

# Some reflections on the role of moral reasoning in economics\*

Bertil Tungodden<sup>†</sup>

June 24, 2004

DP 23/2003. updated June 2004

## Abstract

People seem to be motivated by moral ideas and in this paper I discuss how we should take this into account in positive and normative economics. I review alternative ways of modelling moral motivation and reasoning in positive economics and discuss how the presence of moral motivation may challenge the standard framework of welfare economics. I also discuss the need for invoking non-welfaristic principles in normative economics and whether these principles can be reconciled with the Pareto principle.

JEL Codes: D10, D60, D63, D71, H2, H4

## 1 Introduction

In recent years, we have witnessed an increasing interest in exploring and expanding the role of moral reasoning in economics (see for example Rabin (1993), Hausman and McPherson (1996), Roemer (1996), Fehr and Gächter (2000), Sen (2002), Brekke, Kverndokk, and Nyborg (2003)). To an outsider, this may seem like a very reasonable move, given the fact that moral issues are everywhere in private and public affairs. We are appalled by injustice

---

\*The arguments of this paper were initially presented at The Nordic Journal of Political Economy Conference 2003, on the future of economics. I should like to thank the participants for stimulating comments and discussion.

<sup>†</sup>Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway. E-mail: bertil.tungodden@nhh.no.

institutions, we condemn people for being immoral and we eagerly engage in discussions about the content of a good life. Of course, economists have recognized this for a long time (all the way back to Adam Smith), but still have been reluctant to broadening their analyses beyond the self-interested economic man in positive economics and welfarism in normative economics. Within positive economics, the standard justification has been that even though morality at the surface may seem like an important motivational factor, closer scrutiny will show that the moral dimension only plays a role in cases where not much is at stake (for example when paying tip on a restaurant) or can be completely explained by introducing more advanced models (for example by considering a repeated game). In normative economics, conventionally welfarism has been considered the only approach consistent with the Pareto principle, and hence beyond discussion (see for example Kaplow and Shavell, 2001). In this paper, I will argue that both these views are misleading and that we should welcome the recent attempts to broadening the framework of economics.

Some economists seem to fear that the recent development should make things too easy in positive economics and too vague in normative economics. Within positive economics, there will be no obvious way of restricting the motivational factors that may be included in the analysis and within normative economics we may end up with extensive incompleteness in our evaluations. But should this make us resist a careful analysis of the moral dimension? Isn't our aim to provide the best possible explanations and evaluations of the social phenomena at hand? Moreover, we may wonder what really is the cost of adopting a variety of motivational factors in positive economics. On almost every topic, we already have a myriad of economic theories due to extensive experimentation with the way we model the economic environment, and thus it is hard to see that anything is lost by allowing for experimentation with the preference structures as well. Finally, the possibility of incompleteness in normative analysis is certainly not avoided by restricting ourselves to the welfaristic framework. Incompleteness may also enter the stage within welfarism, if, for example, we only endorse the Lorenz partial ordering in the utility space (see Foster and Sen (1997) for a further discussion of partial orderings).

But some may still insist that by expanding the role of morality in our theories, we are no longer doing economics. To evaluate such a claim, let us take Lazear (2000) as the starting point. He argues that the main features of modern economics are the reliance on maximizing behavior, equilibrium

analysis, efficiency considerations, abstract and formal reasoning, and empirical testing. Do we challenge this perspective by expanding the role of moral reasoning? Not really. There is nothing in such a move that downgrades the role of abstract and formal reasoning and equilibrium analysis (which also should make it clear that we are not making things easy by broadening the framework), and, certainly, empirical testing only becomes more important when the space of theories expands. What about the reliance on maximization behavior and efficiency considerations? Again, there is no reason to give up the modern perspective. Further analysis of the moral dimension may, however, affect the way we interpret these features, and I now turn to a discussion of this issue.

## 2 Maximization and moral reasoning

The proper interpretation of maximization behavior has initiated fierce academic debates both within and outside economics (see Sen, 2002 for an overview). The standard position among economists is given by the revealed preference approach, commonly interpreted as saying that maximization behavior is equivalent to self-interested behavior. Self-interest is defined by the preference structure guiding people's choices, where individual choices are assumed to be the outcome of a maximization procedure. Conceptually nothing is changed by introducing morality as a motivational factor. The fact that some people choose to follow a moral norm, shows, by definition, that this act is in their self-interest.

However, in order to evaluate the claim that people's behavior is governed by self-interest in almost all cases of importance, we need a testable theory of self-interest. That is, we will have to outline an a priori account of self-interest that is independent of individual choices. One possibility is to define self-interest in narrow economic terms, stating that people act in a self-interested manner if they always choose the alternative that maximizes their own income or material well-being. But it is by now well established, for example from Ultimatum Game experiments, that the behavior of people in general do not confirm to this hypothesis. People do seem to pay attention to moral norms, for example fairness norms, in their behavior.

There are, of course, many other *possible* definitions of self-interest going beyond the narrow economic man, and in our discussion it is particularly interesting to consider the possibility of including the satisfaction or pleasure

derived from acting morally. Of course, it is far from trivial to apply such a theory of self-interest in empirical work (because it is not obvious how to measure the degree of satisfaction derived from an act), but still conceptually it is a possibility. The more interesting questions for this discussion, however, are, first, whether this is a good extension of our notion of self-interest and, second, to what extent it captures all there is to say about the role of moral reasoning in individual choices?

First, a problem with a definition of self-interest including the pleasure of acting morally may be that it is not of much interest for normative evaluations. A basic premise in much of modern welfare economics is that we get data for our normative analysis from observing people's behavior. However, this link between positive and normative economics depends strongly on the idea that people act on the basis of a preference structure representing a *normatively relevant notion of self-interest*. Hence, in the present context, we should ask ourselves whether the pleasure derived from acting morally is of importance when evaluating social arrangements (for example various distributive policies). If this is not the case, then the normative status of data revealed by choices relying on this broader notion of self-interest becomes unclear. Some may think that to question the role of revealed preferences in welfare economics is tantamount to rejecting the Pareto principle, but, as I will return to in the next section, it is not. Presently, it is sufficient to notice that the only way of saving the equivalence hypothesis between maximization behavior and self-interested behavior may be to introduce a definition of self-interest that challenges a fundamental part of modern welfare economics.

Second, to focus on the pleasure or satisfaction derived from acting morally when doing positive analysis, may be to overlook the special role of morality in individual choices. Of course, this claim calls for empirical support (showing that this theory of self-interest does not predict individual behavior correctly), but even without such documentation available and in order to inspire empirical investigation of this issue, I believe it is of importance to study alternative perspectives on moral choices. For this purpose, it will be useful to outline briefly a recent formulation of moral motivation in economics (Brekke, Kverndokk and Nyborg, 2003).<sup>1</sup> They consider the problem of providing a public good  $G$ . One possibility is to do this via public contribu-

---

<sup>1</sup>Brekke, Kverndokk and Nyborg discuss a number of other interesting issues not dealt with in the present paper, and hence I outline a stripped down version of their model focussing on the aspects essential for the present analysis.

tion  $G_p$ , another via voluntary contributions  $g_i = \gamma(e_i)$  from all individuals  $i = 1, \dots, N$ , where  $e_i$  is the effort of individual  $i$  measured in time units. Each individual faces the time constraint  $l_i + e_i = T$ , where  $l_i$  is leisure. As is well-known, a system of voluntary contribution implies undersupply of public goods in the case where people act in a narrowly self-interested manner, and the question raised by Brekke, Kverndokk and Nyborg is whether this also happens when people are motivated by moral ideals.

Their basic idea is that people have an understanding of what is a morally responsible contribution  $e_i^*$  and a concern for their self-image  $I_i$ , where their self-image is determined by the distance between their actual contribution  $e_i$  and the morally ideal contribution, i.e.,  $I_i = f(e_i, e_i^*)$ . We may read this as a measure of the pleasure or satisfaction people derive from fulfilling their moral obligations. To make things simple, they work with a specific version  $f(e_i, e_i^*) = -a(e_i - e_i^*)^2$ . The essential features of  $f$ , however, are that the function is increasing and concave in  $e_i$  and that we reach a global maximum in our self-image when we exercise what is considered the morally responsible behavior  $e^*$ .

This formulation of moral motivation raises two important questions. First, how do we determine morally ideal effort  $e^*$ ? Second, how do people incorporate moral ideals in their choices? On the first question, Brekke, Kverndokk and Nyborg suggest that people apply a version of the categorical imperative of Immanuel Kant, saying that morally ideal effort is determined by the action that would maximize social welfare, given that everyone acted in the same way. They simplify further by assuming that everyone is equal and, moreover, agrees that utilitarianism is the correct standard for evaluating outcomes. In this case, the morally ideal effort can be found by maximizing the utilitarian social welfare function.

$$(1) \quad \max_{e_i} W = U_1 + \dots + U_N,$$

where utility  $U_i = u(l_i, G, I_i)$  is assumed to be the relevant expression of individual welfare when calculating morally responsible behavior and we assume that  $e_i = e_j$  for all  $j \neq i$ . The first order condition is given by

$$(2) \quad \frac{\partial u}{\partial l_i} = N \frac{\partial u}{\partial G} \frac{\partial \gamma}{\partial e_i},$$

which says that a person exercises morally responsible behavior when the marginal utility of leisure exactly equals the social benefits (measured in utilitarian terms) of the public good produced by the marginal effort (when everyone else does the same).

On the second question, they follow the standard economic approach and include the concern for self-image (or the satisfaction derived from fulfilling

moral obligations) as an additional motivational factor in people's maximization problem. That is, they assume that  $U_i$  also guides individual behavior, where the utility function has all the standard properties. The first order condition of the individual maximization problem is given by

$$(3) \quad \frac{\partial u}{\partial I_i} = \frac{\partial u}{\partial G} \frac{\partial \gamma}{\partial e_i} + \frac{\partial u}{\partial I_i} (-2a(e_i - e_i^*)).$$

We may read (3) as saying that people in their choices are making trade-offs between the satisfaction they get from leisure, from the public good, and from acting morally. By comparing (2) and (3), it follows straightforwardly that people do not fulfill their moral obligations in equilibrium. If (2) and (3) were equal, then  $e_i = e_i^*$  and  $\frac{\partial u}{\partial I_i} (-2a(e_i - e_i^*)) = 0$ . But this implies that  $(N - 1) \frac{\partial u}{\partial G} \frac{\partial \gamma}{\partial e_i} = 0$ , which is not consistent with the assumptions of the model (if we assume that  $N > 1$ ).

Notice that this is a very general result. We need not assume that people are equal or follow the utilitarian rule when determining morally ideal effort. Actually, we need not even assume that they agree on how to determine morally ideal effort. In any case, as long as they are willing to trade-off their self-image against other dimensions of their welfare and the benefit of a better self-image is on the margin zero when exercising morally responsible effort, then they will not fulfill their moral obligations.<sup>2</sup>

To assume that people are willing to make trade-offs may seem like a trivial assumption, and actually Lazear (2000, page 101) views this as a corollary of maximization. But this is not true, formally speaking. Giving absolute priority to the fulfillment of moral norms can certainly be part of a well-defined maximization procedure. Most economists, however, find this a wildly implausible view on human behavior. Do we really believe that anyone will follow a moral norm without considering the costs of doing so? If not, then we have established that people make trade-offs, and we can disregard the non trade-off view as only covering non-representative people like (maybe) Mother Theresa and Nelson Mandela.

This line of reasoning, however, overlooks the essential distinction between making trade-offs when reasoning about the nature of the morally required act (in the case of Brekke, Kverndokk and Nyborg, when solving (1)), and making trade-offs when considering to what extent we should do

---

<sup>2</sup>In the case where people differ in their opinion on how to calculate social welfare, however, it is not straightforward to say that we get undersupply of the public good. It depends on how the definition of social welfare relates to the various opinions in society. Moreover, notice that the theoretical result does not tell us anything about the magnitude of deviation from the moral norm.

the morally required act (that is, when solving the individual maximization problem). Brekke, Kverndokk and Nyborg assume that people make trade-offs in both cases, but I should like to consider an alternative formulation saying that people only make trade-offs in the first case. There are two reasons for doing this. First, as economists, we should consider it important to understand what is excluded by the framework we usually adopt, and second, I do believe that the alternative perspective (more in line with the way philosophers think on this question) carries some weight.

If we really consider an act as morally required, then isn't it our instinct to follow it without further deliberation? That is, as I see it, part of the meaning of saying that we *ought* to do something. This claim does not imply that we always follow (what we consider as) required moral norms, but rather that our deliberative plan is to act in this manner. We may depart from this behavior due to, for example, weakness of will or insufficient self-command (see for example Sen, 2002), but the real issue is whether these deviations represent a consistent pattern of trade-off behavior.

The alternative perspective does not imply that personal costs do not enter the picture in moral reasoning. They do, but only when reflecting on whether something is a morally required act or not. By way of illustration, we usually consider it as immoral to lie and steal, but certainly not if it is a matter of life and death. Or we may consider it a moral obligation to vote in an election, but probably not if it becomes extremely costly in personal terms. In the model of Brekke, Kverndokk and Nyborg, this implies that it may be appropriate to apply a utility function involving trade-offs in (1), if we believe (as seems quite plausible) that there are trade-offs to make between self-image, leisure and access to public goods when calculating individual welfare. But the very same utility function should not be used when formulating the choice problem at the individual level, if we believe that people pursue what they consider the morally required act without making further trade-offs. This also illustrates the importance of acknowledging a conceptual difference between choice and well-being when formulating a utility function (as pointed out by Sen (1973)).

Notice that the alternative model does not necessarily portray an extremely optimistic picture of human nature. Most importantly, it does not suggest (the wildly implausible position) that people invoke moral reasoning in all their choices. The concern here is how moral reasoning affects the choice process *if* people perceive the situation as within the moral domain, and the alternative perspective may even coincide with a purely narrowly

self-interested view if people consider it as moral always to pursue their own narrow economic interests. Consequently, by way of illustration, it is entirely consistent with low turn-out rates in elections, if these rates reflect that a large portion of the population - for some reason or another - does not consider voting a morally required act.

The perspective on moral behavior outlined in Brekke, Kverndokk and Nyborg implies that people in general do not fulfill what they consider their moral obligations, whereas the alternative perspective outlined above suggests the opposite conclusion. What is the correct view on human nature? As I see it, the alternative perspective should not be rejected out of hand. It is far from trivial to claim that people make trade-offs in choices inspired by moral reasoning, and hence as economists we should avoid considering this as a purely technical issue about the utility function (to wit, whether it is continuous or not). Of course, the trade-off perspective may still be justifiable, but that needs to be established by empirical analysis.

Both perspectives, however, imply that moral values play a role in individual choices, and thus, in any case, it is of interest to discuss the nature of these values and how they are established in society. Brekke, Kverndokk and Nyborg provide a very nice illustration of one possible approach to moral reasoning, by arguing that people solve a maximization problem representing a formal version of Kant's thinking. Their approach is quite general in nature, but it may be instructive first to look at the specific procedure outlined in the paper. They assume that everyone agrees on using utilitarian reasoning as the basis for evaluating goodness in society, whereas they use Kantian reasoning when specifying the relevant assumption to make about others' behavior when solving the moral problem.<sup>3</sup>

Many people do not find utilitarianism - or any other social welfare function - an appropriate basis for moral reasoning. There are two main reasons for this. First, some argue that moral reasoning has to take place in de-

---

<sup>3</sup>Brekke, Kverndokk and Nyborg stress that there is no conflict between utilitarian moral philosophy and the Kantian Categorical Imperative in their model. That is true, if one restricts utilitarian moral philosophy to be about the goodness of social states. However, many utilitarians will insist that utilitarian moral philosophy also has implications for the assumptions we make about others' behavior when determining morally ideal effort. The (act) utilitarian perspective implies that we should solve the moral maximization problem by assuming that people do what they actually do (and not that they do the morally required act, as in Kantian reasoning), and this is what makes utilitarianism such a demanding moral system at the individual level.



liberations with others (see for example Anderson (1993)), and insist that any formal individual procedure of the kind suggested by Brekke, Kverndokk and Nyborg will fail for this reason. I will not pursue this perspective, but rather turn to a second line of criticism arguing that the formulation of Brekke, Kverndokk and Nyborg places too much importance on individual preferences. This may not be an important issue in a model where everyone is equal, but if we extend the analysis to a world with heterogeneous individuals, we get some questionable moral implications within the framework outlined above. By way of illustration, let us assume that people differ in their preferences for leisure. In this case, we cannot solve the moral problem by assuming that everyone else acts in the same manner, but rather have to assume that everyone else acts according to the same rule. This implies that we have to solve the maximization problem for every individual in society, which will give us a vector of effort levels specifying the morally ideal action for each person. It is easily seen that this procedure (whether we use utilitarianism or any other social welfare function) implies that two persons only differing in their preferences for leisure will face different moral requirements. The person valuing leisure strongly (on the margin) is required to exercise less effort than the other person. From a welfaristic point of view, this may seem fair, but I will claim that it is contrary to a widely shared moral intuition saying that people should be held responsible for their own preferences, which in this case implies that the morally ideal effort should be the same for both individuals. An interesting topic for future research would be to construct experiments testing the two alternative hypotheses on moral reasoning within the setting of Brekke, Kverndokk and Nyborg.

The result from such experiments would not only be of importance for positive economics. If it were the case that people are motivated by non-welfaristic ideas, then we also need carefully to study the possibility of incorporating non-welfaristic reasoning in normative economics. I now turn to a discussion of this issue.

### **3 Welfarism and moral reasoning**

Brekke, Kverndokk and Nyborg assume that moral reasoning takes place within a welfaristic framework, as we can see from the formulation in (1).<sup>4</sup>

---

<sup>4</sup>In the following, I only discuss welfarism within a single-profile setting, i.e. I do not consider cases where there is a change in the profile of utility functions. For a brief

But welfarism does not seem to capture the informational richness and focus present in public moral debates, where concepts like human rights, liberties, responsibility, freedom, equal opportunity, basic needs, and so on are commonly used. There is an extensive academic literature, inspired to a large extent by the work of Sen (see Sen (1999) and Sen (2002) for overviews), that aim at establishing alternatives to the welfaristic framework, but these non-welfaristic approaches have been met with skepticism within economics (see for example Kaplow and Shavell, 2001). There are mainly two reasons for this. Many seem to fear that normative economics becomes too vague if we move beyond the welfaristic framework; moreover it is commonly assumed that non-welfaristic approaches violate the Pareto principle. In my view, the problem of vagueness or incompleteness is not a reason for avoiding taking seriously all the valuable dimensions in normative evaluations, even though it may create problems in practical choice situations (see Chang (1997) for a further discussion of practical reasoning in the context of vagueness and incommensurability). As pointed out by Sen on a number of occasions, it is certainly better to be vaguely right than precisely wrong. Hence, the more interesting question is to what extent non-welfaristic principles are in line with the Pareto principle and capture something of normative importance.

Some readers may not consider violations of the Pareto principle too problematic either, because they should like to question the status of a person's preferences as a representation of his or her true interests. This criticism raises an important question about the role of preference formation in normative reasoning, but I will not pursue this issue here. In the following, I will simply assume that people's preferences provide a legitimate expression of their interests, such that if some person prefers social arrangement  $x$  to social arrangement  $y$ , then  $x$  is more in line with the true interests of this person than  $y$ . In this case, the Pareto principle becomes equivalent to what we may name the principle of personal interest, saying that if the true interests of everyone is more in line with  $x$  than  $y$  then  $x$  should be considered better than  $y$ .<sup>5</sup> The important question now is whether it is necessary and possible

---

discussion of the multi-profile setting and welfarism, see Fleurbaey, Tungodden and Chang (2003).

<sup>5</sup>This principle is related to what Broome (1991) calls the principle of personal good, but with one important difference. The principle of personal good says that  $x$  is better than  $y$  if everyone is better off in  $x$  than  $y$ , whereas the principle of personal interest deals with the interests of each person more generally. If the principle of personal good applies in a comparison of two alternatives, then we should expect the principle of personal interests

to combine such a principle with non-welfaristic considerations.

It is important to notice that in general, non-welfaristic considerations do not aim at challenging the Pareto principle (or the principle of personal interest). They are introduced in order to solve cases not covered by the Pareto principle, that is cases where we have a *conflict of interest* in society. There are two reasons for introducing non-welfaristic considerations (see also Brun and Tungodden, 2004), the pragmatic and the fundamental. The underlying idea of the pragmatic argument is that “we must respect the constraints of simplicity and availability of information to which any practical political conception [of justice] is subject” (Rawls, 1993: 182). Welfarism implies that interpersonal comparisons should be based on comparisons of preference satisfaction, which in general is considered to be non-observable. Thus, the welfaristic framework does not provide a practicable public basis for solving interest conflicts. The fundamental critique of welfarism is concerned with the substantive claims of this framework. Rawls (1971, 1993), for example, argues that utility or well-being is not the relevant feature of states of affairs. Appropriate claims should refer to an idea of rational advantage that is independent of any particular comprehensive doctrine of the good, and for this purpose Rawls suggests a list of primary goods. Rawls’ perspective can be considered as a version of equal opportunity ethics, which in its most general form states that society should indemnify agents against poor outcomes that are the consequences of factors that are beyond their control, but not against outcomes that are the consequences of factors that are within their control (Roemer 1998). Hence, if we believe that people should be held responsible for their preferences, then we should not rely on preference information when solving interest conflicts. Similarly, we may interpret a focus on human rights and basic needs (or basic capabilities in the language of Sen (1985, 1992, 1999)) as alternative non-welfaristic perspectives on how to solve interest conflicts without paying attention to individual preferences.

Is it reasonable to ignore individual preferences in cases where we have a conflict of interest? Let me provide a very simple example that should be of interest when discussing this question (for further discussion of this issue, see

---

to apply as well. But the converse implication does not necessarily hold. It may very well be the case that some people prefer  $x$  to  $y$ , even though they are worse off in  $x$  than  $y$  (as we also pointed out in the previous section in the discussion of the distinction between choice and well-being)

Notice that we may state weaker and stronger versions of these principles, see Suzumura (2001) in the case of the Pareto principle.

among others Fleurbaey (1995a,b), Roemer (1996), Cappelen and Tungodden (2002, 2003), and Tungodden (forthcoming)). Consider the case where we have two individuals who have already made their choices with respect to working hours, and where we assume that both value leisure and consumption. We are considering whether we should redistribute consumption (or income) between them. In order to make things simple, we also assume that there are no incentive problems (i.e. they will not make any further choices) and that we have complete information about their utility functions. Hence, we face a first best taxation problem where all redistributive policies are Pareto optimal. What kind of information do we consider relevant in this case? Welfarists will not be concerned with budget information, and consequently the welfaristic redistributive policy will be independent of whether they had equal opportunities or not. Equal opportunity ethics, on the contrary, completely neglects preference information and focuses only on the budget restrictions. If they both had the same opportunities, then this approach sees no reason to redistribute. If the budget restrictions differed, however, then equal opportunity ethics may defend some redistribution.

I believe that many people share the views of equal opportunity ethics in this example. If so, then it illustrates that it is far from obvious that we should apply the welfaristic framework when solving interest conflicts. But what about more complex cases where we face incentive and informational problems? In answering this question, it is important to notice the following asymmetry. The fact that welfarism does not work well in a simple example is sufficient to make clear that there is something problematic with this approach, but the fact that equal opportunity ethics works well in the very same example does not imply that it is a sound approach more generally. Hence, in order to provide a thorough evaluation of alternatives to welfarism, we need to see how they work in all interest conflicts and to what extent they are in line with the Pareto principle. It is beyond the scope of this paper to pursue such an extensive discussion of non-welfaristic reasoning, so let me here only briefly comment on the compatibility with Pareto principle. It turns out that a number of non-welfaristic ways of solving interest conflicts imply a violation of the Pareto principle, if we impose certain consistency requirements on our normative evaluations (see for example Sen (1970), Gibbard (1979), Fleurbaey and Trannoy (2001), and Brun and Tungodden (2004)). These results may be interpreted in two ways. If we considered the consistency conditions as logical requirements, then we can read the results as saying that it is not possible to combine these non-welfaristic approaches with the

Pareto principle. However, if we consider the consistency requirements from a more pragmatic choice perspective, then we can read the results as saying that we will face some choice situations with no best alternative if we endorse both the non-welfaristic framework for solving interest conflicts and the Pareto principle (see also, among others, Sen (1993) and Tungodden (2003) for a discussion of consistency requirements in social choices).

Lazear (2000: 102) stresses that "[Pareto] efficiency is a concept that together with equilibrium, pushes economists to do a particular kind of analysis. When economists model a situation and the resulting equilibrium is inefficient, usually there are trades that could have occurred that are implicitly or explicitly ruled out. The analyst or his critics are induced to ask what the reasons are and what market or other institutions could arise to remedy the situation...The notion that efficiency is a natural outcome motivates a larger series of questions and initiates deeper analysis. It also permits economists to make clear, unambiguous policy statements, although the assumptions that lie behind welfare economics are somewhat controversial." The importance of efficiency analysis is beyond doubt, but the distributive issue is also an extremely important part of economics. In the same way as we pursue remedies for inefficient allocations, we need to consider how to establish institutions that support equilibria with a *just* resource allocation. But this demands a clear understanding of how to properly evaluate social arrangements from the distributive point of view. Lazear seems to suggest that Pareto efficiency goes hand in hand with welfarism (if we read welfare economics as welfarism), but I hope that the present discussion has made clear that this is not the case. Welfarism is an informational perspective with questionable implications in distributive conflicts, whereas the Pareto principle is an (almost) uncontroversial condition on how to solve cases where individual interests overlap. Consequently, we should aim at establishing a non-welfaristic perspective capturing our real concerns in distributive conflicts. Such a perspective, however, may be hard to reconcile with the Pareto principle, and hence we may have to endorse the view of Rawls (1971: 52): "All theories are presumably mistaken in places. The real question at any given time is which of the views already proposed is the best approximation overall."

## 4 Concluding remarks

In my view, there are at least three reasons for welcoming the recent work on morality and economics. First, even though it may turn out that the economic profession is not convinced by the arguments and models outlined within this literature, it will certainly contribute to a deeper understanding of the nature of mainstream economics. It is only by considering alternative approaches that we can really reflect upon the standard methods and solutions within a field. Second, as I see it, much of the recent work on economics and morality does not really challenge the basic framework of economics, but is rather an interesting application of it. Finally, as I have discussed briefly in this paper, the present literature seems to provide us with some important insights in both positive and normative economics, which may change the way we think about maximization behavior and welfarism. There are many unresolved questions within the field, however, and hence it should be a most promising area for future research.

## 5 References

### References

- [1] Anderson, E. 1993. *Value in Ethics and Economics*, Harvard University Press.
- [2] Brekke, K. A., Kverndokk, S. and K. Nyborg 2003. "An Economic Model of Moral Motivation", *Journal of Public Economics*, 87: 1967-1983.
- [3] Broome, J. 1991. *Weighing Goods*, Basil Blackwell.
- [4] Brun, B. C. and B. Tungodden, B. 2004. "The Indexing Impasse: Is 'the intersection approach' a solution?", *Social Choice and Welfare*, 22: 49-60.
- [5] Cappelen, A. and B. Tungodden 2002. "Responsibility and Reward", *Finanzarchiv*, 59: 120-140.
- [6] Cappelen, A. and B. Tungodden, B. 2003. "Reward and Responsibility: How should we be affected when others change their effort?", *Politics, Philosophy and Economics*, 2: 191-211.

- [7] Chang, R. 1997. *Incommensurability, Incomparability, and Practical Reason*, Harvard University Press.
- [8] Fehr, E. and S. Gächter 2000. “Fairness and Retaliation”, *Journal of Economic Perspectives*, 14: 159-181.
- [9] Fleurbaey, M. 1995a. “Equal opportunity or equal social outcome”, *Economics and Philosophy*, 11: 25-55.
- [10] Fleurbaey, M. 1995b. “Equality and responsibility”, *European Economic Review*, 39: 683-689.
- [11] Fleurbaey, M. and A. Trannoy 2000. “The Impossibility of a Paretian Egalitarian”, *Social Choice and Welfare*, 21: 243-263.
- [12] Fleurbaey, M., Tungodden, B. and H. Chang 2003. “Any non-welfarist method of policy assessment violates the Pareto-principle: A comment”, *Journal of Political Economy*, 6: 1382-1385.
- [13] Foster, J. E. and A. Sen. 1997. *On Economic Inequality. Expanded edition*, Clarendon Press.
- [14] Gibbard, A. 1979. “Disparate Goods and Rawls’ Difference Principle: A Social Choice Theoretic Treatment”, *Theory and Decision*, 11: 267-288.
- [15] Hausman, D. M. and M. McPherson 1996. *Economic Analysis and Moral Philosophy*, Cambridge University Press.
- [16] Kaplow, L. and S. Shavell 2001. “Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle”, *Journal of Political Economy*, 109: 281-86.
- [17] Lazear, E. P. 2000. “Economic Imperialism”, *The Quarterly Journal of Economics*, 1: 99-146.
- [18] Rabin, M. 1993. “Incorporating Fairness into Game Theory and Economics”, *American Economic Review* 83: 1281-1302.
- [19] Roemer, J. E. 1996. *Theories of Distributive Justice*, Harvard University Press.
- [20] Rawls, J. 1971. *A Theory of Justice*, Harvard University Press.

- [21] Rawls, J. 1993. *Political Liberalism*, Columbia University Press.
- [22] Sen, A. 1970. "The Impossibility of a Paretian Liberal", *Journal of Political Economy*, 78: 152-57.
- [23] Sen, A. 1973. "Behaviour and the Concept of Preference", *Economica*, 40: 241-259.
- [24] Sen, A. 1985. *Commodities and Capabilities*, North Holland.
- [25] Sen, A. 1992. *Inequality Reexamined*, Harvard University Press.
- [26] Sen, A. 1993. "Internal Consistency of Choice", *Econometrica*, 61: 495-521.
- [27] Sen, A. 1999. *Development as Freedom*, Oxford University Press.
- [28] Sen, A. 2002. *Rationality and Freedom*, Harvard University Press.
- [29] Suzumura, K. 2001. "Pareto Principles from Inch to Ell", *Economics Letters*, 70: 95-98.
- [30] Tungodden, B. 2003. "The Value of Equality", *Economics and Philosophy*, 1: 1-44.
- [31] Tungodden, B. forthcoming. "Responsibility and redistribution: the case of first best taxation". *Social Choice and Welfare*.