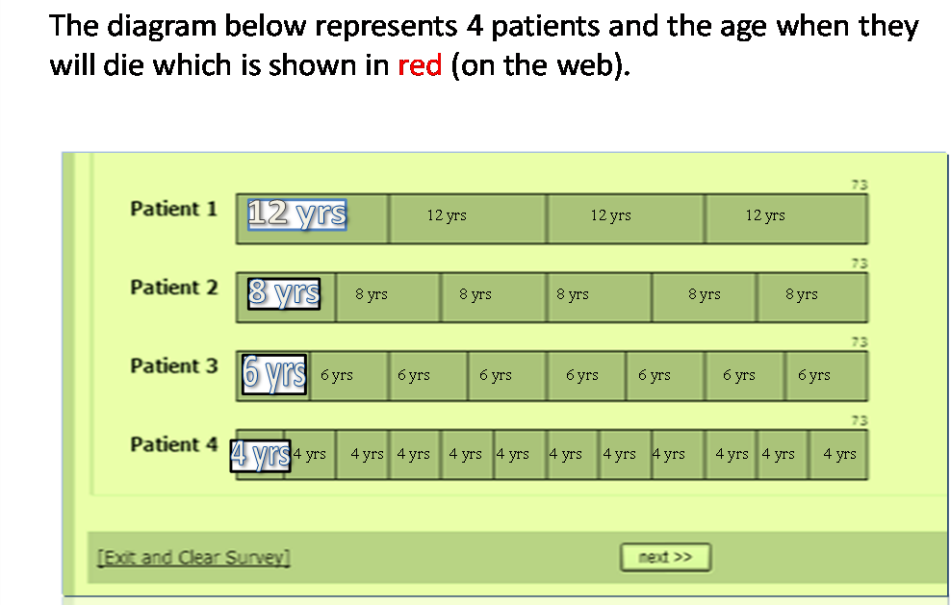
Generally, individuals did not change their views. Only 6 per cent allocated resources to the health maximising option in the final stage.

In Spain, Abellan-Perpiñán, and Prades conducted a similar exercise asking individuals to allocate a budget between two options A and B where the cost of achieving outcome B was double that of achieving outcome A (Abellan-Perpiñán and Prades 1999). Respondents did not allocate all of the resources to option A as predicted by economic theory but allocated them in the ratio 1:2 so that the cost disadvantage of option B was completely offset. Similar results were obtained in a different context by Ubel (Ubel and Loewenstein 1996a), who asked respondents to

allocate resources, this time a limited number of organs, between patients with different prognoses for survival. Rather than allocating entirely to the health maximising group with a good prognosis, respondents allocated some organs to the other groups.

Most recently, the authors conducted a web-based exercise that required respondents to allocate resources on behalf of the Australian National Health Scheme – Medicare. The hypothesis under investigation was that, consistent with previous studies, respondents would not follow a health maximising formula. This study was designed, however, to quantify the deviation from this formula and to test the respective significance of different elements in a respondent's decision.

Figure 3 Web based allocation exercise



Initially respondents were told that they could allocate only $10,000 to one of the four patients shown in Figure 3, which would extend their life by 12, 8, 6 or 4 years respectively (resources could not be shared). Patients were all 25 years old and, without care, faced immediate death. The initial representation of Figure 3 included only the four blocks of years shown in bold and respondents were asked to click on the patient to whom they would give the resources. Respondents were expected to click on the option for Patient 1 (otherwise all subsequent results were deleted). The figure then showed the second block of 12 years for patient 1 and respondents were told that they may/should allocate the second budget of $10,000 to one of the patients. The acceptable choices were now Patient 1 or Patient 2. This procedure continued with new blocks opening on the screen each time a block of resources was assigned to a patient. The exercise terminated when there was no longer any choice. The order in which blocks were clicked – resources allocated – was recorded for analysis.

Figure 4. Orthodox economic ordering

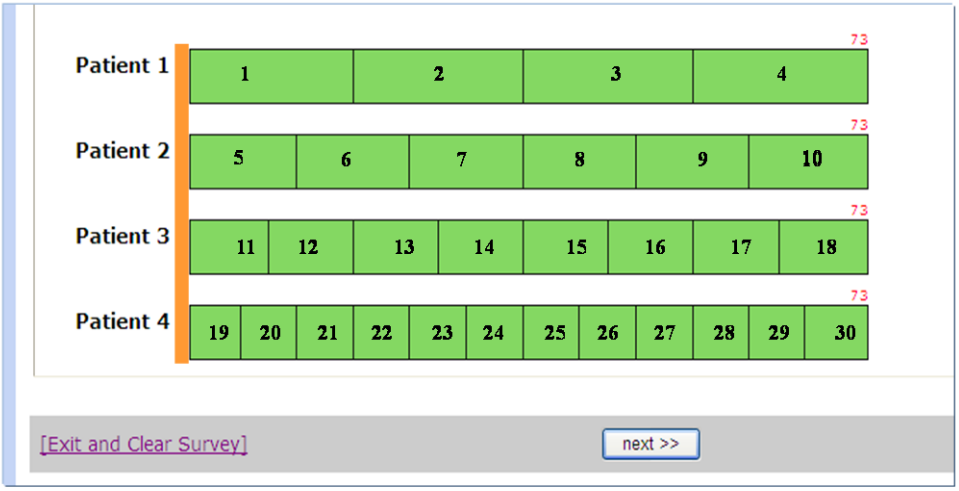
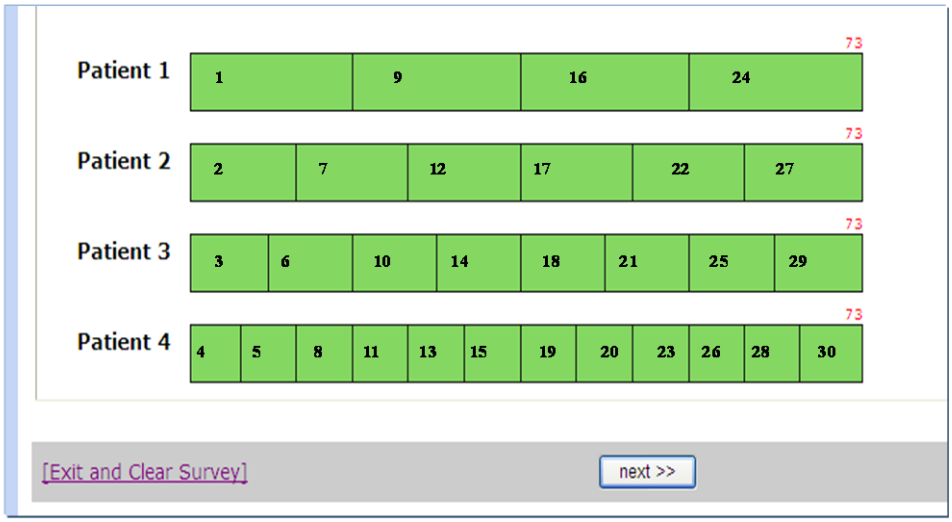


Figure 5. Extreme egalitarian ordering



Two possible outcomes are illustrated in Figures 4 and 5. The first (Figure 4) is the ordering of someone who maximised the number of life years gained. Patient 1 would always receive priority over patient 2 over patient 3 over patient 4. This is very close to the orthodox economic prediction, with some minor variation caused if time discounting was taken into account. The second outcome (Figure 5), in contrast, is the ordering of an extreme egalitarian who allocates resources to the person with the shortest life expectancy and, in cases where patients have equal life expectancy, to the patient gaining most life years. This is tantamount to Rawls' "maximin" principle of justice which would require budgetary allocation be made to the most underprivileged patient. By September 2008 this ongoing project had compiled 14,485 observations from 153 respondents. Of these, 75 per cent were from the general population and 25 per cent from economists specifically recruited from academic departments for this study.

The first analysis conducted was to compare the order in which respondents allocated resources with the order predicted by traditional economic theory and by a perfect egalitarian. To do this, we first calculated the Spearman correlation co-efficient between the actual and predicted order.

Secondly, we constructed a specific "efficiency index" to measure the extent of the deviation of actual choices from the choices that would maximise the number of life years obtained. This was simply defined as the sum of the actual years allocated, divided by the sum of the maximum years that might have been allocated. For example, if a respondent had given patient 2 eight years with their second click and not a second allocation of 12 years to patient 1, then 8 years would be added to the numerator and 12 years to the denominator. A "perfect economist" would obtain a score very close to 1.00 (depending on their rate of time preference) and the perfect egalitarian a score of the sum of the actual years allocated divided by the sum of the maximum years which have been allocated.

Table 2. **Spearmen Correlation Coefficient and Efficiency Index**

Respondent Education/Occupation	No of Individuals	Spearman Corr Coeff		Index (LY Alloc/LY could Alloc)
		Efficiency	Egalitarian	
ECONOMIST	35	0.032	-0.006	0.648
GRADUATE	68	-0.023	-0.019	0.648
OTHER	26	0.137	0.024	0.632
OVERALL	129		-0.007	0.647

Table 2 presents the results of this analysis by type of respondent. The Spearman's correlation coefficient between the respondent's order and the ordering shown in Figures 4 and 5 is reported in columns 3 and 4. As expected, the correlation between the efficient and selected orders was higher for economists than for graduates but, surprisingly, both were lower than for non-graduates. The most striking feature of the results, however, is the small size of the coefficients for all Spearman correlations. The results suggest an almost random association.

The "efficiency index" is essentially the ratio of the years allocated by individuals to the maximum life years that they could have allocated.[4] The average scores are presented in the final two columns of Table 2. Results do not differ significantly by category of respondent. The overall average of 0.647 indicates that respondents allocated 1 - .647 or 35.3 per cent fewer life years than was possible.

In the second analysis we predicted the probability that a particular patient would receive resources. To do this, we used logit regressions (which employ a 0 or 1 as the dependent variable) using the following definitions:

Pr = 1 if a patient received a service
= 0 otherwise
Income = Cumulative No. of clicks (Each click represents a monetary allocation of $10,000 to one of the four patients.)
Cost/LY = 1/ (life years obtained)
LE = life expectancy – ie cumulative life years obtained from received services

[4] For example, if a respondent allocated 4 years to patient *i* at the y^{th} click and there is a possibility of allocating 12 years, then the index for that person on the yth click would be 4/12 and an average of these values would be the overall index score for that person.

For each “click” in the exercise, a respondent created four observations, three for the patients receiving no resources and one for the patient clicked – ie who was assigned the resources. This implies an upper limit of 120 observations (4 x 30 clicks) per respondent.

Logit analyses were carried out to predict Pr from the remaining variables. If all respondents were perfect economist’s, cost/LY would be highly significant and LE insignificant. If all were perfect egalitarians, cost/LY would be insignificant and LE highly significant.

Table 3 reports 3 logit results. The first is for a single individual whose order conforms with the economist prescription in Figure 3. As predicted, the expected costs/life year is highly significant and, along with income, has high predictive power. Life years, the variable of concern to the egalitarian, is statistically insignificant. In contrast, in the second regression, calculated from actual respondents, the significance of cost/life year falls, and most of the explanatory power is transferred to LE. Results for economists (in the third regression) follow the same pattern, although costs/LY increases in significance.

The results indicate that our respondents unambiguously rejected the allocation of resources on the basis of cost minimisation per life year / life year maximisation.

Table 3. **Probability of allocating resources to a patient**

	“Economic Theory” Regression 1			**A-11 – Regression 2**			**Economists Regression 3**		
	b	**z**	**p**	**b**	**z**	**p**	**b**	**z**	**p**
Cost/LY	-76.0	-6.7	0.00	-1.27	-3.29	0.001	-3.12	-3.78	0.00
LE	0.15	0.7	ns	-0.058	-17.69	0.000	-0.07	-10.22	0.00
Income	1.7	11.2	0.00	0.107	18.94	0.000	0.128	10.73	0.000
Constant				-1.02	-14.57	0.000	-0.740	-5.03	0.000
n	72			14,485			3578		
person	1			153			38		
log likelihood				-8319			-2043		

Table 4. **Increase in life years which offsets the effect of the maximum shift in cost/LY years**

All respondents	3.64
Economists	7.45

The quantitative significance of the results is illustrated in Table 4. This reports the increase in life expectancy which the logit equations predict would offset a fall in cost/LY from its maximum to its minimum value, ie from 0.25 (4-1) to 0.0833 (12-1). As life years varied from 0 to 44 the results indicate the relatively small influence of cost/LY. The average result for economists is about double this size and significantly different from the result for the full sample (t = 3.55). While failing to follow economic theory, economists also appear to have different values from the general community.

The implication of these results is that as life expectancy is reduced (all else, including patient age equal) people are more willing to pay higher prices for health services – ie that the threshold cost-per-life-year rises with the severity of a patient's condition.

6. Empirical Ethics

We have argued that orthodox welfare theory has misleading implications for the treatment of both costs and benefits in economic evaluation, at least in the context of an NHS. This leaves open the question of how to determine which costs and benefits to include in evaluation studies and in what way. In very broad terms the answer to the first question is "anything that is of benefit as defined by social goals". This will include the major elements of current economic evaluation – quality and length of life and resource costs, which need little justification. However, determining and quantifying other omitted considerations is more problematical.

There is no technically correct metric for measuring health-related quality of life and, for similar reasons, there can be no technically correct criteria for determining what should be included in an economic evaluation. Welfare theory purports to provide this basis, but does so by adopting a variety of assumptions that are deeply problematical in the health context. In particular, these assumptions deal inadequately with the ethical reasons for the adoption of an NHS.

Most national health schemes are created to achieve objectives relating to fairness that go beyond the simple re-distribution of initial resources. Perceptions of fairness differ between people and countries, particularly with respect to the all-important question of their quantitative importance. Different countries have different institutions and methods of decision making and, once again, there is no technically correct mechanism for social choice. Some countries may elect to use articulated social preferences as the basis for a decision algorithm; others may delegate major social policy decisions to economists or politicians.

We suggest that one important role of economic research should be to identify social objectives, to create and pilot ways in which these might be measured and quantified, to engage in technical debate concerning their validity, reliability, etc., and to engage in ethical debate with the broader community about the inclusion of results in the algorithm used for prioritising health services and programs. The part of this program that deals with eliciting, clarifying and critiquing the community's values we have described elsewhere as "empirical ethics" (Richardson and McKie 2007b).

7. Conclusion

In the present chapter we have argued that economic evaluation has excluded a number of considerations that are of fundamental importance in a NHS, namely NHS-specific objectives and, most obviously, those associated with fairness. Orthodox neoclassical theory has justified its focus on efficiency by appeal to the Pareto principle and the Kaldor-Hicks criterion. These imply that achieving potential Pareto efficiency is a satisfactory basis for recommending policy, given a fair initial distribution of resources. When it comes to prioritising health care, the Pareto principle and the Kaldor-Hicks criterion are largely irrelevant: there will inevitably be winners and losers when health services are prioritised and often compensation will be impossible.

The dimension of fairness that has been most neglected in orthodox theory is the inequality that arises as a direct result of differences in costs. Somewhat inconsistently, economists who have investigated the reasons for deviating from the orthodox formulae have been willing to contemplate public support for higher costs per unit of health gain when this will benefit the more

severely ill, different age groups, and the disabled (McKie and Richardson 2005a; McKie and Richardson 2005b), but not because costs per se are higher. However, the result of differing costs per se may be similar in its impact upon the fairness of priority setting decisions.

The anomaly is almost certainly attributable to the orthodox practice of considering both costs and benefits as "disembodied" units of "value" that may be summed, netted out, and subsequently re-distributed, in the spirit of Kaldor-Hicks, without reference to the different individuals affected. However, costs and benefits are not disembodied in the health sector. They are indissolubly attached to individuals. An "efficient" re-allocation of resources also re-distributes benefits between people and our results clearly indicate that the public is concerned with this. In some cases, a fair distribution of a lesser benefit is perceived as superior to an unfair distribution of a larger benefit, even when this difference is attributable to costs.

Each of the studies that have found a public rejection of the orthodox role of costs has had a common element: variation in costs has explicitly affected the distribution of benefits, and the conclusion to be drawn from these studies is that the public is concerned with the sharing, not just the maximisation of benefits. This is a highly unsurprising result in the context of an NHS, but one that contradicts both the current theory and practice of much economic evaluation.

The evidence we cite here demonstrates the need for considering elements other than health gain and utility in economic evaluations in the health sector, and the importance to the public of sharing resources fairly, even when the sharing does not result in the minimisation of cost per unit of benefit. This is consistent with the ethics that might be expected from a race that evolved in small social units and where survival depended upon social relationships and cooperation.

We have suggested that, more generally and for a variety of reasons, orthodox welfare theory provides an inadequate basis for economic evaluation in the health sector. The alternative which we have described as "empirical ethics" will obviously endorse the major elements of current economic evaluation as of fundamental importance. However, there is a significant task ahead to determine the extent to which present results need to be modified and present methods expanded.

References

Abellan-Perpiñán, J.-M. and J.-L. P. Prades (1999). 'Health State After Treatment: A Reason for Discrimination?' *Health Economics* **8**(8): 701-707.

Andreoni, J. (1995). 'Cooperation in Public Goods Experiments: Kindness or Confusion?' *American Economic Review* **85**(4): 891-904.

Ariely, D., G. Loewenstein and D. Prelec (2003). 'Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences.' *Quarterly Journal of Economics* **118**(1): 73-105.

Beinhocker, E. D. (2006). *The Origin of Wealth: Evolution, Complexity and the Radical Remaking of Economics*. Boston, Massachusetts, Harvard Business School Press.

Bolton, G. E. and A. Ockenfels (2000). 'ERC: A Theory of Equity, Reciprocity, and Competition.' *American Economic Review* **90**(1): 166-193.

Bownds, L. (2003). 'Consumer-Driven Health Plans: More Choice Is Not Always Better.' *Journal of Economic Issues* **37**(2): 425-432.

Dolan, P. and M. P. White (2007). 'How Can Measures of Subjective Well-Being Be Used to Inform Public Policy.' *Perspectives on Psychological Science* **2**(1): 71-85.

Gintis, H., S. Bowles, R. Boyd and E. Fehr (2005). *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, Massachusetts, Massachusetts Institute of Technology Press.

Kahneman, D., E. Diener and N. Schwarz, Eds. (1999). *Well-Being: The Foundations of Hedonic Psychology*. New York, Russell Sage Foundation.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1986a). 'Fairness as a Constraint on Profit Seeking: Entitlements in the Market.' *American Economic Review* **76**(4): 728-741.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1991). 'Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias.' *Journal of Economic Perspectives* **5**(5): 193-206.

Kahneman, D., P. Slovic and A. Tversky, Eds. (1982). *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge, Cambridge University Press.

Keser, C. (1996). 'Voluntary Contributions to a Public Good when Partial Contribution is a Dominant Strategy.' *Economics Letters* **50**(3): 359-366.

Krebs, D. (1982). 'Prosocial Behavior, Equity, and Justice'. *Equity and Justice in Social Behavior*. J. Greenberg and R. L. Cohen. New York and London, Academic Press: 261-308.

Lindsay, C. M. (1969). 'Medical Care and the Economics of Sharing.' *Economica, N.S.* **36**(144): 351-62.

McKie, J. and J. Richardson (2005a). Neglected Equity Issues in Cost Effectiveness Analysis - Part 1: Severity of Pre-Treatment Condition, Realisation of Potential for Health, Concentration and Dispersion of Health Benefits, and Age-Related Social Preferences. Research Paper 2005 (7), Centre for Health Program Evaluation, Monash University, Melbourne.

McKie, J. and J. Richardson (2005b). Neglected Equity Issues in Cost Effectiveness Analysis - Part 2: Direct and Indirect Costs, the Preservation of Hope, the Rule of Rescue, Patient Adaptation, and the Ex Ante/Ex Post Distinction. Research Paper 2005 (8), Centre for Health Program Evaluation, Monash University, Melbourne.

Ng, Y.-K. (2004). *Welfare Economics : Towards a More Complete Analysis*. New York, Palgrave Macmillan.

Nord, E., J. Richardson, A. Street, H. Kuhse and P. Singer (1995b). 'Who Cares About Cost? Does Economic Analysis Impose or Reflect Social Values?' *Health Policy* **34**(2): 79-94.

Richardson, J. and J. McKie (2007a). 'Economic Evaluation of Services for a National Health Scheme: The Case for a Fairness-Based Framework.' *Journal of Health Economics* **26**(4): 785-799.

Richardson, J. and J. McKie (2007b). 'Economics, Political Philosophy and Ethics: The Role of Public Preferences in Health Care Decision-Making'. *Principles of Health Care Ethics*. R. E. Ashcroft, A. Dawson, H. Draper and J. R. McMillan. Chichester, John Wiley and Sons.

Schwartz, B. (2004). *The Paradox of Choice: Why More is Less*. New York, HarperCollins.

Sen, A. (1976). 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory.' *Philosophy and Public Affairs* **6**(4): 317-344.

Shiell, A., P. Hawe and J. Seymour (2000). 'Will Our Understanding of Completeness Ever be Complete?' *Health Economics* **9**(8): 729-731.

Shiell, A., J. Seymour, P. Hawe and S. Cameron (2000). 'Are Preferences Over Health States Complete?' *Health Economics* **9**(1): 47-55.

Smith, A. (1982). *The Wealth of Nations: Books I-III*. London, Penguin Classics.

Sugden, R. (2004). 'The Opportunity Criterion: Consumer Sovereignty without the Assumption of Coherent Preferences.' *American Economic Review* **94**(4): 1014-1033.

Tversky, A. and D. Kahneman (1986). 'Rational Choice and the Framing of Decisions.' *The Journal of Business* **59**(4): S251-S278.

Tversky, A., P. Slovic and D. Kahneman (1990). 'The Causes of Preference Reversal.' *American Economic Review* **80**(1): 204-217.

Ubel, P. A. and G. Loewenstein (1996a). 'Distributing scarce livers: The moral reasoning of the general public.' *Social Science and Medicine* **42**(7): 1049-1055.