On the problematic link between fundamental ethics and economic policy recommendations

Olof Johansson-Stenman

Abstract This paper provides a systematic survey of major simplifying assumptions that economists make, and often have to make, in order to obtain a useful theory for policy recommendations in practice. The aim is to consider the whole chain of assumptions with an emphasis on such simplifications that economists sometimes tend to ignore (at worst), or at best often tend not to take very seriously. The paper concludes that the link from fundamental ethics to economic policy recommendations is often very fragile, but that this is neither a convincing argument for economists to ignore normative issues, nor a reason to pretend that policy recommendations can be derived without value judgements. In order to improve the link it is recommended that welfare economists sometimes move beyond reductionist individualism and learn more from other social sciences regarding preference formation, decision behaviour, and the creation of values, institutions and social norms.

Keywords: ethics, moral philosophy, normative economics, welfare economics, policy recommendations, simplifying assumptions

1 INTRODUCTION: POSITIVE VERSUS NORMATIVE ECONOMICS

Policy recommendations deal with the normative issue of how one (e.g. the government) should act, and not just with how people and other agents actually do act, or why they act as they do. This difference is fundamental, although it is sometimes argued, rightly, that there is no objective value-free scientific theory either, since analysis has to be undertaken by a scientist who can never completely be independent of his or her experience, the cultural and moral climate in which they live, and so on. As expressed by Myrdal (1969: 74):

no social science or particular branch of social research can ever be ‘neutral’ or simply ‘factual’, indeed not ‘objective’ in the traditional meaning of these terms. Research is always and by logical necessity based on moral and political valuations, and the researcher should be obliged to account for them explicitly.

Furthermore, as noted by Solow (1994: 243): ‘Where powerful interests are at
stake, some research will be consciously or unconsciously perverted, and the
critical mechanism will be diverted or dulled. Still, there is a fundamental
difference between positive and normative analyses, since the expressed aim
of positive analysis is to be objective and value-neutral, even if this aim can
never be fulfilled. In the words of Friedman (1994: 181):

Positive economics is in principle independent of any particular ethical
position or normative judgments ... In short, positive economics is, or
can be, an 'objective' science, in precisely the same sense as any of the
physical sciences. Of course, the fact that economics deals with the
interrelations of human beings, and that the investigator is himself part of
the subject matter being investigated in a more intimate sense than in the
physical sciences, raises special difficulties in achieving objectivity at the
same time that it provides the social sciences with a class of data not
available to the physical scientist. But neither the one nor the other is, in
my view, a fundamental distinction between the two groups of sciences.
(emphasis added)

For example, most people seem to think that economists (and other social
scientists) should try to evaluate a proposed tax-reform as objectively as
possible. What will the consequences be for the environment, employment,
equity, national income, terms of trade, etc.? We know that this ideal
objectivity is often far from fulfilled, and that it is never completely fulfilled,
but most people probably agree that it is a reasonable goal to strive for. Still, it
may sometimes be advisable for social scientists to follow Myrdal's advice
and explicitly account for their moral and political values, since economists'
values and 'ideologies' may strongly influence analysis which is normally
characterized as positive and descriptive.

On the contrary, normative analysis can by definition never be value-free.
The normative question: 'how should the government reform its tax system?'
can never be answered without a set of value judgements. Trying to answer
this question solely on the basis of descriptions of how the society works, and
why people and other economic agents act as they do, would be an attempt to
derive an 'ought' from an 'is'. As clearly expressed by Nath (1969: 2):

No prescriptions can ever be made without starting explicitly or
implicitly with some notion of what is desirable, good, or, simply, the
social objective. Again, no objective can be formulated, or questions of
desirability settled, without implicitly or explicitly assuming some value
judgments.

This is of course not to say that normative economics should be less objective
than positive economics in the sense of replicability. Given the ethical
premises, the analysis should ideally end up with the same conclusions
irrespective of the values held by the researcher, in exactly the same way as
these values should ideally not affect the outcome of positive analysis.
(Bromley 1992). The choice of ethical premises has often been considered to be the choice of policy makers, and hence beyond the scope of economists. This view, although still common, is, however, rarely possible to maintain in practice; see e.g. Hausman and McPherson (1993, 1996). For example, it is not very common that policy makers express their inequality aversion in quantitative terms. In any case, normative analysis which leads to policy conclusions regarding what the government should or should not do involves ethical values, irrespective of whether these values are held by the researcher, the policy maker, or someone else.

The argument that one cannot derive an ‘ought’ from an ‘is’ was made famous by David Hume in the eighteenth century, but it is still a controversial statement. For example, Dasgupta (1993: 7) writes: ‘The “descriptive” and “evaluative” components of concepts like destitution cannot be separated. They are entangled’ (cf. Putnam 1993). Another even more problematic example is the proposition ‘it is a sin to kill’. This, in a sense factual proposition, seems clearly to imply the normative proposition ‘one should not kill’. The difficulty clearly arises for semantic reasons since the word ‘sin’ is defined with the use of values, e.g. as something we should not do. A similar example which may be more important in economics, but is perhaps less obvious, is the statement: ‘the social costs of reducing greenhouse gas emissions further exceed the benefits’. This may also seem to be a factual proposition which may be either correct or incorrect. However, even disregarding the tremendous uncertainties and economic interests involved in this issue, it is still a value-laden proposition in the sense that the costs are derived from an implicit or explicit social objective, which (by definition) is ethically based. Thus, the separation between intrinsically value-laden and purely empirical issues is not always straightforward; see also Blaug (1992: chapter 5).

Another reason why the separation between factual and normative propositions is often criticized is that many social scientists argue that economics as an academic discipline will never be value-free, or a solely positive science. But, as argued by Weston (1994), even if this certainly is true, it is not a valid argument for not retaining the positive/normative distinction in economics, since there are no valid reasons to consider normative issues to be less relevant for the academic discipline economics. Even though the language is never completely value-neutral, and we have to be aware of the fact that seemingly descriptive matters may sometimes be derived from value judgements, it seems in many cases to be both possible and useful to characterize an issue as either normative or positive, and either intrinsically value-laden or descriptive, in the senses discussed above.

This paper will focus on the normative and intrinsically value-laden part of economics and discuss the link from the most fundamental ethical premises step by step all the way towards policy recommendations. However, since many assumptions are needed on this way, there are of course many different
ways to end up with policy conclusions. This paper will focus on what may be seen as the main track in normative economics, including consequentialism, anthropocentrism, welfarism, utilitarianism and consumer sovereignty, and not follow up all other possible tracks. For example, it will not discuss which additional assumptions one has to make if one rejects welfarism and revealed preference economics. For more general recent surveys of ethics and economics, see Hausman and McPherson (1993, 1996), Hamlin (1996: introduction), Sen (1987) and Vickers (1997). For a thorough collection of papers on this subject, see Hamlin (1996).

In Section 2, we will discuss the ethical foundation of normative economics further, and discuss several common simplifying assumptions on the way towards a more ‘operational’ theory, in the sense that it is useful for policy purposes. Section 3 links the theory of consumer behaviour to these ethical foundations, and Section 4 briefly discusses the historical controversies on how to deal with normative issues in welfare economics. Section 5 draws conclusions.

2 THE ETHICAL FOUNDATION OF NORMATIVE ECONOMICS

The betterness relation and consequentialism

Questions related to what should be decided or done are of a moral nature, and the analysis must then be based on some ethical grounds. A fundamental ethical criterion is the betterness relation (Broome 1991a) that an action should be undertaken, and a decision taken, if they are better than every other action or decision, in the sense that they are more ‘good’ than any alternative. This is also often denoted a teleological or consequentialist theory as opposed to a deontological theory. Although this goodness-related ethical guideline for actions is indeed very general, and might seem obvious, it is not completely uncontroversial. For example, one may argue that some actions are ‘right’ or ‘wrong’ independently of how ‘good’ they are, based on rights-based ethics (e.g. Kantian duty-based moral). Dasgupta (1993: 27–8) remarks: ‘The key to deontological reasoning lies in a recognition of the priority of the right over the good. The hallmark of consequentialism is just the reverse: it acknowledges the priority of the good over the right.’ Economic theory normally assumes consequentialism, i.e. that what makes an action good or bad depends solely on how good or bad the consequences of the action are. However, it should be noted that consequentialism as such does not rule out that actions may have intrinsic values, since the action is a consequence in itself of the action; see e.g. Broome (1991a) and Scheffler (1982). Still, in economics one typically ignores intrinsic values in addition to consequentialism. Although there are no a priori arguments why this would follow from the betterness relation, many philosophers and economists seem
to agree (implicitly or explicitly) on the usefulness of this assumption.

A popular counter-argument, however, is based on whether it can be considered good to lie, or not. A consequentialist ethical theory which ignores intrinsic values implies that it would be justifiable to lie if the overall consequences (except from the action itself) were better than if one had told the truth. But with a partially deontological view there might be an intrinsic value in telling the truth. Hence, even if the consequences of lying would be slightly better than by telling the plain truth, the latter might be preferable in such a view. This argument obviously has some appeal, and most people probably do value truthfulness intrinsically, at least to some extent. However, most of this appeal may be superficial, due to the fact that there exist long-run dynamic effects from honesty and other social norms in a society. For example, if people can trust each other, obviously considerable control effort and costs are saved. Hence, there is no contradiction between consequentialism (in the stronger form which ignores intrinsic values) and arguments in favour of telling the truth also when the direct foreseeable effects are better than if one lies, due to possible long-run effects which may be difficult to identify. Hence, what may seemingly be intrinsic values may instead be valuations of indirect, dynamic and complicated consequences, see e.g. Ng (1981a, 1983). Still, to rule out non-consequentialist ethical theories, and that actions may have intrinsic values, are serious restrictions.

**Anthropocentrism**

A consequentialist ethical theory is not very ‘operational’ in itself, since we have no guidelines or criteria yet on what makes an action better than other actions. As we will see, there often seems to be a trade-off between the degree of ‘operationality’ and the ‘quality’ of the goodness measurement obtained. In order to make an ethical theory more operational or, loosely speaking, more practical and useful, some further simplifying assumptions are needed. For the theory to be able to say whether an action is better than another, we need to know which consequences should matter, and what makes a consequence good.

It is often (almost always) implicitly assumed in economics that it is solely the consequences for human beings that matter intrinsically. The consequences for animals and plants may also matter, but only indirectly through valuation by humans. For example, in environmental economics, ‘existence values’ of animals and plants are generally assumed to reflect individual human well-being measured as a willingness to pay for the environmental good (Freeman 1993; Mitchell and Carson 1989), but it is very rare that one tries to value nature intrinsically (e.g. based on some naturalist ethics), i.e. irrespective of human welfare. Many moral philosophers, such as Singer (1975, 1979, 1980), criticize anthropocentrism, in the sense that the goodness measure of an action would solely reflect the consequences (in terms
of welfare) for human beings, and argue that animals' welfare should in principle have the same weight as humans' welfare.\(^4\) Ng (1983) agrees in principle with Singer and others, but also notes that the application of such a theory 'may require man to sacrifice an enormous amount of his welfare' (1983: 165). He continues: 'Most people (myself included) are not prepared to sustain such sacrifices. What justification can we provide to reject man–animal parity? I can think of a pure and simple one – self-interest. To those who object that this is hardly an acceptable justification, I have to agree' (1983: 165). Furthermore, as shown by Spash and Hanley (1995), there is evidence that many individuals (including so-called 'deep ecologists', cf. O'Neill (1997)) think that nature should be valued intrinsically (at least to some extent). But is it really possible to talk about values in any meaningful way without considering human beings? Alternatively, should we value the welfare of other species also after the extinction of the human race? Norton (1987) proposes an intermediate way, entitled 'weak anthropocentrism', between anthropocentrism and 'ecocentrism'. An object can then have a value if and only if there is a valuer. But given that there is a valuer, the object may have an intrinsic value; cf. Aldred (1994).

If we still adopt an anthropocentric perspective (in the stronger sense) a natural question follows: which human beings? The ethically most-justifiable answer would perhaps be: all individuals, born and unborn. But, obviously, the task of valuing the consequences for people living 10,000 years from now would most often be rather problematic. Therefore, one often limits analysis to the currently living population. The consequences for future generations may be reflected to some extent, but only through the valuation of the current generation (as for animals and plants). For comparison of some actions, such as the choice between different unemployment insurance systems, this is not very controversial. But the issue becomes more complicated when dealing with actions which may cause long-term environmental effects, such as emissions of greenhouse gases; see e.g. Spash (1993). In such cases it becomes obvious that one has to undertake some kind of explicit intertemporal analysis, and include also the welfare of future generations. However, it is still common to count the welfare of future generations less than the welfare of the current generation, primarily through the choice of discount rate. This choice appears to be a typical example where (some) economists try to be value-neutral by simply choosing the actual observable interest rate (or some average of existing interest rates which are considered relevant) and applying this in a cost–benefit analysis; see e.g. Nordhaus (1991, 1994). But this is of course an illusion since the discount rate is a calculation instrument which should be derived from the social objective. Consequently, the common practice of discounting utilities (and not just consumption) through the so-called 'pure rate of time preference', without explicitly stating that the welfare of future generations should count less than that of the current generation, has been criticized for a long time by many authors including Ramsey (1928) and
Harrod (1948); see Broome (1992) and Azar (1995) for recent treatments of this issue in relation to the greenhouse effect. Hence, one cannot observe from actual behaviour which interest rate one should apply in cost–benefit analysis in general, and in particular not in cases which have consequences for hundreds or even thousands of years. There is also an ethical problem associated with the simple fact that it is this generation which decides how the welfare of different future generations should count; weighing these non-uniformly may be seen as a form of ‘deep’ anthropocentrism. In a similar way, but in another dimension, one often ignores (or puts less weight on) consequences for people in other countries, or in other areas.

Furthermore, some decisions may not only affect a large number of individuals, but they may also affect the number of people (living in different time periods or in different areas). Blackorby and Donaldson (1984, 1991), Broome (1996a, b), Hammond (1988) and Ng (1983, 1986) provide some possible criteria for evaluating such population changes, but due to the obvious ethical difficulties and complexity of this issue, it is neglected in most applied analysis.

**Welfarism**

A further restriction is that the overall goodness of different actions is often assumed to be reflected by an individualistic social welfare measure, which we call a social welfare function (SWF). In western societies of today such an assumption may seem natural and straightforward, but in a large part of the world this is not at all uncontroversial. As recently expressed by Chow (1997: 324):

> The society is more than a collection of individuals. Hence the welfare of the society is more than the sum of the welfares of its individual members. People in many developing countries are striving for nationalism and may consider the common good and national unity more important than individual rights.

Furthermore, as discussed by Thurow (1996), individualism of the western type is a very recent phenomenon in human history.

An SWF which depends solely on individual utilities is what Sen (1979) denotes a welfaristic social welfare function. Utility is here taken to be a measure in broad terms of individual subjective welfare or well-being; see e.g. Haslett (1990). The social welfare W can then, in a society with n individuals, be written:

\[ W = W(u^1, u^2, \ldots, u^n) \]  

where \( u^i \) is the utility of individual \( i \), and where social welfare \( W \) is normally assumed to be an increasing function in all its arguments \( u \). Hence, at this stage we do not rule out that an individual \( k \)’s utility depends on individual \( j \)’s
consumption. Sen has frequently argued for a broader ‘information base’
for the social objective or SWF, and his focus on capabilities and
functionings (Sen 1985, 1992, 1993) can be seen as such attempts. Ng
(1981a) provides a powerful defence of welfarism and argues that most (if
not all) alternative arguments (such as a fundamental human right of not
being tortured) in the SWF are unnecessary since these arguments are
basically motivated by welfarist reasons. He asks rhetorically: ‘Why do
we hear about “human rights”, “animal rights”, etc., but no “stone rights”? An obvious answer is that stones do not feel pleasure and pain’ (Ng 1981a:
530). Sen (1981) in a reply agrees that welfarist arguments may be
important for the creation of human rights and other social rules, but he
still argues that these welfarist arguments are far from sufficient: ‘Cases of
exploitation, sexual discrimination, racial asymmetry, etc., are standardly
criticised on grounds that do not depend exclusively (or perhaps at all) on
utility considerations. Traditional welfare economics is not robust enough to
accommodate this type of values’ (Sen 1981: 532). However, given a
sufficiently broad definition of utility, it appears possible to capture some, if
not all, of these features in a welfaristic framework, cf. Hammond (1991). One
could simply construct a function which depends on happiness as well as
various other elements which one would think contribute intrinsically (and not
only instrumentally) to a person’s well-being, and then call this measure
‘utility’. Broome (1991b, c) argues against such broad interpretation of
utility to mean something rather close to good: ‘it is perfectly redundant. We already have an excellent word with the meaning of good:
“Good”’ (Broome 1991b: 11). Instead, Broome argues that the notion of
utility should be reserved for a representation of preferences. However, Sen
(1991) in a comment argues against such an abandoning since the term
‘utility’ is used, and has for a long time been used, with many different
meanings and, even if it would be better if this were not so, the important
point is to be as clear as possible about which meaning one is currently
attaching to the term. In particular, it is important not to use the term with
different meanings simultaneously.

Since it is common practice to focus on utilities as the only arguments in the
SWF, it appears reasonable to include as much as possible which corresponds
to a person’s well-being in the utility function for the social welfare function to
be a measure which corresponds as closely as possible to an ideal goodness
measure. Still, as frequently pointed out by Sen, there may be other elements
of ethical relevance which are not covered by the concept of well-being (as it is
most often interpreted). However, since it is difficult enough to consider
individual subjective well-being, one typically ignores other possible
elements of what constitutes a person’s interest.

If we still accept an individualistic welfaristic SWF, we can deduce the
Paretian concept that an increase in the utility of one individual, without
decreasing the utility of any other individual, unambiguously increases social
welfare. If a welfaristic SWF is our measure of goodness, the action should correspondingly be undertaken. Note that the same cannot yet be said about an increase in consumption for one individual, even if utility is increasing in the individuals’ own consumption, since utility may decrease as a function of others’ consumption.

The Bergson–Samuelson social welfare function

The SWF is often assumed to fulfil the more restrictive properties of a Bergson–Samuelson social welfare function (B-S SWF). This means that social welfare is a function of individual utilities, which in turn depend (positively) solely on their own consumption of different goods. In addition, the utilities can depend on many other variables which, however, are assumed to be independent of consumption. The SWF can then be written:

\[ W = W(u^1(x^1), u^2(x^2), \ldots, u^n(x^n)) \] (2)

where \( x^i \) is \( i \)'s consumption vector of different goods. This may perhaps look like a rather general and uncontroversial way of writing what the government should try to maximize. However, the weak separability property of (2) is crucial both for positive and for normative analysis. The separability implies that individual \( i \)'s utility from consuming different goods is independent of individual \( j \)'s consumption. However, we can still not say that the individual choice between different goods would not be affected, because we have said nothing yet about the connection between individual utility and choice of consumption. The operationality of the theory increases drastically with this assumption. For example, we are now able to propose that Pareto improvements in terms of consumption should be undertaken, provided that income is increasing in consumption. Thus, if the consumption of a specific good increases for one individual and is unchanged for all other individuals, social welfare increases. Unfortunately, the degree of correspondence with the original goodness measure has correspondingly decreased. Indeed, this may be one of the weakest links on our way from the original betterness relation to actual policy conclusions.

Externalities, such as air pollution and noise, constitute obvious examples where the theory is too restrictive. But even if we disregard conventional externalities, the theory has been restricted to a large extent. For example, it precludes the influence of relative consumption on individual well-being. In a widely quoted paper, Easterlin (1974) surveyed several studies which tried to measure how happy people were in different countries, in different time periods, and in different circumstances. The result for a specific country in a single year supports the view of conventional economic welfare theory: richer people tend, on average, to be happier than poorer people. However, if one compares the result in one country for different years, people seem, on average, to be about equally happy over time, even though income has
increased dramatically during the same period. The same holds when comparing the degree of happiness in different countries, i.e. people seem to be about equally happy in different countries, although the average income in those countries may differ dramatically. Abramovitz (1979) has called this result the 'Easterlin paradox'. The most obvious explanation is that people's well-being, or happiness, depends to a large extent on their relative income, i.e. their own income compared to others' income, as proposed by Duesenberry (1949). This is discussed extensively by Abramovitz (1979), Frank (1985a, b), Layard (1980) and Scitovsky (1992), among others. In a recent paper Easterlin (1995: 35) concludes: 'Will raising the incomes of all increase the happiness of all? The answer to this question can now be given with somewhat greater assurance than twenty years ago... It is "no".'

A weak point in these empirical happiness studies is the link from stated happiness in various kinds of surveys and actual happiness. As observed by Brekke (1997) and Osmani (1993), it may be that people answer these types of questions in relation to what they consider to be an average norm of happiness, and that this norm may depend on income too. If so, happiness may depend primarily on the absolute level of income even though this is not reflected by the answers in surveys.

Still, psychologists have often been aware of these problems and tried to compare with alternative measures such as asking friends of the persons how happy they seem to be. Furthermore, as Frank (1985a) has observed, the hypothesis that relative income matters appears to be more compatible with evolutionary biology than the more conventional view that utility depends solely on the absolute level of income. This is because in an evolutionary perspective, the probability (and extent) of reproduction has historically generally been positively related to the control of resources. For this reason, as evolutionary biologists such as Alexander (1987) explain, our utility is often positively related to income or consumption in a given context. Alternatively speaking, there is evolutionary value in controlling more resources than do others in the neighbourhood. According to Frank (1985a) it seems most reasonable to assume that individual subjective well-being depends positively both on absolute income (particularly at low income levels) and relative income. See also the recent policy forum in the Economic Journal on economics and happiness (Dixon 1997; Frank 1997; Ng 1997; Oswald 1997).

A related issue, introduced by Hirsch (1977) and discussed by Frank (1985a, b), is the existence of so-called 'positional goods', in which the relative component is more important than for other (non-positional) goods. For example, the utility derived from a new car or jewellery may depend strongly on others' consumption of similar goods, whereas the utility derived from consumption of leisure may depend less on the amount of others' consumption of leisure. This may have consequences, e.g. in the derivation of optimal taxes as well as for the appropriate provision of public goods. For
example, Ng (1987) showed that pure public (and hence non-positional) goods should in general be over-provided relative to the basic Samuelson (1954) rule (EMRS = MRT) in the presence of such relative income effects.

Furthermore, a B–S SWF implicitly neglects all kinds of altruism, benevolence or malevolence. Boulding (1969) remarks on this assumption:

The plain fact is that our lives are dominated by precisely this interdependence of utility functions which the Paretian optimum denies. Selfishness, or indifference to the welfare of others, is a knife edge between benevolence on the one side and malevolence on the other. It is something that is very rare. [. . .] We either rejoice when they rejoice, or we rejoice when they mourn.

(1969: 6)

However, in the last decades there has been an increased interest in economics on interdependent utility functions, and in particular on economics and altruism.10

Utilitarianism

Since the functional form of the SWF is not yet defined, we can still make no general judgement of what to do in situations where utility increases for some individuals but decreases for others. This is obviously a very serious limitation, since almost all public decisions have such consequences. One way to deal with this problem is to assume that social welfare can be represented by a weighted sum of individual cardinal utilities, what is sometimes called ‘generalized utilitarianism’, which is of course a further restriction compared to the more general B–S SWF.11 This procedure is proposed by Diamond (1967) among others, and the SWF can then be written:

\[ W = \sum_{i=1}^{n} \alpha_i u_i(x_i) \]  
(3)

where the \( \alpha \)'s are welfare weights which can reflect the government’s distributional concern in terms of utilities. For example, social welfare may be higher for an equal distribution of utilities compared to the same amount of total utility more unequally distributed. In addition, distributional concerns can also be reflected by the concavity of the utility functions themselves, i.e. through the fact that a unit rise in income may increase utility (or well-being) more for a poor individual than for a rich.

Rawls’s (1973) theory of justice is often considered to be incompatible with utilitarianism (even by himself), although it, interpreted in a very narrow sense, is sometimes seen as a special case of (3), where the weights are zero for all but the most unfortunate member of the society; see e.g. Atkinson and Stiglitz (1980).12
Harsanyi (1975), Ng (1981b) and more recently Kaplow (1995), have criticized this kind of weighted utilitarianism for being inconsistent with the Pareto principle in the presence of uncertainty (as opposed to unweighted utilitarianism). Their main argument, which follows Harsanyi (1955), can easily be illustrated as follows. Consider a society consisting of two individuals, Allan and Boris, who are assumed to maximize their own expected utilities. The government can choose between two feasible alternatives, I and II. In state I Allan receives 10 utility units and Boris 5 utility units with probability 0.5, and vice versa, i.e. Allan receives 5 utility units and Boris 10 utility units with probability 0.5. In state II Allan and Boris receive 7 utility units each. Should the government choose I or II? According to the Pareto principle the government should choose I because both prefer I over II, since both Allan’s and Boris’s expected utilities are larger in I compared to II. Still, if the SWF is according to (3), it would for sufficiently large inequality aversion choose II, and hence violate the Pareto principle.\footnote{Sen (1976, 1977a, 1986) has frequently objected that Harsanyi’s theorems are not really about utilitarianism, since the utilities are not defined independently of the individual choice; see also the rejoinders by Harsanyi (1975, 1977), and Weymark (1991) who largely seems to support the view of Sen.} Sen (1976, 1977a, 1986) has frequently objected that Harsanyi’s theorems are not really about utilitarianism, since the utilities are not defined independently of the individual choice; see also the rejoinders by Harsanyi (1975, 1977), and Weymark (1991) who largely seems to support the view of Sen.\footnote{The classical unweighted utilitarian SWF in the Benthamian tradition, supported by Harsanyi and others, is given by:}

\[
W = \sum_{i=1}^{n} u^i(x^i)
\]

Both of these SWFs are frequently used in economic analysis. A common critique of utilitarianism deals with the fact that some individuals may have a much higher marginal utility of consumption than other individuals with the same income (Friedman 1947; Nozick 1974).\footnote{Consequently, some people’s well-being may increase more from an additional unit of income than may others’. Consider two individuals with the same income. Al is a modest stay-at-home kind of guy who likes to go fishing in his local pond more than anything else, and whose utility does not vary much with his consumption. His utility would not increase much with extra income. Bob is a party-guy who loves spending money and hates not to. His utility would increase rapidly with an increase in income. Everything else kept constant, would it then be a good thing to redistribute income from Al to Bob? Most people would probably not think so, even though it would increase our measure of social welfare. Then we clearly have a problem, since there seems to be a contradiction between people’s intuitive perception of what society should do, on the one hand, and our definition of social welfare on the other.} Consequently, some people’s well-being may increase more from an additional unit of income than may others’. Consider two individuals with the same income. Al is a modest stay-at-home kind of guy who likes to go fishing in his local pond more than anything else, and whose utility does not vary much with his consumption. His utility would not increase much with extra income. Bob is a party-guy who loves spending money and hates not to. His utility would increase rapidly with an increase in income. Everything else kept constant, would it then be a good thing to redistribute income from Al to Bob? Most people would probably not think so, even though it would increase our measure of social welfare. Then we clearly have a problem, since there seems to be a contradiction between people’s intuitive perception of what society should do, on the one hand, and our definition of social welfare on the other.\footnote{However, in practice this seems not to be a major problem in applied economics, because these kinds of individual differences are generally not easily measurable at the necessary disaggregated level at which the economic analysis is undertaken; cf. Lerner (1944) and Sen (1973: 81–5). Hence, in
applied economic studies one typically either assumes an identical utility function for all individuals, or imposes some specific restrictions such as equal marginal utility of consumption for equal income. Consequently, differences in marginal utility of income for people with identical incomes do not seem to play any role for actual policy recommendations. This illustrates the fact that, sometimes, the quality of the goodness measure may actually increase with additional restrictions.17

3 CONSUMER BEHAVIOUR

Consumer sovereignty and revealed preferences

So far, nothing has been said about the relation between individuals’ well-being and how they act. In economics, the dominating doctrine is simply that individuals do what is best for themselves. Such an assumption simplifies the analysis tremendously. We would then simply know that, for a given amount of resources, individuals would maximize their own well-being. Alternatively speaking, individuals’ preferences are assumed to be revealed by their actions, following Samuelson (1938). With this assumption, together with equation (2), we are able to say that the individual’s consumption choice is independent of other individuals’ consumption (provided that prices will not be affected by their consumption). This assumption is crucial in most normative economic analysis, such as optimal-taxation literature. For example, it implies that an individual is always better off (or equally well off) from paying a certain amount of money in a lump-sum manner compared to income or consumption taxes.

Although this may often be a very well-motivated simplification, it is of course not in any absolute sense correct. As expressed by Sen (1987: 11): ‘The coolly rational types may fill our textbooks, but the world is richer.’ There are several reasons for this, including obvious cognitive limitations; see Conlisk (1996) for a recent survey on bounded rationality. Scitovsky (1974) provides several examples based on ‘money illusion’, ‘effort illusion’ and ‘time illusion’ in a paper entitled ‘Are men rational or economists wrong?’. Furthermore, psychologists have presented much evidence that expected utility theory is often unsatisfying as a descriptive model under risk; see e.g. Tversky and Kahneman (1974) and Kahneman and Tversky (1979), who show that people instead often tend to apply some simplistic heuristic decision criterion. It is also well documented that individuals often tend to overestimate small probabilities and underestimate large probabilities (Thaler 1992; Viscusi 1992).

We have already noted that individuals may be altruistic and derive utility or well-being from others’ utility or consumption. In addition, as pointed out by Sen (1977b), individuals may take others’ well-being into account through some kind of commitment or social norms without
deriving any utility (or well-being) from this choice. In economics, this type of assumption, where utility maximization cannot fully describe individual behaviour, is still quite rare.\textsuperscript{18} Sen (1977b) denoted people who are always selfishly rational as 'rational fools,' since such behaviour will often imply a sub-optimal outcome for all involved. As argued by Thaler (1992: 20) after reviewing much experimental evidence that people often choose the cooperative alternative, even in situations where economic theory would predict them not to: 'Perhaps we need to give more attention to "sensible cooperators"'.

**Exogenous tastes**

Implicit in (2) we have that the utility functions *per se* are exogenously given and will not change with regard to consumption of various goods, i.e. preferences or tastes are assumed to be fixed. This common assumption has been questioned by Galbraith (1991), among others. For example, he argues (1991: 129) that the 'central function [of modern advertising and salesmanship] is to create desires – to bring into being wants that previously did not exist'. In a formalization of similar ideas, Dixit and Norman (1978) argue that the amount of advertisement seems to be socially too large under a quite general set of assumptions.

In their famous paper 'De gustibus non est disputandum', Stigler and Becker (1977) argue, on the contrary, that it is not necessary, or useful, to explain the obvious influence of advertisement on behaviour as change in tastes.\textsuperscript{19} Instead, they argue that the changed consumer behaviour can be explained by the changed information (true or false) that the consumers receive, and by changes in the 'technology' through a special kind of household production function in order to produce 'commodities' such as prestige.\textsuperscript{20} However, as explained by Pollak (1978), these arguments may be useful in the *explanation* of consumer behaviour, but the analysis needed is much more complicated in *normative* analysis, e.g. when we are dealing with the welfare effects of advertisement. The reason is of course that if utility is seen as a measure of what the government should maximize, it is important how utility is defined. Consequently, if the relevant measure for what the government should maximize is changing, the utility measure also needs to change. If, on the other hand, utility is not primarily seen as a measure for what the government should maximize, then we need to define a new independent concept for this purpose (e.g. individual welfare or well-being), since the maximization of an SWF that solely depends on utilities would not follow from any (reasonable) ethical principles.

Another economist who has frequently questioned the assumption of exogenously given tastes is Kenneth Boulding. Several years before the Becker and Stigler article he wrote (Boulding 1969: 1–2): ‘The most absurd of all pieces of ancient wisdom is surely the Latin tag *de gustibus non*
disputandum. In fact, we spend most of our lives disputing about tastes.' He then goes on exploring this:

One of the most peculiar illusions of economists is a doctrine that might be called the Immaculate Conception of the Indifference Curve, that is, that tastes are simply given, and that we cannot inquire into the process by which they are formed. This doctrine is literally 'for the birds', whose tastes are largely created for them by their genetic structures, and can therefore be treated as a constant in the dynamics of bird societies. In human society, however, the genetic component of tastes is very small indeed.

Akerlof and Dickens (1982) and Elster (1982) provide other examples of preference formation, as a result of cognitive dissonance and similar arguments, which may have important policy-relevant welfare effects; see also Veblen (1899) for an early and powerful discussion.

Maybe a change in attitude towards the assumption of exogenous tastes is under way. Recently, in his AEA presidential address, Victor Fuchs (1996: 16) stated: 'I believe there is an analogy between the economics of values and the economics of technology. Over the past several decades some economists have begun to treat technology as endogenous. Now, a similar effort must be undertaken for values.' Interestingly, one of the founders of the endogenous growth theory, Paul Romer, discusses in a new book (Romer 1999) the importance of changing preferences, both for explaining human behaviour and for policy recommendations.21

We have seen that the social norms may be important for individual behaviour and choice. But these norms, of course, are not constant over time. The sociologist Coleman's (1990) most fundamental critique of neoclassical economics seems to be the neglect of the formation of social norms and the corresponding build-up of a 'social capital'.22 In frequently discussed work, Putnam et al. (1993) and Helliwell and Putnam (1995) emphasize the importance of social capital to explain a large part of the big difference in economic development between southern and northern Italy.

There seems to have been an increased interest in the importance of social norms in economics for the last fifteen years or so. Elster (1989a, b), Frank (1988) and Sugden (1986, 1989) provide thorough explanations of why and how social norms occur. Lindbeck (1995, 1997) discusses the role of social norms as incentives in the labour market, Ostrom (1990) gives several empirical examples of successful local norms to overcome 'prisoner's dilemma' and 'tragedy of the common' situations in relation to natural resources management in a third world context, and Akerlof (1997) and Cole et al. (1998) discuss social norms and the existence of distinct social classes. Kreps (1997: 363) concludes a recent paper on the interaction between norms and economic incentives: 'The results are likely to be messy. They will...
involve activities unfamiliar to economics (e.g. theories of how preferences are formed and reformed). But messy or not, they are important and must be pursued.

**Well-behaved preferences**

In order to end up with an operational ethical theory we must typically impose some structure on individual preferences, such as reflexivity, transitivity and local non-satiation; see, for example, Varian (1984). However, although routinely made, these assumptions have in many cases been demonstrated not to hold. For example, Thaler (1980) presented evidence supporting the existence of an *endowment effect*, i.e. that people often demand much more in compensation to give up a good, than they would be willing to pay to get it, and that this seems to be the case also for small amounts of money where a marginal analysis could be argued to be a reasonable approximation. Kahneman and Tversky (1984) and Samuelson and Zeckhauser (1988) discuss the related concepts of *loss-aversion* and *status quo bias*, respectively. Slovic and Lichtenstein (1983) and Tversky *et al.* (1990) demonstrate and discuss *preference reversals*, that people in a choice between goods A and B prefer A, but that they still would be willing to pay more for good B than for good A.

**Chaos and evolution**

The contributions in evolutionary economics and chaos theory often focus on the dramatic outcome that may result from small changes in a crucial variable, due to strong non-linearities and unstable equilibria; see Boulding (1991) for an introduction to evolutionary economics and Kemp (1997) for a non-technical introduction to chaos and non-linear dynamic systems. But in a complex dynamic system such phenomena are often extremely difficult to predict before they occur. Consequently, economists tend to assume they will not occur, at least not due to a discussed policy measure. Again, stress of ecological systems due to human activities constitutes a perhaps obvious example when such assumptions may be doubtful. But, as history shows, human systems, cultures and political systems may also be very unstable, and relatively small changes in the environment may cause dramatic effects, such as a switch to a development path towards democracy instead of dictatorship (or, of course, the other way around). In a recent and thoughtful book Thurow (1996) discusses the rapidly increasing inequalities in the capitalist world, and in particular in the US, during the last two decades: 'How far can inequality widen and real wages fall before something snaps in a democracy? No one knows, since it has never before happened. The experiment has never been tried' (1996: 261). Cf. Ayres (1998) on the same issue.

Basically all decisions taken by a government at a national or local level will affect a large number of individuals, and it is almost always possible to find both people who gain and people who lose from political decisions. To deal with such issues, one has to trade what is good for some against what is bad for others. This was the natural conclusion drawn by classical economists and philosophers in the utilitarian tradition, such as Mill, Bentham and Edgeworth.

The revolution of the new welfare economics

However, in the twentieth century, this view has been considered 'unscientific' and impossible to test and become gradually less common. Marshall, for example, who was first a convinced utilitarian became more influenced by these 'new' ideas. In particular, ethics and moral values were considered to be unscientific, and should therefore be removed from economics. Robbins (1932: 148) argued that 'it does not seem logically possible to associate the two studies [economics and ethics] in any form but mere juxtaposition'. As noted by Sen (1987: 2), 'he was taking a position that was quite unfashionable then, though extremely fashionable now'.

At the core of this debate between the old welfare economics and the new welfare economics lies the concept of utility. As we have seen, the term 'utility' is used with many different meanings. There are at least two fundamentally different meanings and uses of the concept utility frequently used in economics; they are sometimes called the experience model and the preference model, see e.g. Kahneman and Varey (1991). The different meanings have often caused, and still cause, great confusion. Briefly, the experience model derives utility only through individual experience, i.e. through mental states or states of consciousness. Originally, in the Benthamian tradition, these mental states were taken to be pleasure or happiness. Later on, this has been modified to a much broader concept; see e.g. Hausman and McPherson (1994). Even though it is of course difficult to measure (or estimate) such a vague concept as utility per se, there do exist various psychological methods of measuring individual well-being; see e.g. Tversky and Griffin (1991) and Kahneman and Varey (1991).

The preference model instead assumes that what increases an individual's utility is simply what the individual prefers, which is revealed by their actions. In its purest form, no assumption is made of whether the individual will be happier or better off (defined independently of the actions). This model is of course in general much easier to work with empirically when it comes to purely explanatory models, if we accept revealed preferences, since we need then only to observe market transactions. According to the new welfare
economics, it is beyond the domain of economists to ask why an individual prefers a certain good or service to a certain extent.

The statement ‘Individuals will always maximize their own utility’ is then either wrong or a tautology. It is wrong if we define utility as a measure of well-being, as in the experience model, or as some other goodness-related measure defined independently of the action. As expressed by Hausman and McPherson (1994: 263): ‘What people prefer may not be good for them, because people make mistakes, and because they may prefer to sacrifice their own well-being in pursuit of some other end.’ On the other hand, it is a tautology if the maximization of utility is defined solely by the fulfilment of preferences, and the preferences are defined solely by the actual actions undertaken. The new welfare economics, which is sometimes considered to be a basic pillar of neoclassical theory, therefore avoids interpersonal utility comparisons, which are frequently said to be unnecessary, unscientific, useless, etc. Since interpersonal comparisons of utility were ruled out, the Pareto criterion became the only guidance for action. But since almost all decisions imply that some will gain and some will lose, the Pareto criterion provides of course little help in practice.

Some extreme Paretians go even further and argue that one (the government) should not do anything unless a Pareto improvement is achieved. This does not necessarily follow from the Pareto criterion itself, but if it is interpreted in this way it will obviously bias the decisions in a conservative direction. For example, Robinson (1962) pointed out that with this view, the only justifiable economic policy would be to do nothing. Furthermore, it may even be that any economic policy (including doing nothing) would be incompatible with the Pareto criterion, interpreted in this (rather extreme) way. Consider the following tax-collecting paradox which arises from the fact that we act in a historical context. The question ‘Should we collect any income taxes next year?’ is then indeed difficult to answer. We can with reasonable certainty assume that most people would prefer not to pay any income taxes themselves (even though most people would prefer others to pay taxes). For this reason, we should not decide to collect any income taxes for the next year (unless people would like to make any voluntary contributions). However, we do collect such taxes now, and many people, notably the poor and ill, would obviously be much worse off if we did not collect them. Hence, we could then not decide not to collect any income taxes for the next year either. Thus, we can neither collect taxes nor not collect them! Harrod (1938) pointed out early the difficulties for the new welfare economics to say anything about policy.

If the incomparability of utility to different individuals is strictly pressed, not only are the prescriptions of the welfare school ruled out, but all prescriptions whatever. The economist as an adviser is completely stultified, and, unless his speculations be regarded as of paramount aesthetic value, he had better be suppressed completely.

(1938: 397)
Consequently, the huge amount of social choice literature, which, following Arrow (1951), most often rules out interpersonal comparisons of utilities, seems to be largely focusing on negative results and has very little to say which has direct policy relevance; see e.g. Sen (1986) for a survey and Elster and Hylland (1986) for an extensive treatment of social choice literature. As noted by many, this has to do with the fact that the 'information base' on which the decision should be taken is so drastically reduced without interpersonal comparisons of utilities. According to Hammond (1991: 206): 'Indeed, after Arrow's original contribution, not much of this literature seems all that useful in retrospect.'

**Potential compensation criteria**

Since a situation where economists could not give any policy recommendations whatsoever was not considered very appealing, at least by the economists themselves, there was a large demand for some other criterion which would be more useful in practice and which was still 'scientific' and 'objective'. To some, the well-known potential compensation criteria proposed by Hicks (1939) and Kaldor (1939) were the answer to that demand, and they have frequently been associated with the new welfare economics. It is obvious that these criteria proved extremely useful, and they are generally considered to constitute the welfare-theoretic foundation of cost-benefit analysis. For example, in a recent textbook on environmental economics, Freeman (1993: 87) writes about the Hicks–Kaldor criterion:

> This is the efficiency criterion of the new welfare economics. According to the efficiency criterion, the objective of social policy is to maximize the aggregate value of all the goods and services people receive, including environmental and resource services.

Whether this is more 'scientific' and 'objective' than the old welfare economics is not at all clear, however. Samuelson (1950) showed that no policy recommendations could be derived on the basis of the Hicks–Kaldor criteria, or from the extension proposed by Scitovsky (1941). According to Boulding (1969: 8): 'Any decision involving other people obviously involves these interpersonal comparisons.' This is clear since if the losers are not compensated, and we still argue that a decision should be taken, what we are doing is exactly an interpersonal comparison of utilities: we are saying that, from a social perspective, the utility increase from those who gain is worth more than the utility loss from those who lose. Of course, this is by no means more objective, or less controversial, than classical utilitarianism. On the other hand, if we do compensate the losers 'then the overall outcome – after compensation – is a Pareto improvement, and then there is no need for a compensation test as a supplement to the Pareto principle. So the
compensation criteria are either unconvincing or redundant’ (Sen 1987: 33, footnote 4, italics in original).

However, although there is nothing ethically appealing about the compensation criteria per se, they may still be useful in normative economics as approximations to a utilitarian view (which, in turn, is an approximation of our original goodness measure). Under some conditions, such as relatively small distributional consequences, such an approximation may be justified.31

The silent counter-revolution

The new, new welfare economists, sometimes called the ‘new cardinalists’, have returned to the view that interpersonal comparisons of utility, however difficult these may be, are necessary for any policy recommendation regarding more than one individual. The classic paper on optimal income taxation by Mirrlees (1971) is generally considered to be the beginning of the new, new welfare economics; see e.g. Stiglitz (1985) and Arnott (1994). This ‘counter-revolution’ has typically been made in a rather ‘quiet’ and discrete manner, without much rhetoric, at least compared to the ‘revolution’ of the new welfare economics, even though there exist notable exceptions such as Blackorby (1990) and Ng (1997):

if ever we are to have something consistent and correct to say about policy in the real world, we shall have to learn how to make and make use of, interpersonal comparisons of well-being. This is an activity that is often dismissed as simply non-scientific or worse, impossible. The word ‘scientific’ is often invoked as an almost magical incantation. . . . The general view behind this position is that it is more scientific to use efficiency values rather than distributional values. To describe the choice of the former as scientific renders the word practically meaningless. In any case, my claim is that our choice is between making interpersonal comparisons of utility or in having little or nothing to say.

(Blackorby 1990: 749)

Furthermore, as argued by Davidson (1986), we often make interpersonal comparisons of well-being without finding this very difficult:

Some cases are hard, of course; the same can be said concerning some decisions which affect one person only. But on the whole we do not experience the problem of comparing the interests of different people as being harder in kind or degree than comparing the interests of our own. Naturally we are more apt to be ignorant of the interests of others than of our own; but this is a variable we have no trouble in accounting for.

(1986: 195)

A possible difference between the old welfare economics and the new, new welfare economics may be that the latter typically interprets the utility concept
in a broader way, and that it also often simultaneously deals with incentive
effects and distributional issues in a second-best general equilibrium
framework. The objective is then to maximize a (welfaristic) SWF given the
fact that individuals (and other economic agents) will maximize their
objective functions, and given that the government does not have all relevant
information, or access to all possible policy instruments. For example, in the
classic Mirrlees tax problem, the government could not observe individual
abilities to work or the amount of leisure consumed.

However, it should be made clear that although interpersonal comparisons
of utility are in general necessary for deriving policy conclusions, they are not
sufficient, since we still need some social choice rule for how we should act
given this utility information.\(^\text{32}\)

**Consumer surplus and ‘exact’ welfare measures**

In order to measure welfare changes in practice we must typically add a few
standard simplifications including perfectly competitive markets, zero
transaction and search costs, that all other prices (except the ones we focus on)
are non-distorted, and that firms can be characterized as profit-maximizers.
Although these assumptions are never strictly fulfilled, they may (or may not)
serve as useful approximations; cf. Blackorby (1990) and Rappaport (1996).

With *discrete* non-continuous changes, as most changes obviously are, the
most commonly used welfare measure is probably still the change in
Marshallian consumer surplus; i.e. the area between the demand curve and the
price line in a price–quantity diagram. Besides ignoring all distributional
effects, this measure has some other theoretical drawbacks arising from the
fact that individual marginal utility of income is not constant during this
discrete change (of prices or quantities). To deal with these problems, one can
instead use the well-known monetary welfare measures introduced by Hicks
(1943), i.e. the compensating variation (CV) and the equivalent variation (EV)
corresponding to the maximum willingness to pay for an improvement and the
minimum willingness to accept for not having an improvement, respectively;
see e.g. Varian (1984) or Freeman (1993) for definitions.\(^\text{33}\) Willig (1976)
derived bounds for the differences between the Hicksian and the Marshallian
welfare measures for price changes, and concluded that the differences are
most often negligible (if the income effect is not very large). Randall and Stoll
(1980) attempted a similar exercise for quantity changes (of a public good),
and Hanemann (1991) clarified their result and showed that the differences
will be larger if private consumption is far from being a perfect substitute for
the public good; see also Becht (1995) for a survey of related literature. The
current practice is to call these Hicksian measures *exact* welfare measures (see
e.g. Hausman 1981, or Becht 1995) in order to separate them from the
Marshallian consumer surplus, which seems, in the light of the simplifications
made above, somewhat misleading.\(^\text{34}\)
In modern applied welfare economics this problem, arising from differences in marginal utility of income within a single person (due to changes in prices or provision of public goods), is generally emphasized much more than the distributional effects arising from differences in marginal utility of income between different persons. This is so even though the variation of the marginal utility of income, for most reasonable SWFs, will be much larger between the individuals (at least for relatively small changes in prices or provision of public goods). Presumably, this is a result of the still-lasting strong influence of the new welfare economics; perhaps it also reflects a time-lag before an influential breakthrough of the new, new welfare economics in applied welfare economics.

Furthermore, so far we have discussed the problem of welfare change given perfect information. But, of course, in reality the uncertainties involved are often large. Although economists have begun to study imperfect information extensively during recent decades, it is still often neglected in applied studies in order to keep the complexity at a manageable level. Sometimes this may be a reasonable simplification; sometimes it is not. An example where this simplification seems highly inappropriate is cost-benefit analysis related to global warming and the greenhouse effect. Focusing solely on what is thought to be the most likely outcome may seriously underestimate expected costs, since, for example, the expected cost of an unlikely catastrophic event may be much larger than the cost of the most likely outcome. Although most economists seem to agree on this point (at least, few argue oppositely), most work in this area still ignores this problem.

5 CONCLUSIONS

We have seen that the link from fundamental ethics towards policy recommendations is often very weak and that it depends on several very strong assumptions. This is so even though the economics literature is becoming fairly rich in questioning these restrictive assumptions separately. However, in applied work it is unfortunately much more difficult to avoid these strong assumptions. As expressed by Goodwin (1991: 286):

Hence we have a relatively unscientific applied welfare economics, and a scientific theoretical welfare economics, with no better bridges between the two today than was to be found at the beginning of the century in Marshall’s system of inconsistent, or ambiguous, inclusiveness.

What is left of our original goodness measure when arriving at the policy conclusions after all simplifying assumptions is certainly no trivial question, and it is important to realize that the correspondence between the two may be quite poor. Still, this is hardly a good argument for economists not to deal with normative issues, since it is hard to see what could be gained from ignoring a systematic treatment of all the obviously important issues of what the
government should do when more than one single individual is affected. Furthermore, as history shows, economists will not give up the possibility of influencing political decisions and it is then far better if this influence is explicitly built on the only available solid ground, namely on some fundamental moral values which are then linked to economic theory, however loose this link may be, rather than if it is built on the sand of some imagined objective and value-neutral economic science.

Then there is an everlasting process for economists to improve this weak link by choosing more reasonable assumptions and improving our understanding of human behaviour. In this process, it is important to be as explicit as possible concerning basic value judgements and to realize that humans act in a social context, and hence that we may need to go beyond reductionist individualism. For example, a well-working economy needs well-working institutions, which is of course a far from new insight (see e.g. Veblen 1899; or Knight 1924). Nevertheless, it is neglected in much welfare-theoretic work, partly due to the obvious fact that modelling evolution and change of the institutional environment is quite difficult. Again, the seriousness of this neglect varies largely due to the circumstances. There is much evidence that the development of proper institutions is particularly important for the developing world and transitional economies, and hence that neglecting these issues here is often a very serious simplification.

Furthermore, as proposed by Akerlof (1984) and Hirshleifer (1985), we need to learn more from our colleagues in social science, regarding fundamental issues such as preference formation, decision behaviour, and the creation of values, institutions and social norms. Hirshleifer (1985: 52) argued that ‘good economics will also have to be good anthropology and sociology and political science and psychology’. Akerlof (1984: 6) ‘would like to think that psycho-socio-anthropo-economics is at the beginning of a period when many people will be working in this area. Thirty years ago mathematical economics was probably in a similar stage of development.’ Since these words by Akerlof were written, it is probably fair to say that other social sciences have actually influenced economic ideas substantially, but hardly sufficiently.

Finally, even though economics, like all (social) sciences, will always have to rely on a number of strong and stylized assumptions, there are a number of very general points which it might be wise to consider (again and again), even though they may seem trivial:

1. There is a trade-off between simplicity and realism (see e.g. Sen 1985). Sometimes the larger realism is worth the cost in terms of increased complexity, reflected in more difficult analysis and/or less rigour, fewer unambiguous results, and less obvious and straightforward policy conclusions.
2. There is a choice concerning which simplifying assumptions to make.
This choice should preferably be undertaken so as to maximize the usefulness, and degree of realism of the analysis, for a given level of complexity. Hence, it is important not to choose certain assumptions for dogmatic reasons.

It is important to consider how various simplifying assumptions affect the results and hence the policy conclusions. This seems particularly important for assumptions which are often routinely made without much deeper reflection.

There is certainly no reason for economists to stop dealing with normative issues and policy recommendations. But in doing so, perhaps sometimes a little more modesty would not hurt.

**ACKNOWLEDGEMENTS**

I have benefited largely from discussions with Daniel Bromley and from his comments on earlier versions. I have also received valuable comments from Ramon Lopez, Thomas Sterner, Rick Wicks, the editor and anonymous referees. Financial support from the Swedish Transport and Communications Research Board (KFB) is gratefully acknowledged. The usual disclaimer applies.

**NOTES**

1. The term 'operational' will be used with this specific meaning throughout the paper; see Blaug (1992: 87-91) for somewhat different meanings of this term.
2. Unfortunately, the definitions of deontological and teleological theories vary somewhat in the literature. Rawls (1973: 24) defines an ethical theory as teleological if 'the good is defined independent from the right, and then the right is defined as that which maximizes the good'. Howarth (1995) says that an action that is based on duty (a 'Kantian' motivation), instead of preference fulfilment, is 'deontologically motivated'. Etzioni (1988) seems to use 'deontology' in a similar way.
3. However, as pointed out by Hammond (1991: 201), when the space of consequences considered is extended to take everything of ethical significance into account, consequentialism becomes a tautology.
4. See Blackorby and Donaldson (1992) who apply the ethical view of Singer in an economic analysis of animal exploitation, and Ng (1995) for an attempt 'towards welfare biology'.
5. A possible solution to this problem would be to include everything of ethical relevance, and not just well-being, in the utility function; see Hammond (1996). Sen (1987: 40, footnote 13) argues against such very broad utility interpretation: 'As a defence of utility-based calculations this is tautologous and adds little to the discussion.'
6. This seems to be the most common interpretation of the notion B–S SWF. Unfortunately, sometimes different meanings of this term are used in the literature.
7. It is not very difficult to extend the analysis to take 'technological' externalities
into account, which is frequently done in neoclassical economic theory; see e.g. Baumol and Oates (1988).

9 Again, this is not to say that happiness and well-being are taken to be the same thing, but, as argued by Frank (1997), it seems reasonable that happiness is an important element in most people's well-being.

For relation to optimal taxation, see Boskin and Sheshinski (1978), Tuomala (1990) and Persson (1995). See also Kocherlakota (1996) for a recent discussion in relation to 'the equity premium puzzle', i.e. why the rate of return on stocks has been that much larger compared to that on bonds. He discusses a utility concept which depends on relative consumption in the time dimension, i.e. that utility today is a function of both current consumption and yesterday's consumption.


However, it should be noted that utilitarianism is not a subset of B–S SWF. For example, we may perfectly well think of utilitarianism where the utility functions are not independent of each other. Furthermore, utilitarianism can also incorporate animals; see e.g. Singer (1979). The order in which additional restrictions are imposed in this paper is thus not unambiguous, but rather made for illustrative purposes.

12 An often used SWF is characterized by constant social inequality aversion (in terms of utilities) so that \( W = \sum_i u_i^{1-\varphi}/(1 - \varphi) \), where \( \varphi \) is the social inequality aversion. (When the inequality aversion is equal to one, \( W = \sum_i \ln(u_i) \).) It is straightforward to see that this expression simplifies to classical utilitarianism when \( \varphi = 0 \) and to a so-called Rawlsian maxi-min SWF when \( \varphi \) goes to infinity.

Note that we have dealt with distribution of utilities, not consumption. For example, if both individuals have an identical utility function \( u = uc \) then in state I Allan receives 100 consumption units and Boris 25 units with probability 0.5, and vice versa with probability 0.5. In state II, both Allan and Boris receive 49 consumption units each.

See, however, Broome (1991a, 1996c) for a useful reinterpretation of the theorems by Harsanyi.

15 Alternatives to utilitarianism include the commonly discussed contract theories by Rawls (1973) and Nozick (1974).

16 Some utilitarians would argue that these kinds of expressed views or social norms do not reflect fundamental, but rather instrumental, values. For example, a situation where the difference in the marginal utility of income is estimated between different persons may per se result in individual suffering, and hence reduced overall utility.

17 Consider the Al-and-Bob example. Starting with the betterness relation, we assume that it is not 'good' to redistribute money from Al to Bob. But if we apply a utilitarian SWF (either weighted or unweighted) it would increase social welfare to do so. If, on the other hand, we impose an additional simplification that we will count their marginal utility of consumption as equal (if their incomes are equal) then a redistribution will not increase social welfare, which is in accordance with the betterness relation.

18 In the economics of altruism, this type of commitment is sometimes termed...
'genuine altruism' (Edwards 1992; Johansson 1997), as opposed to the more
common pure altruism (altruism with regard to others' utilities) and paternalistic
altruism (altruism with regard to some specific component of others' utilities)
where the altruistic concern is modelled inside the utility function.

19 They provide similar arguments for other cases which may seemingly be
incompatible with the assumption of exogenous and stable preferences, including
the addiction to drugs.

20 However, one needs to be careful when interpreting Becker and Stigler since they
use notions such as 'commodities' with a somewhat different meaning compared
to most economists. As expressed by Ackerman (1997: 661): 'The reader who
lacks an English-to-Becker dictionary must remember that what Becker calls
commodities are what others would call experiences or satisfactions, while the
commodities visible to the rest of us are, for Becker, inputs purchased by
households in order to produce commodities.' Furthermore, as noted by Akerlof
(1997), Becker has in later writings adopted a less rigid position towards
changing preferences.

21 The recent contribution by Hanemann and Kriström (1995) on preference-
uncertainty seems to be a step in that direction in environmental economics.

22 Frank (1992) provides a thorough review of this comprehensive work.


24 The revealed preference assumption is not useful for explaining the choice as
such, since the choice is simply assumed to reflect the preferences (Bromley
1989). However, the assumption is useful to explain preferences, such as whether
people seem to prefer white or brown bread.

25 Still, it may be a both useful and reasonable approximate assumption, but that is
of course a fundamental difference.

26 With the same type of 'logic' we can 'prove' that Bangladesh is a part of Africa
(which, of course, it is not). If Africa is defined as the continent which consists of
the poor countries of the earth, and Bangladesh is considered a poor country
(which it is), then Bangladesh must, by necessity, be a part of Africa!

27 The Pareto criterion, as it is most often interpreted, says that an action should be
undertaken if a Pareto improvement is achieved. It does not say that an action
should not be undertaken if a Pareto improvement is not achieved.

28 The criteria are defined somewhat differently by Hicks and by Kaldor. A common
description of these criteria is that an economic allocation \( X \) is superior to an
allocation \( Y \) if and only if it is possible to reach an allocation \( Z \) through
redistribution from \( X \) such that \( Z \) is preferred to \( Y \) according to the Pareto test. One
problem with this statement is that, when relative prices are affected by the
decision, it may be that two separate allocations can be potentially Pareto
dominated by each other; see e.g. Scitovsky (1941) or Samuelson (1950).

29 See Hubin (1994) for a recent critique of this view.

30 Little (1957) proposed an alternative modification to take distributional concern
into account. An action should then be undertaken if any of the Hicks–Kaldor
criteria and the Scitovsky criterion is fulfilled and at the same time the change in
income distribution is considered positive. However, as discussed, e.g. by Nath
(1969), the Little criterion too many have problems with inconsistency and
contradictions.

31 Alternatively, one may argue that under some conditions there are other possible
measures such as income taxes which are better suited to deal with distributional issues. Christiansen (1981) and Boadway and Keen (1993) derive sufficient conditions when the basic Samuelson rule may be applied in the presence of an optimal non-linear income tax. Of course, the existence of perfect differentiated lump-sum taxes is also a sufficient condition (Atkinson and Stiglitz 1980), but in the public finance literature this very unrealistic assumption is often considered to be of limited policy relevance.


33 During quantity changes, these measures are sometimes known as ‘compensating’ and ‘equivalent surplus’, respectively (Freeman 1993).

34 There is still an on-going discussion about whether the large disparities typically found between willingness-to-pay and willingness-to-accept measures in surveys for public goods can be explained within the framework of conventional economic theory, or whether non-conventional phenomena such as endowment effects play an important role; see Kahneman and Knetsch (1992a, b) and Smith (1992).

35 The fact that uncertainty and imperfect information have played a minor role for a long time in economic theory should, however, not be taken as an indication supporting the view that most economists were unaware of this strong limitation; see e.g. Knight (1921). Hirshleifer and Riley (1992) provides a good textbook treatment of economics and information.

36 From Jensen’s inequality we have that \( E(C(X)) > C(E(X)) \) if \( C(X) \) is a convex function, where \( E \) denotes expected value, \( C \) is the social cost, and \( X \) is a variable describing the outcome in physical terms (e.g. temperature increase due to global warming). Hence, a sufficient condition is that \( C \) is increasing in \( X \) at an increasing rate (i.e. that both the first and the second derivatives are positive).

37 See Bromley (1989) and Hodgson (1988) for good textbook treatments of institutional economics, and North (1990) for an example of the new institutional economics which is generally more closely related to conventional neoclassical economics. Hodgson et al. (1994) provides a good overview of published work in institutional and evolutionary economics.

38 See Goodwin et al. (1997) and Ackerman et al. (1997) for surveys of this promising development.

REFERENCES


Duesenberry, J.S. (1949) Income, Saving, and the Theory of Consumer Behavior,
——— (1991) ‘Interpersonal comparisons of utility: why and how they are and should be made’, in J. Elster and J.E. Roemer (eds) Interpersonal Comparisons of


(1975) ‘Non-linear social welfare functions: do economists have a special exemption from Bayesian rationality?’, Theory and Decision 6: 311–32.


Hicks, J.R. (1939) Value and Capital, Oxford University Press.


——— (1987) ‘Relative-income effects and the appropriate level of public
Ethics and economic policy recommendations

---


---


---


---


---

Articles


