

Egalitarianism versus Utilitarianism

KEN BINMORE

ELSE, University College London

This paper is a comparative analysis of egalitarianism and utilitarianism from a naturalistic perspective that offers some insight into the manner in which we come to make interpersonal comparisons of welfare.

There is no such thing as justice in the abstract;
it is merely a compact among men.

Epicurus

I. INTRODUCTION

To those of us who believe that ethical systems evolved along with the human race, debates between followers of absolute conceptions of the Good and the Right have something of the flavour of the Mad Hatter's tea party. Each side fires off absurdly contrived moral questions¹ and then begins to poke fun at the answer given by their rivals before they have heard what it is. Even if they listened to each other, we moral naturalists cannot see what either side can hope to achieve from such philosophical sparring. If a moral problem is sufficiently outlandish, it will have occurred so infrequently in our evolutionary history that it would be amazing if our moral intuitions were adapted to solving it in a way that made any sense.

In a recent book, I argue that it makes no more sense to dispute whether the right or the good is more fundamental to morality than to dispute whether the feasible or the optimal is more fundamental in decision theory.² Theories of the right are invented by deontologists who focus on the social mechanisms that have evolved to *sustain* a social contract. Theories of the good are invented by consequentialists who focus on the social mechanisms that have evolved to allow a society to *select* a new social contract when the environment changes. On this naturalist view, deontology and consequentialism are repositories of complementary moral intuitions that need to be pruned and then grafted together if a model adequate to describe the moral conventions of a real social contract is to be created.

In this paper, I try to identify the natural source of an intuition that simultaneously inspires both the deontological theory of Rawls and

¹ For example, would we be entitled to steal an eye from a sighted person if it could be used to make a blind man see?

² K. Binmore, *Just Playing: Game Theory and the Social Contract II*, Cambridge, MA, 1998.

the consequential theory of Harsanyi.³ Even if one rejects this attempt at reconciling the two traditions, the argument shows that slanging matches between the two sides are pointless unless the question of interpersonal comparison of utility has been addressed first. Only then does it make sense to ask how egalitarian solutions to moral problems differ from utilitarian solutions.

II. NATURALISM

Moral philosophers typically hold that the purpose of their discipline is to uncover universal principles that we all ought to follow when interacting with our fellows. A naturalist following in the tradition of David Hume denies that such principles exist. He sees himself as a scientist exploring the biological and social facts on which our moral intuitions are based. Such facts are contingent. They would have been otherwise if biological and social history had taken a different course. Moral behaviour in chimpanzees and baboons differs from moral behaviour in humans because their biological history differs from ours. Moral behaviour in other human societies differs from moral behaviour in our society because their social history differs from ours.

Such frank relativism is too much for many to swallow, because it denies that there are any absolute moral standards. Those who wish to enter the pulpit to preach that one society is better than another are therefore not entitled to appeal to naturalistic theories of ethics. Even the wishy-washy liberal doctrine that all societies are equally meritorious receives no support from naturalism. There is no culture-free Archimedean standpoint from which to apply a moral lever to the world. If we could liberate ourselves from all cultural prejudices, we would find that morality no longer had any meaning for us.

Matters seem different to moral absolutists only because they are unable to break sufficiently free of their childhood conditioning to see that the moral absolutes to which they pledge allegiance would have been different if they had been brought up in another place or time. It is a matter of historical record that the moral absolutists of Plato's circle thought that making love to underage boys was entirely admirable. Presumably they thought it always would be. But we naturalists do not blind ourselves to the obvious fact that our distaste for paedophilia is a cultural artefact of the society in which we live. If we had lived in classical Greece, we would also have emulated Socrates in chasing after adolescent boys.

³ J. Rawls, *A Theory of Justice*, Harvard, 1971; J. Harsanyi, *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge, 1977.

Once such unwelcome truths have been assimilated, it becomes possible to speculate about the evolutionary origins of the actual moral systems whose workings provide absolutists with the ammunition for their metaphysical speculations. What function does morality serve? What do the moral systems of historical record have in common? Should we attribute any common properties to our common biological inheritance? Social evolution doubtless matters as much or more than biological evolution. How does it work? How should we expect social evolution to adapt the moral system a society operates as its circumstances change?

III. JUSTICE AND THE ORIGINAL POSITION

In my forthcoming book, I argue that morality serves as an equilibrium selection device in the human Game of Life.⁴ Anthropologists trace the origins of human sociality to primitive insurance contracts in which individuals guarded against privation by reaching an understanding on how food should be shared. The sharing norms that evolved for this purpose are seen as providing the basis for the notion of fairness with which modern man still regulates his societies. From this point of view, the basic function of a fairness norm is to allow a group of individuals to co-ordinate efficiently on a method of exploiting a new source of surplus without costly internal conflict. Intuitions about fairness are derived from seeing such fairness norms in operation.

Consider, for example, what happens when a dish in short supply is shared at a polite dinner party. If things go well, and they usually do, the dish gets divided without any discussion or intervention by the host. When questioned, everybody will agree that each person should take his fair share.

But how do we know what is fair? This is not a simple question. What is judged to be fair commonly depends on a complex combination of contingent circumstances – like who is fat and who dislikes cheese. Moreover, if we observe what actually happens, rather than what people say should happen, we will find that it also depends on how each person at the table fits into the social pecking order. Woe betide the poor relative sitting at the table on sufferance in the last century who helped himself to an over-generous portion of his favourite dish!

Numerous scholars have tried to make sense of the calculations that people must implicitly have made when they co-ordinate on an outcome that they afterwards describe as fair. It surely can be no accident that

⁴ The first volume of *Game Theory and the Social Contract* appeared in 1994 with the subtitle *Playing Fair*. The second volume will appear in 1998 with the subtitle *Just Playing*; see n.2 above.

the consensus is firmly in favour of some type of do-as-you-would-be-done-by principle. Moralists down the ages have offered numerous arguments that seek to explain why it is morally imperative that each person should follow such a golden rule. But none of these traditional arguments is founded on anything solid. I think we get suckered into taking them seriously because we are too ready to confuse a fairly accurate description of *what* we do in certain situations with an explanation of *why* we do it.

One particularly interesting variant of a do-as-you-would-be-done-by principle is described by Rawls in his *Theory of Justice*. Others have independently proposed conceptually similar schemes, notably Harsanyi. Rawls's *original position* is a hypothetical standpoint to be used in making judgements about how a just society should be organized. Each citizen is asked to envisage the social contract to which he would agree *if* his current role in society were concealed from him behind a *veil of ignorance*.

I agree that the device of the original position does generate compromises that would commonly be regarded as fair, but the Kantian arguments that Harsanyi and Rawls offer when urging its use seem entirely empty to me. I believe that the real reason it appeals so strongly to our intuition is because we *already* use something very like it when making fairness judgements. After all, it is simply a stylized do-as-you-would-be-done-by principle that does not evade objections like: don't do unto others as you would have them do unto you – they may have different tastes from yours.

Formulating the do-as-you-would-be-done-by principle in terms of the original position makes it necessary to model the empathetic preferences that we use when comparing our lot with others. Empathetic preferences need to be distinguished from personal preferences. I reveal a personal preference when I choose to wear a fig leaf rather than eat an apple. I reveal an empathetic preference when I say that I would prefer to be Eve eating an apple than Adam wearing a fig leaf. If I am behind the veil of ignorance, thinking it equally likely that I will turn out to be Adam or Eve when the veil is lifted, there is no way I can assess the possible agreements I might make with my partner in ignorance unless I am equipped with empathetic preferences.

Harsanyi is the latest of a number of scholars who have modelled empathetic preferences in the same way that economists standardly model personal preferences. If the players empathize perfectly with each other – so that Adam agrees that, if he were Eve, he would have Eve's personal preferences – then Harsanyi shows that one can completely summarize a person's empathetic preferences over Adam and Eve simply by stating the rate at which Adam's utils are to be traded for Eve's. In a society operating a common standard of interpersonal

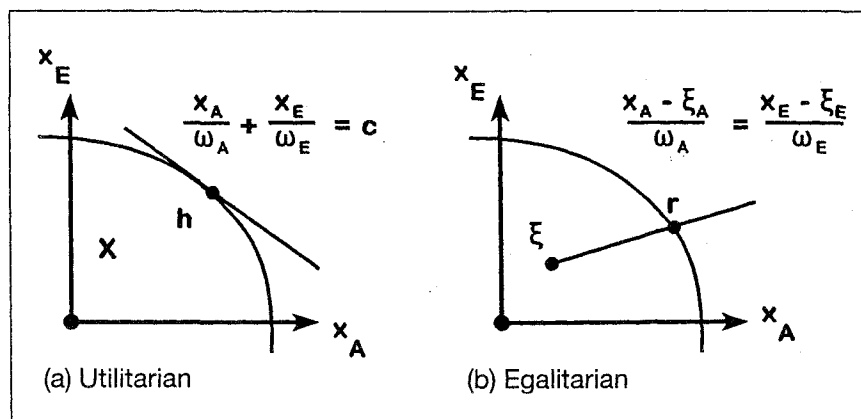


Figure 1

comparison, it is then only necessary to label each person i with a *worthiness* coefficient w_i . One of Adam's utils is then taken to be equivalent to w_E/w_A of Eve's.

Behind the veil of ignorance, each possible agreement can be abstracted to a pair (x, y) in which I get x if I turn out to be Adam and y if I turn out to be Eve. Harsanyi shows that my preferences over such agreements can be represented by a utility function of the form

$$U_i(x, y) = \frac{u_A(x)}{w_A} + \frac{u_E(y)}{w_E},$$

where u_A and u_E are Adam and Eve's personal utility functions. Since both people behind the veil of ignorance will have the same preferences, Harsanyi therefore concludes that we will settle on the utilitarian outcome h illustrated in Figure 1(a), in which X represents the set of available pairs of personal payoffs and $\xi = (x_A, x_E)$ represents our current *status quo*.

Rawls escapes Harsanyi's utilitarian conclusion by the iconoclastic expedient of ditching orthodox decision theory in favour of the maximin criterion. But such iconoclasm is unnecessary. The intuition that led Rawls to his Difference Principle can be vindicated by taking his concern about the 'strains of commitment' to its logical conclusion. In *Just Playing*, the original position is analysed without any commitment assumptions being made at all – especially in respect of the fall of the phantom coin that supposedly decides who will be Adam and who will be Eve.

Following Diamond, imagine that Adam and Eve are both in need of

a heart transplant but that only one heart is available.⁵ Behind the veil of ignorance, it will be agreed that the heart should be assigned to Adam and Eve at random, but Eve would regard it as grossly unfair if the heart were then given to Adam on the grounds that he is a man. She would not be mollified by being told that she had an equal chance of finding herself male or female when the veil was lifted. Nor that she had an equal chance of becoming a man or a woman when her egg was fertilized in her mother's womb. On the contrary, she will insist that there is no reason why she should regard herself as being committed to any randomizing that does not involve a real coin being tossed right now. And if it does not fall to her liking, what is to prevent her demanding that it be tossed again?

In defending their commitment assumptions, Harsanyi talks of *moral* obligations and Rawls of *natural* duties. However, I agree with Hume that matters are not advanced by paraphrasing the claim that we ought to do something because we ought to do it.⁶ Removing the commitment assumption has the additional virtue that it becomes possible to see the fairness norm modelled by the original position as a device to co-ordinate Adam and Eve's behaviour on an *equilibrium* in their Game of Life. No commitment assumptions or appeals to external enforcement agencies are necessary to maintain such an equilibrium, because equilibria are *self-policing* – after an agreement on an equilibrium, each player will find it *optimal* to honour the agreement, provided nobody else cheats first.

Refusing to regard the hypothetical deal reached in the original position as binding has far-reaching implications. In particular, Harsanyi's argument that fairness requires the use of a utilitarian social welfare function can no longer be sustained. Without commitment assumptions, the set X of feasible outcomes must be reinterpreted to be the set of equilibrium outcomes in Adam and Eve's Game of Life. Moreover, instead of maximizing the *sum* of x_A/w_A and x_E/w_E , fairness requires maximizing the egalitarian social welfare function obtained by taking the *minimum* of $(x_A - \xi_A)/w_A$ and $(x_E - \xi_E)/w_E$. We therefore recover the maximin component of Rawls's Difference Principle without needing to overthrow orthodox decision theory.

Economists will recognize the egalitarian social welfare function as the proportional bargaining solution r of co-operative game theory. When X is convex, r is the point at which a straight line of slope w_E/w_A through the *status quo* ξ cuts the boundary of X , as illustrated in

⁵ P. Diamond, 'Cardinal Welfare, Individualistic Ethics and Interpersonal Comparison of Utility: Comment', *Journal of Political Economy*, lxxv (1967).

⁶ D. Hume, *A Treatise of Human Nature*, 2nd edn., ed. L. A. Selby-Bigge, rev. P. Nidditch, Oxford, 1978 (first published 1739).

Figure 1(b). The proportional bargaining solution has been extensively studied⁷ because it is the result to which one is inexorably led once full interpersonal comparison of utility has been assumed.

Once expressed in terms of the proportional bargaining solution, philosophers will recognize the egalitarian social welfare function in Aristotle's dictum: 'What is just ... is what is proportional.'⁸ Psychologists will recognize the egalitarian social welfare function in their empirically based law of equity.⁹ Written as an equation, the law asserts that

$$\frac{G_A}{W_A} = \frac{G_E}{W_E},$$

⁷ See e.g. J. Isbell, 'A Modification of Harsanyi's Bargaining Model', *Bulletin of the American Mathematical Society*, lxvi (1960); E. Kalai, 'Solutions to Bargaining Situations: Interpersonal Utility Comparisons', *Econometrica*, xlv (1977); R. Myerson, 'Two-person Bargaining and Comparable Utility', *Econometrica*, xlv (1977); H. Peters, *Bargaining Game Theory*, PhD thesis, Proefschrift Universitat Nijmegen, 1986; H. Raiffa, 'Arbitration Schemes for Generalized Two-person Games', in *Contributions to the Theory of Games II*, ed. H. Kuhn and A. Tucker, Princeton, 1953; A. Roth, *Axiomatic Models of Bargaining*, Berlin, 1979.

⁸ Aristotle, *Nicomachean Ethics*, trans. T. Irwin, Indianapolis, 1985.

⁹ For example, J. Adams, 'Towards An Understanding of Inequity', *Journal of Abnormal and Social Psychology*, lxvii (1963); J. Adams, 'Inequity in Social Exchange', in *Advances in Experimental Social Science*, vol. II, ed. L. Berkowitz, New York, 1965; J. Adams and S. Freedman, 'Equity Theory Revisited: Comments and Annotated Bibliography', *Advances in Experimental Social Science*, vol. IX, ed. L. Berkowitz, New York, 1976; W. Austin and E. Hatfield, 'Equity Theory, Power and Social Justice', in *Justice and Social Interaction*, ed. G. Mikula, New York, 1980; W. Austin and E. Walster, 'Reactions to Confirmations and Disconfirmations of Expectancies of Equity and Inequity', *Journal of Personality and Social Psychology*, xxx (1974); J. Baron, 'Heuristics and Biases in Equity Judgments: A Utilitarian Approach', in *Psychological Perspectives on Justice: Theory and Applications*, ed. B. Mellers and J. Baron, Cambridge, 1993; R. Cohen and J. Greenberg, 'The Justice Concept in Social Psychology', in *Equity and Justice in Social Behaviour*, ed. R. Cohen and J. Greenberg, New York, 1982; L. Furby, 'Psychology and Justice', *Justice: Views from the Social Sciences*, ed. R. Cohen, Cambridge, MA, 1986; G. Homans, *Social Behaviour: Its Elementary Forms*, New York, 1961; B. Mellers, 'Equity Judgment: A Revision of Aristotelian Views', *Journal of Experimental Biology*, ci (1982); B. Mellers and J. Baron, *Psychological Perspectives on Justice: Theory and Applications*, Cambridge, 1993; D. Messick and K. Cook, *Equity Theory: Psychological and Sociological Perspectives*, New York, 1983; R. Pritchard, 'Equity Theory: A Review and Critique', *Organizational Behaviour and Human Performance*, iv (1969); G. Wagstaff, 'Equity, Equality and Need: Three Principles of Justice or One?', *Current Psychology: Research and Reviews*, xiii (1994); G. Wagstaff, J. Huggins and T. Perfect, 'Equal Ratio Equity, General Linear Equity and Framing Effects in Judgments of Allocation Divisions', *European Journal of Social Psychology*, xxvi (1996); G. Wagstaff and T. Perfect, 'On the Definition of Perfect Equity and the Prediction of Inequity', *British Journal of Social Psychology*, xxxi (1992); E. Walster, E. Berscheid and G. Walster, 'New Directions in Equity Research', *Journal of Personality and Social Psychology*, xxv (1973); E. Walster and G. Walster, 'Equity and Social Justice', *Journal of Social Issues*, xxxi (1975); E. Walster, G. Walster and E. Berscheid, *Equity: Theory and Research*, London, 1978. For a user-friendly book in draft that sets the philosophical scene, and reviews the history and current status of modern equity theory, see G. Wagstaff, *Making Sense of Justice*, Psychology Department, University of Liverpool, 1997.

where g_A and g_E are the respective gains to Adam and Eve, and w_A and w_E quantify how worthy they are. To see the connection between the law of equity and the proportional bargaining solution, simply write $g_A = x_A - \xi_A$ and $g_E = x_E - \xi_E$.

The psychological formulation makes it necessary to face the problem of how worthiness coefficients are to be calculated. Are they measured in terms of social status, merit, effort, need, or what? Harsanyi evades this issue altogether by supposing that its solution is somehow part of Adam and Eve's common heritage in the original position. Rawls invents an index of primary goods for a similar purpose. My own view is that it is a mistake to suppose that some absolute standard for making interpersonal comparisons of utility can be laid down for all times and places. I believe that the way we measure worthiness when making fairness judgements is a function of the way our societies have evolved and are evolving. One therefore cannot evade some attempt to model social evolution if the problem of interpersonal comparison is to be addressed adequately.

IV. EMPATHY EQUILIBRIUM

A standard for making interpersonal comparisons is fundamental to any ethical theory, but I do not know of a treatment in which the subject is dealt with adequately. Rawls plucks his commonly accepted index of primary goods from the air, while Harsanyi's appeal to some rational algorithm that makes the standard a function of what we all have in common is unhelpful without some clue as to the nature of the algorithm. I think safer ground is to be found in the factual observation that there is indeed a substantial level of consensus within a particular society on how utils should be compared across individuals.

A recognition that such a *de facto* consensus exists is implicit in the approach of those philosophers who continue to nurture the impartial spectator tradition of Adam Smith's *Theory of Moral Sentiments*. Their ideal observer need no more be seen as real entity than the invisible hand of Adam Smith's *Wealth of Nations*.¹⁰ He can be viewed instead as a personification of a set of general cultural attitudes that we feel no need to defend because we unconsciously take for granted that they are shared by all. But if we all share common cultural attitudes, it is because we have a common cultural history. So why not try to build the notion of a common cultural history directly into the analysis?

In following up this approach to interpersonal comparison, our

¹⁰ A. Smith, *The Theory of Moral Sentiments*, ed. D. Raphael and A. Macfie, Oxford, 1975 (first published 1759); A. Smith, *The Wealth of Nations*, Indianapolis, 1976 (first published 1776).

surrogates for the relevant aspects of an evolving culture are the empathetic preferences that Adam and Eve bring with them into the original position. It is important that these empathetic preferences are not confused with personal preferences. To keep things simple, it will be assumed that the latter change only in the long run, which I identify with biological time. Personal preferences are therefore fixed in the medium run, which I identify with the amount of time that social evolution needs before an equilibrium is reached. Since empathetic preferences are seen as social constructs, they are then able to adjust to the environment in the medium run. The short run corresponds to economic time, during which preferences are fixed. Since any decisions are made in the short run, they can be studied using standard economic techniques. In particular, decisions made in the original position can be studied using the economic theory of bargaining.

The previous section discusses two polar cases. When Adam and Eve are entirely committed to the deal reached in the original position, their agreement is utilitarian. When they feel entirely uncommitted, their agreement is egalitarian. Taking the set X of feasible outcomes as an idealization of the set of historically possible social contracts, *Just Playing* argues that the players' empathetic preferences will adjust in the medium run until an equilibrium is reached. To test whether a pair of empathetic preferences constitutes an empathy equilibrium each player should be asked the following question: 'Suppose that you could deceive everybody into believing that your empathetic preferences are whatever you find it expedient to claim them to be. Would such an act of deceit seem worthwhile to you *in the original position* relative to the empathetic preferences that *you actually hold*?' The right answer for an empathy equilibrium is, 'No'.

At a symmetric empathy equilibrium, the worthiness coefficients are easily computed. Whether in the utilitarian or egalitarian case, one first locates the symmetric Nash bargaining solution n for the bargaining problem in which the set of feasible payoff pairs is X and the *status quo* is ξ . Figure 3(a) illustrates this notion. Observe that Adam and Eve's worthiness coefficients are not used in locating n . In fact, one of Nash's characterizing axioms insists that his solution be independent of any standard for making interpersonal comparisons of utility.¹¹ It therefore has no virtue as an ethical concept. Indeed, Nash intended his solution to predict the bargaining outcome when two players negotiate face-to-face using whatever bargaining power they may have at their disposal. One might therefore say that social evolution will erode all moral content from a utilitarian or egalitarian

¹¹ J. Nash, 'Non-cooperative Games', *Annals of Mathematics*, liv (1951).

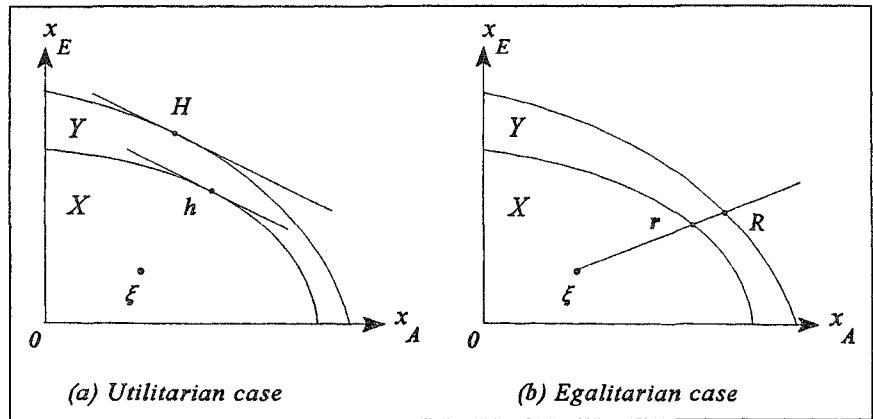


Figure 2

fairness norm in the medium run. However, it is important to recall that decisions are taken in the *short run*. If the set of feasible payoff pairs expands from X to Y , the short run will see Adam and Eve still operating their old fairness norm with worthiness coefficients adapted to X as illustrated in Figure 2(a) for the utilitarian case and in Figure 2(b) for the egalitarian case.

V. COMPARING EGALITARIANISM AND UTILITARIANISM

Rawls's treatment of his index of primary goods reveals that he thinks achieving a consensus on how interpersonal comparisons should be made is relatively unproblematic compared with the other issues he discusses. Harsanyi is equally sanguine about the prospects of consensus when he observes that: 'In actuality, interpersonal utility comparisons between persons of similar cultural background, social status, and personality are likely to show a high degree of inter-observer validity.'¹² I share Harsanyi's and Rawls's impression that we do in fact see a substantial level of agreement on how to make interpersonal comparisons when a group of people are working smoothly together. But my theory predicts that the consensus should *vary with the context*.

The view that fairness judgements are made in different ways in different circumstances has been defended by numerous authors. In recent years, Elster and Young have urged this point particularly strongly.¹³ However, it usually seems to be taken for granted that the

¹² Harsanyi, *Rational Behaviour*, p. 60.

¹³ J. Elster, *Local Justice: How Institutions Allocate Scarce Goods and Necessary Burdens*, New York, 1992; P. Young, *Equity*, Princeton, 1994.

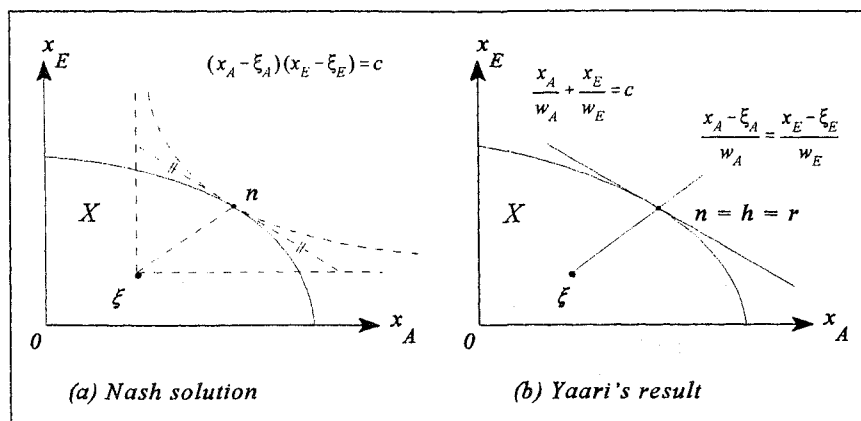


Figure 3

standard for making interpersonal comparisons does not vary much between contexts, leaving differences in what is deemed to be fair to be explained by variations in the fairness norm in use. By contrast, my theory assumes that the fairness norm is always the device of the original position, no matter what the context may be. Differences in fairness judgements are therefore attributed to contextual variations in the standard for making interpersonal comparisons of utility.

My claim that differences in fairness judgments are not caused by changes in the fairness norm in use looks as though it needs to be qualified with the reservation that the device of the original position leads to different results depending on whether or not Adam and Eve regard themselves as being committed to the hypothetical deal reached behind the veil of ignorance. However, it turns that we are led to the *same* worthiness coefficients whether we study a utilitarian or an egalitarian fairness norm at a symmetric empathy equilibrium. Figure 3(a) illustrates various characterizations of the Nash bargaining solution. The useful characterization for the purpose at hand is the observation that the tangent to X at n and the ray from ξ to n make the same angle to the vertical. As Yaari has pointed out, it follows that the worthiness coefficients necessary to make the weighted utilitarian solution h coincide with the Nash bargaining solution are the same as the worthiness coefficients necessary to make the proportional bargaining solution r coincide with the Nash bargaining solution.¹⁴ Figure 3(b) illustrates this result by showing the worthiness coefficients w_A and w_E that identify n , h and r .

¹⁴ M. Yaari, 'Rawls, Edgeworth, Shapley, Nash: Theories of Distributive Justice Re-examined', *Journal of Economic Theory*, xxiv (1981).

In the *medium run*, an egalitarian and a utilitarian will therefore find themselves justifying precisely the same social contract with an appeal to precisely the same standard of interpersonal comparison!¹⁵ But when decisions are taken, it is the *short-run* implications of a moral doctrine that are significant, and changes in w_A and w_E move the utilitarian outcome H and the egalitarian outcome R in *opposite* directions in Figures 2(a) and 2(b).

Adam wishes that w_E/w_A were *bigger* when the weighted utilitarian solution is applied, and *smaller* when the proportional bargaining solution is applied. To understand this point, it is helpful to think of an egalitarian and a utilitarian handing out morsels of the surplus one by one. An egalitarian takes each morsel and splits it between Adam and Eve in proportion to their respective worthiness coefficients, w_A and w_E . A utilitarian gives the morsel to whoever is perceived to gain the most utility thereby. He therefore favours players that get a lot of utility from a small share of the surplus. But such players are precisely the folk who are deemed unworthy by an egalitarian. Since I believe that our evolutionary history has led us to make fairness judgements on egalitarian principles, utilitarians might therefore be characterized by saying that they are in favour of giving priority to the needs of the unworthy, while neglecting the just claims of worthy folk like ourselves.

VI. COMPARATIVE STATICS

What makes one person more worthy than another? Economists regard such a question as raising a problem in comparative statics. Although I am no admirer of Karl Marx, it has to be admitted that he put his finger on all the characteristics of a person that seem to be of interest in studying how a society identifies the worthy. Recall that, after the revolution, workers were to be rewarded according to their labour. The Marxist labour theory of value is certainly no jewel in the crown of economic thought, but it is nevertheless true that the relative levels of *effort* required from Adam and Eve in creating a source of surplus must be one of the major parameters requiring attention in determining their worthiness when it come to splitting the surplus they have jointly created. Of course, according to Marx, labour was to provide only a stop-gap measure of worthiness. In the socialist utopia that would ensue after the state had eventually withered away, the

¹⁵ Even if one rejects my evolutionary approach, it still follows that disputes over policy between utilitarians and egalitarians that do not begin by clarifying how interpersonal comparisons are to be made are pointless, since the argument shows that a standard exists that will lead to their making the same policy recommendations.

rule was to be: from each according to his ability – to each according to his need. Human nature being what it is, such an incentive scheme seems designed to convert the able of a large society into the needy overnight. Nevertheless, both *ability* and *need* are parameters that must be taken into account when evaluating worthiness.

Finally, we need to add *social standing* or *political pull* to effort, ability and need. For Marx, social standing was a feudal survival to be swept away along with all of history's other failed experiments in human organization, but a realistic social contract theory cannot ignore the fact that the power structures of today evolved from the power structures of the past. A person's social standing, as measured by the role assigned to him in the social contract currently serving a society's *status quo*, is therefore highly relevant to how his worthiness is assessed by those round him.

The arguments offered in *Just Playing* suggesting that worthiness *increases* with need and ability and *declines* with effort and social standing do not find ready acceptance. The problem is partly one of separating the meanings of concepts that are usually bundled together. For example, I assume that a needy person lacks something important to him, but not that he is necessarily powerless to get it. Nor do I assume that the effort a person contributes to a project provides any guarantee that he has any power over the division of the surplus. A second problem lies in the fact that worthiness is not the only factor that affects who gets what. For example, although social standing decreases worthiness, it is not a handicap, since the fact that the high-and-mighty are expected to take a smaller share is more than compensated for by their being favoured in the *status quo* that serves to determine what counts as being divisible.

However, the major problem that critics have with my comparative statics lies in their unwillingness to distinguish between their utopian ideas about the way that they feel worthiness *ought* to be evaluated, and the manner in which worthiness *is* evaluated by real people. Utopians find it particularly upsetting that power should be relevant to the latter calculation. Nor are they mollified when it is explained that modern social contracts distribute power in complex and surprising ways. They are insistent that fairness should be determined *independently* of the balance of power. But I am the kind of practical democrat who believes that it is a waste of time to propose reforms that real voters are going to reject as unfair.

VII. SOCIALISM VERSUS CAPITALISM

Social standing is a tricky concept with complex social dynamics, but the dynamic effects of the other factors that influence worthiness are

much more straightforward. Their study serves to sharpen the consensus view that old-style socialism stifles enterprise and initiative.

A utilitarian society neglects the needy because their willingness to take risks makes them able to look after themselves. It handicaps the able because they don't have to work so hard as others to produce the same output. Entrepreneurs who are willing to take risks and investors who are willing to sink money into new enterprises are punished for the same reasons. Getting more education than others amounts to sinking an investment in oneself, and so scholarship is similarly disadvantaged. More generally, utilitarianism frowns on anything whatever that *empowers* people to get on in the world under their own steam. Self-help, thrift, ingenuity and enterprise are discouraged by steering resources away from those who display such whiggish virtues to unthinking work-horses like Boxer in Orwell's *Animal Farm*.

If I were offered the role of an omnipotent philosopher-king in a *static* society and were confident of my ability to resist the corruption traditionally associated with the exercise of absolute power, I would impose a utilitarian constitution on the society under my control, for the reasons given by Harsanyi. But we cannot count on those who offer themselves as philosopher-kings to possess either the power or the inclination to enforce the commitments that we would make behind Harsanyi's veil of ignorance. Nor does utilitarianism offer the right incentives to a society whose cultural survival depends on how well it competes on economic terms with other societies in *changing* world.

Rather than belabouring the fall of the Soviet Empire and the retreat from social democracy elsewhere, let me offer a slightly doctored extract from the doggerel poem in which de Mandeville expresses the moral to be drawn from his fable of the *Grumbling Hive*, in which the bees destroy their society by actually adopting the utopian virtues that they had previously honoured only in the breach:¹⁶

So Vice is beneficial found,
When it's by Justice lopt and bound;
Nay, where the People would be great,
As necessary to the State
As Hunger is to clear a plate.

Modern utilitarians commonly accept de Mandeville's prophetic observation that socialism cannot compete with capitalism in the creation of economic wealth, but they see no reason to give any ground to egalitarians on this count. Why should the economic success of capitalism be any more relevant to the Right or the Good than the fact

¹⁶ B. de Mandeville, *The Fable of the Bees – or Private Vices, Publick Benefits*, ed. F. Kaye, Indianapolis, 1988. (First published 1714.)

that money can be made by mugging tourists in back alleys? In this at least, postwelfarist egalitarians agree. After all, the Rawlsian program of maximizing some objective measure of the well-being of the least successful class of society is not much better at providing incentives for wealth creation than utilitarianism.

I also share the view that wealth creation is irrelevant to the battle between the Right and the Good, but for the very different reason that such debates over invented Moral Absolutes have nothing to do with the real world. Morality evolved in the human race to co-ordinate human behaviour on Pareto-improving equilibria in the Game of Life. The versions of egalitarianism and utilitarianism being compared in this paper respect this insight of David Hume, and hence represent alternative approaches to the problem of how a society can or could realistically reform its social contract. They remain relevant to much of the traditional debate between egalitarians and utilitarians because absolutist intuitions about the Right and the Good actually derive from seeing genuinely workable social contracts in action. But I have to part company with moral absolutists when they insist that Morality sometimes demands that opportunities to share a surplus should be neglected. At this point, absolutists lose contact with the intuitive understanding of the workings of real social contracts that keeps them on the right track much of the time. Fairness norms evolved as a way of co-ordinating among the *Pareto-efficient* equilibria of the Game of Life. When properly reformulated on a Humean basis, egalitarianism respects this function, but utilitarianism does not.¹⁷

else@ucl.ac.uk

¹⁷ The paper is a compilation from my book *Just Playing: Game Theory and the Social Contract II*, to be published in 1998 by MIT Press. I am grateful to the Economic and Social Research Council and to the Leverhulme Foundation for funding the research through the Centre for Economic Learning and Social Evolution at University College London.