Fair Air: Distributive Justice and Environmental Economics

Olof Johansson-Stenman · James Konow

Abstract   Are fairness concerns of relevance to environmental economics and, if so, are they sufficiently structured to improve analysis in this field? On both of these questions, we answer in the affirmative, arguing that people's fairness views are based on both general rules and the context, where context refers to the set of variables and persons employed to interpret and apply the principles. The fairness rules analyzed are accountability (i.e., rewards that are proportional to contributions individuals control), efficiency, need and equality. We conclude that stakeholders typically exhibit a "fairness bias", i.e., they tend, consciously or not, to interpret and apply fairness principles in a self-serving manner, whereas the views of spectators, or impartial third parties, tend to converge significantly more. Further, we argue that fairness considerations are relevant to both descriptive and prescriptive analysis in environmental economics. These fairness concerns are reflected in the behavior of private and public decision-makers and have potentially important policy implications through the overall social objective function.

Keywords   Fairness · Justice · Equity · Environmental policy · Behavioral economics · Experimental economics

O. Johansson-Stenman (✉)
Department of Economics, University of Gothenburg, Box 640, 40530 Gothenburg, Sweden
e-mail: Olof.Johansson@economics.gu.se

J. Konow
Department of Economics, Loyola Marymount University, One LMU Drive, Suite 4200, Los Angeles, CA 90045-2659, USA
e-mail: jkonow@lmu.edu

J. Konow
Erik Malmsten Guest Professor, Department of Economics, University of Gothenburg, Box 640, 40530 Gothenburg, Sweden

1 Introduction

A critical issue faced by policy makers across the world is how to distribute the costs and benefits of policies designed to address environmental problems, whether dealing with automobile emissions, water pollution, toxic substances, solid waste or global climate change. Whereas fairness concerns are frequently discussed in many different public policy areas, they appear to occupy a particularly prominent place in environmental issues. In fact, environmental policies, such as green taxes, fuel economy standards and tradable emissions permits, differ dramatically in the distribution of their costs and benefits.

In this paper, we conceive of fairness in rather broad terms as distributive justice, i.e., distributive preferences as distinct from self-interest and other social preferences, such as altruism, reciprocity and honesty. We focus on the contextual factors and the rules, including accountability, efficiency, need and equality, that guide those preferences. Moreover, we note that solutions to real environmental problems are often related to the kind of justice concerns we address here.

Although environmental economics, as well as economics more generally, is largely based on the normative criterion of economic efficiency, it is clear that many people have strong opinions about the perceived fairness of different environmental policies, including policies related to climate change and to the implementation and design of tradable permit systems. These opinions are reflected in the political debate, where fairness arguments are clearly much more prominent than in the environmental economics literature. Thus, although fairness is not a unique concern to environmental issues, it is a prominent theme in real world discussions in this area, even if it has been relatively neglected in the environmental economics literature.

Why, then, do economists not integrate fairness more frequently into their models? There are (at least) three reasons:

1. Such concerns are often seen as highly subjective and amorphous, i.e., they are not (sufficiently) structured to be productively incorporated into such models.
2. They are invoked chiefly in order to benefit the agent employing them, i.e. they are consciously or subconsciously chosen for self-interested reasons.
3. They complicate the models and make the analysis less tractable.

We think there is merit in these arguments but argue that each of these concerns can be satisfactory addressed. First, we certainly do not deny that there are subjective elements in people's fairness perceptions. Nevertheless, even these subjective views obey rules that vary predictably across contexts. Second, it is clear that people often exhibit what we will refer to as a *fairness bias*, that is, they construe fairness in self-interested ways. Stakeholders are frequently focused on the fairness of policies from a biased perspective, and proponents and even opponents of environmental measures often appeal to such arguments, e.g., the latter sometimes claim that gasoline taxes would harm the poor. Thus, even when self-interest insinuates itself into these deliberations, the arguments are often seemingly framed around a shared set of fairness principles (e.g., Lange et al. 2010). Moreover, these views can also be observed empirically under conditions of impartiality and reveal considerable consensus under such conditions. Third, as with all generalizations, it follows that fairness considerations do add a layer of complexity to formal models. But we maintain that the costs of incorporating fairness concerns into the analysis in terms of greater complexity are often outweighed by the benefits, which include improved empirical predictive power, richer descriptive theories, and greater policy relevance. We also believe that environmental economics would have greater

influence on the actual environmental policy if fairness considerations were to be taken more seriously.

The reminder of this paper is organized as follows: Sect. 2 discusses different means for measuring fairness concerns with particular attention to the role of fairness biases. Section 3 discusses different fairness rules that people apply and the associated empirical evidence, whereas Sect. 4 discusses further interpretive issues, such as how people handle situations involving tradeoffs between different fairness principles and context dependencies. Section 5 specifically discusses implications for environmental economics, including normative analysis. For example, one position is that fairness principles should inform the overarching goals of environmental policy, if one takes some version of the justice values of real people as relevant to normative foundations (e.g., Konow 2003). Section 6 concludes the paper.

2 Can Fairness Concerns be Measured, and, if so, How?

The fact that fairness judgments require various interpretative decisions introduces the possibility of arriving at very different conclusions. This means that there is rarely unanimity in views of justice, even though impartial judges often exhibit a high degree of consensus. The situation with stakeholders, however, is considerably more difficult, often with polarized views and sometimes even modest levels of consensus a remote ideal. Stakeholders frequently advocate and pursue goals that favor their material interests but that depart from the outcomes preferred, on average, by impartial spectators. We refer to this discrepancy between spectator and stakeholder views as a "fairness bias." This bias can be decomposed into at least two parts.

First, there is a "self-centered bias," which is that part of the discrepancy that is acknowledged by the stakeholder, at least to him- or her-self. For example, a polluter seeking to avoid regulation might recognize its efforts as serving primarily to protect its own profits, even if it does not cite that fact in making its case to regulators. Arguments motivated by this bias might or might not involve fairness, but, if they do, those people advancing them do not actually believe the arguments are legitimate but instead might simply be advancing them to manipulate and exploit the fairness beliefs of others. Second, there is a "self-serving bias," which involves a distortion of the agent's own beliefs about what is fair in the direction of his or her material interests. That is, the self-serving bias is due to self-deception about what is fair. This idea can be found in cognitive dissonance theory (e.g., Festinger 1957), which states that, when people are faced with opposing goals, they can relieve the resulting disutility by altering their beliefs. For example, if the executive leadership of a firm faces a conflict between maximizing profits and bearing their fair share of the costs of reducing emissions, they are likely to underestimate their fair share of abatement costs relative to the view third parties hold.

Babcock and Loewenstein (1997) report the results of a series of studies with their collaborators that investigate self-serving bias in the laboratory and the field. In particular, they find that this bias is significantly correlated with bargaining impasse, that this relationship is causal, and that it affects behavior not only in the laboratory but also has important repercussions in real world bargaining situations with high stakes, e.g., in teacher contract negotiations. In their experimental studies, subjects bargained within a contextually rich framework, a novel departure from the minimalist scenarios of most economics experiments. Research in psychology indicates that such richness feeds self-serving bias: Dunning et al. (1989), for example, find that self-serving evaluations increase with the number of criteria at one's disposal. This suggests that the numerous interpretative decisions inherent to most fairness

evaluations facilitate this bias. But the self-serving bias can manifest itself even in relatively sterile contexts. For example, the fairness bias is decomposed in a simple dictator experiment in Konow (2000). Based on actual behavior in economic experiments, around 60% of the overall bias in dictator decisions can be attributed to self-serving bias, whereas about 40% is due to unadulterated self-interest (i.e., the self-centered bias). Such subconscious biases can be expected frequently in environmental contexts, which are often characterized by countless specifics and informational asymmetries. In addition, these same properties can lead to conscious strategic manipulation of fairness concepts. For example, stakeholders, such as polluting firms, frequently possess better information about their costs of complying with environmental objectives than, say, their competitors or government regulators, and one might expect the former to exaggerate these costs and the latter to discount them.

Lange et al. (2010) study fairness bias in international climate policy. They report a survey of almost two hundred individuals actually involved in negotiating climate policy across three continents. Specifically, respondents state their views of the fairness of four rules informed by different interpretations of the fair distribution of obligations to reduce greenhouse gas emissions. They find that individuals in different geographical regions generally favor rules that impose lower costs on their region and higher costs, by implication, on those in other regions, apparently in support of a fairness bias. The data in this study do not allow a decomposition of this bias, but one piece of evidence is suggestive of some role for self-serving bias: the respondents tended to view their own preference of rules as less self-interested than the preferences of others. The latter fact is consistent with a large literature suggesting that most people seem to consider themselves as "better" than the average person in dimensions such as social responsibility. For example, Carlsson and Johansson-Stenman (2009) found that people on average think that others vote more for selfish reasons than they do themselves.

To the extent stakeholders impacted by environmental policy are subject to a self-serving bias, the results from fairness research are not auspicious. In particular, the stakes in real world environmental problems are much higher than in an economics experiment. Yet even in experiments it is quite difficult to dislodge this bias. Babcock and Loewenstein (1997) report various interventions to mediate or purge the self-serving bias among subjects, mostly to little effect. Informing subjects of the bias had no effect, nor did writing an essay arguing the opponent's case, although informing subjects of the bias combined with having them list weaknesses in their own case did significantly reduce the bias. The study of Oberholzer-Gee et al. (1997) suggests that, ultimately, solutions to such conflicts might reside in lessons from procedural justice. They survey a sample of the general Swiss population on their views of the fairness of different procedures for the siting of hazardous waste facilities, viz., waste from nuclear power plants, and also report facts about the actual siting of such a facility. They find strong support for processes that respect impartiality, information, consent and fairness, as well as a greater willingness for people to accept a burden under such conditions.

Given the frequent evidence of fairness biases one may wonder whether fairness perceptions are inherently subjective, and whether it is at all meaningful to talk about objective or impartial measures of fairness? Yet, we argue that it is both meaningful and possible to analyze measures of fairness objectively. We maintain, however, that such attempts should eliminate, or at least minimize, the direct material stakes of the subject whose fairness preferences are measured in order to reduce fairness biases. There are primarily two kinds of broad measures for doing so: (1) survey studies where subjects are asked to state which of different alternatives is the fairest one (e.g. Konow 1996), and (2) experimental studies with real monetary incentives where the subjects act as third party decision-makers who choose allocations for other subjects, implying that their decision will have real effects for the other subjects but not for themselves (e.g. Konow 2000).

3 Fairness Rules

Distributive justice, which we use here interchangeably with fairness, concerns moral preferences over the distribution of social and economic benefits and burdens among a group of individuals. As a *moral* preference, it is distinct from self-interest. Viewed as *distributive* preferences, justice also differs from altruism, which in economics typically characterizes an unconditional preference, such as pure altruism or warm glow (see Andreoni 1990). Instead, justice takes a more specific form and might involve some general relationship between preferred outcomes and other factors such as desert, endowments, productivity, need or effort. As behavioral economics has developed over almost three decades, fairness concerns have figured prominently throughout, for example, fairness has frequently been suspected in the results of experimental bargaining games beginning with Güth et al. (1982). But it is only relatively recently that fairness has been analyzed as a possible force in environmental economics, e.g., Brekke and Johansson-Stenman (2008), Lange and Vogt (2003). Numerous experimental and field studies indicate that such preferences have frequent and quantitatively large impacts on economic behavior. Nevertheless, despite certain patterns, including the frequent incidence of equal splits in bargaining experiments, the findings defy explanation according to a single rule or principle. In this paper, we examine four rules that we believe characterize actualized distributive preferences: accountability, efficiency, need and equality.

3.1 The Accountability Principle

The first rule we discuss is a relative concept of fairness, i.e., it pertains to the allocation of a benefit or burden of a given size among a set of persons. The *accountability principle* states that a person's allocation should be in proportion to the relevant variables he/she controls, but this rule does not hold the person accountable for other differences (Konow 1996). For example, if worker A is twice as productive as worker B, A should earn twice as much as B, if the productivity difference is due to A working more hours or exerting greater effort. But they should earn the same, if the productivity difference is due entirely to factors they do not control, such as innate ability or differing work conditions. Thus, accountability integrates two concepts: proportionality and responsibility.

Justice as proportionality is often referred to as *equity theory* in psychology, sociology and political science. Equity theory states that individuals are motivated to establish or maintain proportionality of agents' inputs (e.g., their contributions to production) to the outputs they receive (e.g., their earnings), although, as usually formulated, equity theory does not restrict the variables that may count as inputs. Empirical studies substantiate the economic importance of proportionality, e.g., Gächter and Riedl (2005), Güth (1994), Selten (1978) . Specifically, proportionality emerges as a quite robust finding when impartial preferences are elicited. In Gächter and Riedl (2006) , survey participants respond that the fairest division of earnings in a vignette is in proportion to historical performance. In the dictator experiments of Konow (2000), Konow et al. (2009) , third parties allocate proportionately to the individual performance of other subjects in a task.

The other part of accountability is responsibility, which is sometimes associated with the concept of just deserts, e.g., Nozick (1974) , and is often considered central to normative theories of fair reward, e.g., Bossert and Fleurbaey (1996), Cappelen and Tungodden (2003), Roemer (1998). Alesina and Angeletos (2005), Fong (2001) find that popular support for income redistribution is strongly related the perceived control of recipients over their circumstances. The accountability principle merges proportionality and responsibility:

fair outputs are proportional to inputs agents' control. This principle has been corroborated in experiments: in the aforementioned dictator study of Konow (2000), third parties allocate proportionally, when productivity is chosen, but equally, when productivity is due solely to arbitrary differences between subjects. The accountability principle is also related to important work in normative ethics, e.g., Cappelen and Tungodden (2009), which potentially informs policy. In the context of environmental policy, it implies, for example, that those who reduce emissions should benefit proportionately or, conversely, that polluter costs be proportionate to their emissions. Thus, even if the conditions of the Coase Theorem apply and pollution abatement can be as efficiently achieved by assigning property rights to emitters as to the parties affected by the externality, this principle suggests that people are not neutral with respect to the implied distribution of burdens. This is consistent with the survey results in Konow (1996) showing that, even when it is considerably more expensive for the emitter to eliminate the pollution, 84% of 219 respondents find it unfair when the court assigns property rights to the emitter, whereas 80% of 240 respondents find it fair when the court assigns these rights to the affected party.

In terms of burden sharing for global carbon emission reduction, this seems to imply that every person in the world should have an equal initial right to pollute, since it is hard to argue that some individuals have earned the right to pollute more than others. Although such a global distribution of emissions would, of course, imply large inefficiencies, it is, in principle, straightforward to combine it with a market-based instrument that deals with the efficiency aspects. One possibility, for instance, is a global system of tradable permits, where the initial allocation is proportional to the population size in each country. Alternatively, a system with very similar distributional implications is a global tax on emissions, whereby the revenues are distributed back in proportion to population size.[1] Of course, policies that seem ethically appealing (e.g. Kverndokk 1995) are not necessarily politically feasible, and there are also further practical hurdles, such as widespread political corruption in many poor countries, which complicates the picture.

3.2 The Efficiency Principle

The efficiency principle simply calls for maximizing total surplus. In economics, fairness and efficiency are often considered to be at odds, and much discussion concerns the equity-efficiency trade-off. To be sure, the two can conflict, as with a lump sum tax that impacts living standards of low income individuals disproportionately. However, they need not necessarily conflict, for example if pollution permits are allocated so as to favor developing economies as mentioned above. Moreover, survey evidence shows that views of fairness encompass not only equity (i.e., accountability) but also efficiency; indeed, fair allocations reflect trade-offs between competing distributive rules (see Konow 2001). Experimental studies corroborate the trade-off between efficiency and equity; see e.g. Engelmann and Strobel (2004), Fehr et al. (2006), Stahl and Haruvy (2009). In fact, many subjects exhibit a remarkable willingness to sacrifice personal earnings in order to increase the total, e.g., in the dictator games reported in Andreoni and Miller (2002), Charness and Rabin (2002).

The importance of efficiency among distributive preferences is good news for environmental economics, most of which has, to date, stressed efficiency as the normative goal of

---

[1] Furthermore, one could even argue that poorer countries should have the right to emit more than the richer countries in the future, in order to compensate for their lower emissions in the past.

policy.[2] Indeed, the news only seems to get better, when one considers recent experimental evidence suggesting that surplus maximization also affects the willingness to cooperate. In a voluntary contribution game, Oxoby (2006) allows subjects to constrain the choices of others to more efficient outcomes and finds that, if they can dictate cooperation by others, subjects then voluntarily choose themselves to contribute rather than to free-ride. Nevertheless, collectively, the evidence suggests that efficiency does not exhaust the set of preferences relevant to cooperation. Stahl and Haruvy (2009) argue that the results of their extensive-form games are best reconciled with a behavioral model that incorporates self-interest, efficiency and inequality aversion.

3.3 The Need Principle

The need principle of distributive justice requires the satisfaction of the basic needs of all individuals. Similar to the efficiency principle, it is an absolute standard, but it pertains to minimum allocations to individuals as opposed to maximization of aggregate allocations. Survey results indicate that people associate fairness with the satisfaction of basic needs, but when needs satisfaction conflicts with accountability or efficiency, there is a trade-off among these goals (e.g., Konow 2001). Some dictator experiments are suggestive of a concern for need. Eckel and Grossman (1996) find that student dictators are less generous toward fellow students than toward the presumably more needy recipients of their donations to the American Red Cross. Crumpler and Grossman (2008) present subjects with a choice of charities, and dictators seem to choose most frequently ones serving those whose basic needs or health is threatened. The results of Konow (2010) provide stronger support for a concern for basic needs: in one treatment, dictators donate to one of two charities that explicitly provide for basic needs and, in other treatments, to fellow students. Transfers in the former significantly exceed those in the latter and are the largest of any laboratory dictator study (the modal gift was the dictator's entire endowment). Moreover, in post-experimental questionnaires, roughly two-thirds of the more generous dictators explain their motivation in terms of a concern for the needs of recipients. Interestingly, many of the more selfish dictators in this study also explain their allocations in terms of need but by appeal to their own needs.

Despite the prominent place given to needs in the early days of economics, e.g., by Malthus (1798) and George (1879), it occupies a considerably less visible position in most modern economics, excepting, to some degree, in development economics. In particular, most economics experiments and social preference models that have accompanied them do not consider need among the motivations. Perhaps this is related to the fact that most of this research has been conducted in the developed world, where such concerns are not salient in the populations usually studied. Nevertheless, with an estimated 40% of the world's population living on less than $2 per day (World Bank Development Indicators, 2005), need certainly cannot be ignored in an international analysis of justice. In fact, discussions about the fairness of the burden of climate change policy often centers on differences between developed countries and those that have to cope with high rates of poverty. In a survey of agents involved in negotiations over climate change policy, Lange et al. (2010) find strong support for the "ability to pay" rule that favors less developed countries by assigning responsibility for abatement costs based on GDP.

---

[2] Nevertheless, as discussed further in Sect. 5.3, it is possible that some people value efficiency without considering it to be an element of fairness.

3.4 The Equality Norm

Probably the most striking stylized fact of experimental studies involving social preferences is the high incidence of equal splits. A frequent, and often the modal, decision is one that creates equal earnings in designs as varied as the dictator game, ultimatum game, and trust game. Most of the major theories of social preferences that have been developed to account for experimental results incorporate equality as an argument, including Bolton and Ockenfels (2000), Charness and Rabin (2002), Fehr and Schmidt (1999), although see also Cox et al. (2007) for one exception. It is usually the case that results of public good experiments are also characterized by equality, but of a different kind, viz., equal contributions. Buckley and Croson (2006), Cherry et al. (2005), for example, conduct linear public good experiments with endowment heterogeneity. To equalize net earnings, those subjects with high endowments should contribute more than those with low endowments. But, instead, there is a pronounced tendency in these studies for subjects to match absolute contributions irrespective of endowments or net earnings. Indeed, such matching of transfers is a stylized fact of many experiments involving cooperation, including Charness et al. (2007), Croson et al. (2005), Fischbacher et al. (2001). These two goals (equalizing earnings vs. equalizing contributions) have very different implications for the distribution of the costs and benefits of environmental policy.

We believe these conflicting goals can be better understood by distinguishing impartial fairness preferences from the potentially biased views of stakeholders. The former are important as guides to normative economics and economic policy and, to the extent they are shared by stakeholders, to descriptive analysis of fairness. The latter are crucial to understanding and predicting behavior in many real world situations in which self-interest is implicated but agents might also have to factor in their distinct fairness preferences from those of spectators. Let us consider and clarify these two perspectives in turn.

Evidence on impartial preferences comes from survey studies of fairness and from experimental studies employing third party decision-makers who choose allocations for other subjects. Based on these, Konow (2003) concludes that equality is not a *principle* of justice, i.e., a general distributive rule endorsed by impartial third parties. The argument is that equality does often emerge for a variety of reasons, but not because of a general, impartial preference for it. One reason for equality is that it is sometimes a *special case* of a general principle that otherwise produces inequality. For example, equality is a special case of accountability, if the contributions individuals control are equal, of efficiency, if equal rewards maximize the total, and of need, if the resources needed to satisfy basic needs are the same across individuals. Often, information about differences that are relevant to general principles is unavailable or unreliable. In such cases, people usually make the *ceteris paribus assumption* about such differences and favor equality by default. A corollary of this last statement is that, when information relevant to other principles becomes available, people will abandon equality in favor of those principles. This is consistent with the survey results on fairness attitudes in Konow (2001) and with the allocation decisions of experimental subjects who reason impartially in Herne and Mård (2006). The ceteris paribus assumption is a plausible explanation for the prominence of equal splits in many economics experiments: in order to focus on a particular question or to avoid inadvertent cues, many of these studies have a very lean design, but then they often lack the contextual elements or means for subjects to apply principles like accountability, efficiency or need. When the information and opportunities exist to act on these principles, subjects typically depart from equality, as seen for the sake of efficiency in Andreoni and Miller (2002), for accountability in Gächter and Riedl (2006), and for need in Konow (2010).

Thus, the evidence casts doubt on equality as a principle of justice, i.e., a general rule endorsed from an impartial or impersonal perspective. Nevertheless, it is clear that equality is frequently employed by stakeholders in personal (as opposed to impartial) relationships, such as with the treatment of co-workers, duties among family members, earnings of counterparts in bargaining experiments, or sharing of a dinner bill among friends. Equality is, therefore, a distributive rule but one of a different kind, which is why we refer to the equality *norm* instead of the equality *principle*. As used here, norm connotes a moral rule distinct to personal relationships, whereas principle denotes a general and impartial rule. A principle might also be valued, of course, in personal relationships, e.g., co-workers might be motivated not only by equality but also by accountability, but a norm is relevant only in personal contexts. This typology is consistent with the results of Konow et al. (2009). Their dictator study involves allocations by third parties, or impartial spectators, as well as by standard dictators, or stakeholders, who vary across treatments in how personal their relationships are to their paired recipients. Spectators allocate strictly according to accountability, whereas stakeholders shift toward equality, and the shift is greater, the more personal their relationship to recipients. These patterns are remarkably robust with respect to a wide range of variables, including nationality (US or Japanese), culture, race, income and gender.

Contributors to a public good are also in a personal relationship: even if they are anonymous, they are matched with other subjects and share earnings in a manner that depends on their individual decisions. Public good experiments often produce marked patterns of equality, but this typically takes the form of a tendency to equalize absolute contributions as in Croson (2007). Cherry et al. (2005) find a similar tendency: high endowment subjects do not contribute proportionately more than low endowment subjects, but the former do contribute significantly more in absolute terms, suggesting that there is an attempt, on average, partially to offset differences in endowments. These results relate easily to debates that occur over the fairness of environmental policies. For instance, negotiations about global warming often revolve around the fairness of rules that require equal contributions, such as the egalitarian rule, versus those that tend to equalize incomes after the costs of abatement, such as the ability to pay rule.

4 Tradeoffs and Interpretive Issues

The fairness rules described in the previous section are conceptually quite simple, so it seems that applying them should be relatively straightforward. Nevertheless, the frequency and ferocity of debates over fairness make it obvious that their application is anything but straightforward.

A little reflection reveals, however, that, even if there were complete agreement on and clarity about the rules of justice, people would not be able to judge the fairness of allocations without specification of additional variables. What is required is a series of interpretative decisions about contextual factors. First, what is the relative importance of each rule, when the pursuit of one rule conflicts with the satisfaction of another? For instance, how does one balance the efficiency of reducing toxic industrial waste with the potential loss of employment? Indeed, the experimental literature is replete with examples of heterogeneity of types according to their social preferences or their willingness to act on them. Andreoni and Miller (2002), for example, report that about 47% of their dictators are selfish, 30% want equality and 23% prefer efficiency. Studies of cooperation in public good games and other social dilemmas similarly suggest that people fall into different categories, including defectors,

conditional cooperators and unconditional cooperators (see Chaudhuri 2009). Agents also differ in their motives, which can include, among others, altruism and warm glow (Andreoni 1990), or reciprocity and conformity (Alpizar et al. 2008).

Even if there is agreement on the first question about how to weigh fairness rules, however, the second set of questions reveals even deeper interpretative issues. Specifically, which people and what variables should be included when evaluating fairness? This question can be contentious, even if one is trying to apply only a single rule. For example, when attempting to determine fair country requirements for the reduction of greenhouse gases according to the accountability principle, should nations with historically higher emissions bear greater costs given that their situation resulted from their choices of inefficient vehicles and dirty technologies, or should their burden be moderated by claims that their higher emissions are partially due to factors outside their control, such as greater travel distances or an outdated infrastructure that is expensive to renovate? As Cappelen et al. (2007) demonstrate in their dictator game study, people sometimes arrive at very different conclusions about what variables are relevant when assigning responsibility. The other question is which people to include when considering the repercussions for fairness of any policies. For example, should one consider the effects of a sulfur dioxide tradeable permits program only on those directly affected, or should one adopt a broader perspective and also consider its effects on their incomes relative to those of others in the region, country or world? With global climate change, the choice seems more obvious, at least for normative matters, that one should include peoples of all countries affected. But then does one also try generally to tackle broader questions of international inequities in income? The evidence on stakeholder behavior suggests people tend to choose a narrower, local view. Ruffle and Sosis (2006), for example, find significant in-group-out-group bias in a field experiment on cooperation among kibbutz members and city dwellers.

Nevertheless, a considerable volume of survey and experimental research suggests that impartial judges often achieve a fairly high of agreement about what is fair, even if stakeholders more frequently disagree (see Konow 2003, for examples, including for the following conclusions). What seems to guide stakeholder views and often create spectator consensus are the particulars provided by the context. First, people typically work with the available information in evaluating the relative importance of fairness rules. Thus, for example, if people in a developed country are presented with the question of how to divide a given salary pool between two workers who have chosen to work different hours, they will apply accountability and favor proportionality, ignoring both efficiency, since the pool is fixed, and need, since the basic needs of workers in developed countries are not typically endangered. Similarly, as previously mentioned, student subjects in dictator experiments are more generous toward charities, responding to the need of their beneficiaries, than toward fellow students, where the equality norm usually applies and requires only equal splits. Second, impartial judges usually exhibit a high degree of agreement about the choice of variables and people relevant to fairness evaluation. Again context provides most direction on and, consequently, answers to these questions. For example, survey respondents view a CEO's high salary as fair, when compared to other similarly paid executives, but as unfair, when compared to the much lower compensation to workers. As this last example illustrates, however, not all information is created equal, and judgments can be sensitive to framing effects. This is useful to know for descriptive analyses of fairness but perhaps troubling for prescriptive work on justice. Nevertheless, even the latter concern seems less serious in light of evidence that spectators can identify morally relevant information and, when they do, arrive at a consensus about fair allocations Konow (2009).

5 Fairness Concerns and Environmental Policy

The studies cited above provide ample support for the claim that people care about fairness. But does this fact have any implication for environmental policy and for how we should systematically analyze environmental problems? We answer these questions in the affirmative for a variety of reasons.

5.1 Cooperation Possibilities and Fairness Concerns

The climate can be seen as a global public good, since we can all benefit from it, but we cannot exclude others from also benefiting. This is, of course, also the core of the problem, since what is rational for a single country in isolation is globally suboptimal. If each country has to pay for its own abatement costs of reducing greenhouse gas emissions, while all countries now and in the future share the benefits, there is clearly room for free-riding, such that each country continues to emit more than what is globally optimal. In order to prevent free-riding, we need agreements on some kind of a cooperative solution, e.g., in terms of coalitions. There is a relatively large theoretical literature on negotiations related to transnational pollution, e.g., Carraro and Siniscalco (1998), Asheim et al. (2006), but almost all of this literature assumes that each negotiating country cares solely about its own material payoff (cf. Barrett 2005). Some of this literature concerns repeated games, i.e., negotiations occur not once, but rather multiple times, and asymmetries between, negotiating parties, i.e., they take account of the fact that some countries are much larger and more powerful than others. Other approaches, so-called differential games, deal both with strategic interaction and dynamic optimisation simultaneously. Moreover, parts of the literature allow for collusion, i.e., the possibility that some countries cooperate with one another against other countries.

Carraro and Siniscalco (1993), Barrett (1994) provide the standard approach to studying coalition formation in the climate context. They use a two-period non-cooperative framework in which countries first decide whether or not to join a coalition. Those who join will then behave cooperatively with one another in a second stage. Both the coalition (as an entity) and the remaining countries choose their emission levels non-cooperatively. As shown by Carraro and Siniscalco (1993), Barrett (1994, 1997), the resulting coalition size tends to be rather small.

Lange and Vogt (2003), on the other hand, show that, when countries care not only about their own absolute payoffs but also about fairness, the coalition size tends to be larger, and even the grand coalition (that includes all countries) can be stable. In contrast to Lange and Vogt (2003) who employ a symmetric model, Lange (2006) considers the situation with asymmetries, in which case it is no longer obvious that fairness concerns increase the coalition size.

The above examples show that fairness concerns can, in principle, have significant implications for negotiating parties, providing a rationale for incorporating such concerns into formal environmental economics models. However, the major entities making and attempting to influence environmental policy are often large organizations, such as governments and corporations, rather than the individuals whose decisions often form the basis for the behavioral findings reported previously. That individuals seem to care about fairness does not, of course, imply that the Board of Directors of Exxon and Archer Daniels Midland or the leaders of industrial nations are similarly motivated. Indeed, the fact that decisions about environmental policy are often made by large groups does not seem to bode well for fairness and for international cooperation on the environment, given evidence that groups tend to act

in a more selfish way than individuals (e.g., Luhan et al. 2009). Thus, one might expect resistance to environmental initiatives from countries and large corporations. On the other hand, groups, in the end, consist of individuals, who bring their own values into these interactions, as the aforementioned results of Lange et al. (2010) suggest.

Moreover, these organizations are often constrained by the values of others, even if they do not place great value on fairness themselves. Governments (even undemocratic ones) are usually subject, to some degree, to the views and passions of their constituents, and fairness likely plays a significant role in the motives of the latter. Given the negligible instrumental value of voting relative to its cost, Brennan and Hamlin (1998) conclude that an important reason people vote in the first place is for expressive purposes. This claim finds support in field and experimental studies, e.g., see Sobel and Wagner (2004), Tyran (2004), respectively. The putative importance of this expressive motive suggests that voting is a reflection less of a concern for material payoffs than of values, such as fairness. Similarly, corporations must, of course, pay attention to their potential customers, who can engage in boycotts and appeal to state intervention, as evidenced by their public relations campaigns and by adjustments of business practices in response to expressed or expected public reactions.

Thus, although self-interest will certainly influence group behavior, as it does individual behavior, we can also expect fairness to play a role. Moreover, the project of determining the influence of fairness on environmental issues is not aimed at proving that fairness always dominates such issues, but rather at understanding better its distinctive and often significant role, both for descriptive as well as prescriptive analysis.

5.2 Environmental Valuation and Fairness

In survey-based environmental valuation investigations, such as contingent valuation studies, people are implicitly or explicitly asked to value an environmental improvement, or the avoidance of an environmental damage, in terms of their maximum willingness to pay (or WTP). It is normally assumed that these values reflect their preferences in a manner that corresponds to their individual well-being. There is also evidence, however, from "think-aloud" methodology that the results of environmental valuation studies appear to be influenced by what people believe their fair share of the costs associated with an improvement would be (Schkade and Payne 1994).

Note that this motive is distinct from other non-selfish motives, such as those based on non-paternalistic or paternalistic altruism, cf. McConnell (1997), Johansson-Stenman (2000). Although fairness-based WTP statements might provide the authorities with some valuable information in terms of priorities (see Sect. 5.3 below), it does not make sense to incorporate such responses uncritically into a conventional cost-benefit analysis as if they reflected the respondents' true values in terms of their well-being. There are at least two reasons for prudence in this regard: first, there are questions about the contingent valuation method that have yet to be answered, and, second, there are questions about interpreting preferences (even fairness preferences) as well-being, which we will address in the following sub-section.

On the other hand, these facts do not rule out the usefulness of such studies for purposes of measuring people's preferences over the distributive consequences of solutions to environmental problems, such as to the problem of climate change. For example, Cai et al. (2008) provide a carefully executed web-based contingent valuation study of a hypothetical referendum type that measures willingness to pay for global warming mitigation. They found that people are typically willing to pay more, when the payment vehicle implies that larger cost shares are borne by parties deemed to bear a greater responsibility for mitigation, and when

respondents believe that the effects of climate change might be borne disproportionately by the world's poor. These results are, of course, entirely consistent with the *accountability* and *need* principles discussed above. Nevertheless, as useful as stated preference methods may be, it is important to keep in mind that, when it comes to welfare analysis, it is inappropriate to equate people's willingness to pay with individual measures of welfare change.[3]

5.3 The Social Objective

Let us suppose that people have fairness preferences as outlined above. Does that mean that the social objective function in economic analysis should change accordingly? We argue against the a priori imposition of some social objective, including that of the scientist. Thus, economists should be able to analyze normative problems based on an ethical foundation (e.g. some version of utilitarianism) that the economist does not share. Hence, in our view it makes perfect sense to conclude that based on the ethical assumption A, behavioral assumptions B, and market equilibrium assumption C, that conclusion D follows, even if the economist undertaking the analysis does not share the views underlying the ethical assumption A. Yet, in practice we often have to make some underlying normative assumptions, and we find it different to defend the a priori exclusion of fairness considerations in normative analysis. How incorporating fairness considerations affects the normative analysis depends on at least two broad considerations. Put simply, the one consideration concerns what, exactly, to incorporate, and the other how, exactly, to integrate it. These are addressed in turn below.

Fairness terminology (including the terms justice and equity) is employed with differing levels of specificity, both in theoretical work and in common usage, a fact that has been noted from Aristotle's *Nicomachean Ethics* to modern social science research. That is, at times, it signifies the whole of goodness or virtue, including reciprocity, honesty, charity, etc., whereas, at other times, it represents values that are a proper subset of goodness. Hence, the question "What is fair?" is sometimes the same as the question "What is right?" or "What is good?" In other cases, it signifies something more specific, whereby there is survey evidence that the most specific sense of fairness is equivalent to accountability alone (Konow 2001). We have chosen, in this paper, to define fairness in an intermediate sense as distributive preferences, partly because we believe this represents a close approximation of everyday usage of fairness terminology, which reflects a combination of general and specific meanings. But this implies that evidence on fairness preferences will give disproportionate weight to more specific meanings, like accountability.

To get an intuitive sense of this last claim, consider the following thought experiment. Suppose two workers, A and B, have the same ability, work equally hard, and produce the same amount of output. Which allocation, I or II, below is most *fair*?

I. A earns 100 USD per day, B also earns 100 USD per day
II. A earns 200 USD per day, B earns 150 USD per day

Since both workers have contributed equally to output, allocation I seems fairer, consistent with the accountability principle, which calls for equal allocations in such a case. But consider instead the following question: Which allocation do you think *should* be implemented? Now allocation II seems a more acceptable response. Framed this way, the second question explicitly elicits more general moral preferences, and efficiency, therefore, receives increased weight in the respondent's moral calculus. There is no inconsistency, therefore, in

---

[3] At least, this caveat applies, if one means by 'welfare change' changes in actual or expected well-being; if 'welfare change' means something else, that must be specified differently.

maintaining that I is fairer than II but that II is better than I. In the framework of both welfare economics and consequentialist ethics more generally, goodness can be the overarching goal, even if fairness is an element of goodness. Hence, it does not follow that, if people believe that Policy I is fairer than Policy II, society should prefer Policy I to Policy II.

Thus, any attempt to incorporate moral preferences into a social objective function must first address the question of the level of specificity in those preferences. We advocate a very general approach that incorporates goodness in the broadest terms and believe that the burden of proof, in this matter, must fall on claims for a more narrow definition that excludes some moral preferences. Thus, we choose to treat fairness broadly as distributive justice here, since this encompasses many, if not most, moral preferences in this context. Nevertheless, in the interests of brevity and tractability, we have ignored other important moral preferences, such as reciprocity and honesty, that eventually might be included in a more general analysis.

So, focusing on fairness as distributive justice, should this be incorporated into welfare analysis, and, if so, how? In addressing these questions, we proceed first from the most conservative conclusions implied from the foregoing analysis. The behavioral and survey evidence on fairness strongly suggests that people value it subjectively. That is, put into economic terms, people derive utility (in terms of well-being) from fairness, or equivalently experience disutility from unfairness. If this is the case, fairness has implications for allocations even under standard analysis.

Formally, we may think of a conventional welfaristic social welfare function as follows:

$$W = w\left(U^1, U^2, \ldots, U^n\right), \tag{1}$$

where U^i reflects the utility or well-being of individual i. Assume that each individual's utility depends on his/her own income, x^i, as well as the perceived overall fairness, F^i, so that

$$U^i = u^i(x^i, F^i) \quad \forall i. \tag{2}$$

Let us next assume that i's perceived fairness depends on the distribution of income in the society generally, so that

$$F^i = f^i(\mathbf{x}) \equiv f^i(x^1, \ldots, x^n) \quad \forall i. \tag{3}$$

Note that f^i can reflect issues such as accountability, efficiency and need, and it is, therefore, not at all obvious that it is symmetric with respect to the income levels. For example, i may perceive the allocation to be fairer if individual 2 has twice the income of individual 1 relative to the case in which both have the same income. Substituting (2) and (3) into (1) we obtain

$$W = w\left(u^1(x^1, f^1(\mathbf{x})), u^2(x^2, f^2(\mathbf{x})), \ldots, u^n(x^n, f^n(\mathbf{x}))\right). \tag{4}$$

We can next calculate the corresponding social marginal rate of substitution between an income change for individual i and j, as follows:

$$SMRS^{ij} = \frac{dW/dx^i}{dW/dx^j} = \frac{\partial w/\partial U^i \left(\partial u^i/\partial x^i + \sum_{k=1}^n \partial u^k/\partial F^k \, \partial f^k/\partial x^i\right)}{\partial w/\partial U^j \left(\partial u^j/\partial x^j + \sum_{k=1}^n \partial u^k/\partial F^k \, \partial f^k/\partial x^j\right)} \quad \forall i, j. \tag{5}$$

Thus, (5) reflects the social welfare increase of giving an additional dollar to individual i, relative to the welfare increase of instead giving the dollar to individual j. Clearly, this ratio depends on several components. First, it depends on the differences in the extent to which

incremental utility contributes to social welfare. For example, a frequently made assumption is that w is utilitarian. In that case, $\partial w / \partial U^i = \partial w / \partial U^j$ and (5) reduces to

$$SMRS^{ij} = \frac{\partial u^i / \partial x^i + \sum_{k=1}^{n} \partial u^k / \partial F^k \; \partial f^k / \partial x^i}{\partial u^j / \partial x^j + \sum_{k=1}^{n} \partial u^k / \partial F^k \; \partial f^k / \partial x^j} \quad \forall i, j. \tag{6}$$

If w is not utilitarian, it is typically assumed that w is strictly quasi-concave, implying that the same utility increase raises welfare more the lower the initial utility level is. Next, there is a direct effect through the changes in utility levels. As we have specified the utility function very generally, we can, of course, not say anything in general about comparisons between i's and j's marginal utilities of income, but a commonly made assumption in applied work is that the marginal utility of income is decreasing in the income level.

In addition to these welfare effects, we have the effects through the fairness perceptions. Note that we here have welfare effects for all individuals, since each individual's fairness perception includes the incomes of i and j. The size of these welfare effects, in turn, depends on the weight the fairness concern carries in each individual's utility function, $\partial u^k / \partial F^k$, as well as on how much each individual k's fairness perception is affected by a small income increase of i (or j), i.e., on $\partial f^k / \partial x^i$. Thus, even under very modest assumptions acceptable within standard welfare analysis, including utilitarianism, fairness has implications for social welfare.

An alternative approach is to assume that fairness is an intrinsic argument in the social welfare function (SWF), as follows:

$$W = w \left(U^1, U^2, \ldots, U^n, F \right). \tag{7}$$

Note that (7) constitutes a *non-welfaristic* social welfare function, following the terminology of Sen (1970, 1979), since it does not solely depend on individual utilities. That is, people might think that the society should value fairness intrinsically. In this case, it can be shown that fairness matters for allocations, even if people individually do not derive utility from fairness. That is, let us here assume that people's utilities (reflecting well-being) are not affected by any fairness concerns, so that we have

$$U^i = u^i (x^i) \quad \forall i, \tag{8}$$

whereas the fairness concerns at the social level can be written as a social fairness function

$$F = f(\mathbf{x}). \tag{9}$$

Substituting (8) and (9) into (7) we obtain

$$W = w \left(u^1 (x^1), u^2 (x^2), \ldots, u^n (x^n), f(\mathbf{x}) \right). \tag{10}$$

The social marginal rate of substitution between an income change for individual i and j can then be written as:

$$SMRS^{ij} = \frac{\partial w / \partial U^i \; \partial u^i / \partial x^i + \partial w / \partial F \partial f / \partial x^i}{\partial w / \partial U^j \; \partial u^j / \partial x^j + \partial w / \partial F \partial f / \partial x^i} \quad \forall i, j. \tag{11}$$

Note that we cannot analyze a utilitarian case here, since utilitarianism is a special case of a welfaristic SWF, and the SWF in (7) is non-welfaristic. Nevertheless, we can still analyze the symmetric case in which each individual's utility carries the same weight in the SWF, similar to the utilitarian case. Then, we have $\partial w / \partial U^i = \partial w / \partial U^j = \partial w / \partial U$, implying that

(11) simplifies to

$$SMRS^{ij} = \frac{\partial u^i/\partial x^i + \frac{\partial w/\partial F}{\partial w/\partial U}\partial f/\partial x^i}{\partial u^j/\partial x^j + \frac{\partial w/\partial F}{\partial w/\partial U}\partial f/\partial x^j} \quad \forall i, j. \tag{12}$$

Ignoring the effects through the SWF, welfare effects emerge through the marginal utility of income, as above, and through the social fairness function. The size of the fairness related welfare effect depends on how much fairness is affected by the income change, $\partial f/\partial x^i$, and on the relative weight that is given on the margin to fairness concerns relative to utility concerns in the SWF, i.e., $\frac{\partial w/\partial F}{\partial w/\partial U}$.

We can, of course, also think of the case in which fairness is an argument both of individual utility functions and of the social welfare function. In this case, we can substitute (2), (3) and (9) into (7) and obtain

$$W = w\left(u^1(x^1, f^1(\mathbf{x})), u^2(x^2, f^2(\mathbf{x})), \dots, u^n(x^n, f^n(\mathbf{x})), f(\mathbf{x})\right). \tag{13}$$

Then, we can write the social marginal rate of substitution between an income change for individual i and j as follows:

$$SMRS^{ij} = \frac{\partial w/\partial U^i \left(\partial u^i/\partial x^i + \sum_{k=1}^{n} \partial u^k/\partial F^k \, \partial f^k/\partial x^i\right) + \partial w/\partial F \partial f/\partial x^i}{\partial w/\partial U^j \left(\partial u^j/\partial x^j + \sum_{k=1}^{n} \partial u^k/\partial F^k \, \partial f^k/\partial x^j\right) + \partial w/\partial F \partial f/\partial x^j} \quad \forall i, j, \tag{14}$$

which, in the symmetric case where $\partial w/\partial U^i = \partial w/\partial U^j = \partial w/\partial U$, can be simplified as

$$SMRS^{ij} = \frac{\partial u^i/\partial x^i + \sum_{k=1}^{n} \partial u^k/\partial F^k \, \partial f^k/\partial x^i + \frac{\partial w/\partial F}{\partial w/\partial U}\partial f/\partial x^i}{\partial u^j/\partial x^j + \sum_{k=1}^{n} \partial u^k/\partial F^k \, \partial f^k/\partial x^j + \frac{\partial w/\partial F}{\partial w/\partial U}\partial f/\partial x^j} \quad \forall i, j. \tag{15}$$

Thus, in addition to the welfare effects in (12), we have here also the effects through the individual fairness concerns in each individual's utility function.

Should the government respect people's preferences in these respects? Starting with the effects through people's utility functions, we find it hard to argue that these effects should not count. Indeed, why would utility derived from perceived fairness count less than the same amount derived for any other reason?

The intrinsic fairness effect through the social objective function is perhaps less straightforward. But there are compelling reasons for basing social objective functions on the ethical values that people in the society have. By one view, this is justified as a representation of the moral preferences of an idealized agent. In fact, one can imagine a more radical departure from welfaristic social objective functions that abandons utility as an argument altogether. The social objective function might be written solely as a function of the various functions that, in turn, reflect moral preferences, including, inter alia, accountability, efficiency and need. The arguments of these moral preference functions could be income, as formulated above with the social fairness function, or utility could be reinstated by making the moral preference functions themselves functions of individual utilities.

A related, but mostly separate, discussion has arisen regarding the implications of findings about bounded rationality for social welfare. A growing literature, motivated by behavioral economics insights, points out the hazards of adopting "consumer sovereignty" as a universal guiding rule for welfare analysis. One approach is to assume that what really matters is well-being rather than choice, following, for example, Broome (1999), Ng (1999), O'Donoghue and Rabin (2006), Johansson-Stenman (2008), or, using the terminology of Kahneman et al.

(1997), Kahneman and Thaler (2006), to assume that what matters is *experienced* utility rather than *decision* utility.

In conventional welfare analysis, the ultimate goal is often expressed as the maximization of social welfare or well-being, and if we believe that respecting people's preferences, as revealed by their choices, is an effective way of obtaining this goal, then it follows that it would indeed be a good idea for policy makers to respect the principle of consumer sovereignty. This is contingent, however, on the assumption that people do know, and act in accordance with, what is best (in terms of their well-being) for themselves. But suppose, for example, that people have self-control problems, which imply time inconsistency, such that they fail to act in accordance with their own interests. Then the link from consumer sovereignty to welfare maximization does not follow.

As a result, policy measures based on different kinds of paternalism has been proposed. For example, Gruber and Köszegi (2001) argue in favor of cigarette taxation, not in order to internalize *externalities* (which they argue are rather small, anyway), but rather in order to internalize what they denote *internalities*, i.e., in order to help them act consistent with their own ultimate will and interest. Similarly, O'Donoghue and Rabin (2006) argue in favor of "fat taxes," and other "sin taxes." Camerer (2003), Thaler and Sunstein (2008) provide good overviews of such arguments, whereas Sugden (2004), Bernheim and Rangel (2007), Bernheim and Rangel (2009) provide alternative choice-based approaches when people make mistakes.

Concerns for moral preferences and bounded rationality problems sometimes intersect. Bazerman refers to the phenomenon of "bounded ethicality," i.e., the idea that even good people often violate moral norms unawares and as the result of ordinary and predictable psychological processes. For example, Gino and Bazerman (2009) find that people are more likely to accept unethical behavior when allowed to adapt to it more slowly. One can easily imagine serious ramifications of moral preferences under bounded rationality, e.g., if people express political preferences in favor of institutional arrangements that are perceived to be fair but that imply large inefficiency losses, which people systematically tend to underestimate. In fact, there is evidence of this, e.g., Caplan (2002) finds that people, on average, tend to underestimate the benefits from trade and globalization. The question then arises whether the social objective should be based on the values people express or rather on their "informed values."

Thus, considerations of fairness and bounded rationality raise interesting, important and difficult challenges for welfare analysis. Nevertheless, these facts can be seen as a basis for improving and making more relevant this analysis. In the partial equilibrium context of cost-benefit analysis, that is precisely the conclusion at which Bazerman and Greene (2010) arrive, and we believe the same will be found to be true for general equilibrium welfare analysis.

6 Conclusion

This paper presents empirical evidence that people care about fairness and that it is possible to analyze fairness concerns systematically. Although views of fairness often differ, there is much less heterogeneity regarding impartial fairness concerns, i.e., those expressed by spectators.

We conclude that fairness preferences have important implications for the field of environmental economics through the analysis of (1) individual behavior, (2) the behavior of other decision makers such as policy makers and international negotiators, and also through (3) the

overall social objective function. We believe that the benefits in terms of greater empirical predictive power, as well as greater theoretical and policy relevance, of incorporating fairness concerns in the analysis often dominate the associated costs in terms of greater complexity. We also believe that environmental economics would have greater influence on environmental policy if fairness considerations were taken more seriously.

Thus, we think that theoretical and empirical research of fairness in environmental economics is fertile ground for future work. One particularly important and potentially impactful avenue of investigation is empirical analysis of fairness perceptions about alternative policy instruments. Many environmental policy instruments, such as tradable permit systems, can be quite complex and cognitively demanding for laypersons. This contextual complexity, often combined with the previously discussed self-serving biases, present challenges to reaching a consensus on environmental policy that is the result of an informed and thoughtful process. Thus, future work might build on experimental studies, such as those of Frohlich and Oppenheimer (1992), which engage people in discussions aimed at learning and moral reflection and which have sometimes been shown to generate consensus about policies with important distributive consequences.

References

Alesina A, Angeletos G-M (2005) Fairness and redistribution. Am Econ Rev 95(4):960–980
Alpizar F, Carlsson F, Johansson-Stenman O (2008) Anonymity, reciprocity and conformity: evidence from voluntary contributions to a natural park in Costa Rica. J Public Econ 92:986–997
Andreoni J (1990) Impure altruism and donations to public goods: a theory of warm-glow giving. Econ J 100:464–477
Andreoni J, Miller J (2002) Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica 70(2):737–753
Asheim GB, Froyn C, Bretteville HJ, Menz FC (2006) Regional versus global cooperation for climate control. J Environ Econ Manage 51:93–109
Babcock L, Loewenstein G (1997) Explaining bargaining impasse: the role of self-serving biases. J Econ Perspect 11(1):109–126
Barrett S (1994) Self-enforcing international environmental agreements. Oxf Econ Pap 46:878–894
Barrett S (1997) Heterogeneous international environmental agreements. In: Carraro C (ed) International environmental agreemens: strategic policy issues. Edward Elgar, Cheltenham, pp 9–25
Barrett S (2005) The theory of international environmental agreements. In: Mäler K-G, Vincent JR (eds) Handbook of environmental economics, vol. 3. Elsevier, Amsterdam, p 28
Bazerman M, Greene J (2010) In favor of clear thinking: incorporating moral rules into a wise cost-benefit analysis. Perspect Psycol Sci 5(2):209–212
Bernheim D, Rangel A (2007) Toward choice-theoretic foundations for behavioral welfare economics. Am Econ Rev 97(2):464–470
Bernheim D, Rangel A (2009) Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. Q J Econ 124(1):51–104
Bolton G, Ockenfels A (2000) A theory of equity, reciprocity, and competition. Am Econ Rev 90(1):166–193
Bossert W, Fleurbaey M (1996) Redistribution and compensation. Soc Choice Welf 13(3):343–355
Brekke KA, Johansson-Stenman O (2008) The behavioural economics of climate change. Oxf Rev of Econ Policy 24(2):280–297
Brennan G, Hamlin A (1998) Expressive voting and electoral equilibrium. Public Choice 95(1):149–175
Broome J (1999) Ethics out of economics. Cambridge University Press, Cambridge
Buckley E, Croson R (2006) Income and wealth heterogeneity in the voluntary provision of linear public goods. J Public Econ 90:935–955

Cai B, Cameron TA, Gerdes GR (2008) Distributional preferences and the incidence of costs and benefits in environmental policy. In: Paper presented at the annual European conference of environmental and resource economics in Gothenburg 2008

Camerer C (2003) Behavioral game theory: experiments in strategic interaction. Princeton University Press, Princeton

Caplan B (2002) Systematically biased beliefs about economics: robust evidence of judgmental anomalies from the survey of Americans and economists on the economy. Econ J 112:1–26

Cappelen A, Tungodden B (2003) Reward and responsibility: how should we be affected when others change their effort. Polit Philos Econ 2(2):191–211

Cappelen AW, Hole AD, Sørensen EO, Tungodden B (2007) The pluralism of fairness ideals: an experimental approach. Am Econ Rev 97(3):818–827

Cappelen AW, Tungodden B (2009) Rewarding effort. Econ Theory 39(3):425–441

Carlsson F, Johansson-Stenman O (2009) Voting motives, group identity, and social norms. Working paper in Economics no. 366, Department of Economics, University of Gothenburg

Carraro C, Siniscalco D (1993) Strategies for the international protection of the environment. J Public Econ 52:309–328

Carraro C, Siniscalco D (1998) International environmental agreements: incentives and political economy. Eur Econ Rev 42:561–572

Charness G, Rabin M (2002) Understanding social preferences with simple tests. Q J Econ 117(3):817–869

Charness G, Frechette GR, Qin C-Z (2007) Endogenous transfers in the prisoner's dilemma game. Games Econ Behav 60(2):287–306

Chaudhuri A (2009) Experiments in economics. Routledge, London and New York

Cherry TL, Kroll S, Shogren JF (2005) The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. J Econ Behav Organ 57(3):357–365

Cox JC, Friedman D, Gjerstad S (2007) A tractable model of reciprocity and fairness. Games Econ Behav 59:17–45

Croson R, Fatas E, Neugebauer T (2005) Reciprocity, matching and conditional cooperation in two public goods games. Econ Lett 87(1):95–101

Croson R (2007) Theories of commitment, altruism and reciprocity: evidence from linear public goods games. Econ Inq 45(2):199–216

Crumpler H, Grossman PJ (2008) An experimental test of warm glow giving. J Public Econ 92(5–6):1011–1021

Dunning D, Meyerowitz JA, Holzberg AD (1989) Ambiguity and self-evaluation: the role of idiosyncratic trait definitions in self-serving assessments of ability. J Pers Soc Psychol 57(6):1082–1090

Eckel CC, Grossman PJ (1996) Altruism in anonymous dictator games. Games Econ Behav 16:181–191

Engelmann D, Strobel M (2004) Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. Am Econ Rev 94(4):857–869

Fehr E, Schmidt KM (1999) A theory of fairness, competition and cooperation. Q J Econ 114(3):817–868

Fehr E, Naef M, Schmidt KM (2006) Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: comment. Am Econ Rev 96(5):1912–1917

Festinger L (1957) A theory of cognitive dissonance. Stanford University Press, Palo Alto

Fischbacher U, Gachter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. Econ Lett 71(3):397–404

Fong C (2001) Social preferences, self-interest, and the demand for redistribution. J Public Econ 82(2):225–246

Frohlich N, Oppenheimer JA (1992) Choosing Justice. An Experimental Approach to Ethical Theory. University of California Press, Berkeley

Gächter S, Riedl A (2005) Moral property rights in bargaining with infeasible claims. Manage Sci 51(2):249–263

Gächter S, Riedl A (2006) Dividing justly in bargaining problems. Soc Choice Welf 27:571–594

George, H (1879) Progress and poverty. Random House: San Francisco, CA

Gino F, Bazerman M (2009) When misconduct goes unnoticed: the acceptability of gradual erosion in others' unethical behavior. J Exp Soc Psychol 45(4):708–719

Gruber J, Köszegi B (2001) Is addiction 'rational'? Theory and evidence. Q J Econ 116(4):1261–1303

Güth W (1994) Distributive justice: a behavioral theory and empirical evidence. In: Güth W, Brandstätter H (eds) Essays on economic psychology. Springer, Berlin, pp 153–175

Güth W, Schmittberger R, Schwarze B (1982) An experimental analysis of ultimatum bargaining. J Econ Behav Organ 3(4):367–388

Herne K, Mård T (2006) Three versions of impartiality: an experimental investigation. Department of Political Science: University of Turku, Finland

Johansson-Stenman O (2000) On the value of life in rich and poor countries and distributional weights beyond utilitarianism. Environ Resour Econ 17:299–310

Johansson-Stenman O (2008) Mad Cows, terrorism and junk food: should public policy reflect subjective or objective risks?J Health Econ 27(2):234–248

Kahneman D, Wakker P, Sarin R (1997) Back to bentham? Explorations of experienced utility. J Health Econ 112(2):375–405

Kahneman D, Thaler R (2006) Anomalies: utility maximisation and experienced utility. J Econ Perspect 20:221–234

Konow J (1996) A positive theory of economic fairness. J Econ Behav Organ 31(1):13–35

Konow J (2000) Fair shares: accountability and cognitive dissonance in allocation decisions. Am Econ Rev 90(4):1072–1092

Konow J (2001) Fair and square: the four sides of distributive justice. J Econ Behav Organ 46(2):137–164

Konow J (2003) Which is the fairest one of all? A positive analysis of justice theories. J Econ Lit 41(4):1186–1237

Konow J (2009) Is fairness in the eye of the beholder? An impartial spectator analysis of justice. Soc Choice Welfare 33(1):101–127

Konow J (2010) Mixed feelings: theories of and evidence on giving. J Public Econ 94(3–4):279–297

Konow J, Saijo T, Akai K (2009) Morals versus mores: experimental evidence on equity and equality. Manuscript

Kverndokk S (1995) Tradeable CO2 emission permits: initial distribution as a justice problem. Environ Values 4(2):129–148

Lange A, Vogt C (2003) Cooperation in international environmental negotiations due to a preference for equity. J Public Econ 87(9–10):2049–2067

Lange A (2006) The impact of equity-preferences on the stability of international environmental agreements. Environ Resour Econ 34:247–267

Lange A, Loschel A, Vogt C, Ziegler A (2010) On the self-serving use of equity in international climate negotiations. Eur Econ Rev 54:359–375

Luhan WJ, Kocher M, Sutter M (2009) Group polarization in the team dictator game reconsidered. Exp Econ 12(1):26–41

Malthus T (1798) An essay on the principle of population. J.M. Dent & Sons, London

McConnell KE (1997) Does altruism undermine existence value. J Environ Econ Manage 32:22–37

Ng Y-K (1999) Utility, informed preference, or happiness: following Harsanyi's argument to its logical conclusion. Soc Choice Welfare 16:197–216

Nozick R (1974) Anarchy, state, and Utopia. Basic Books Inc, New York

O'Donoghue T, Rabin M (2006) Optimal sin taxes. J Public Econ 90:1825–1849

Oberholzer-Gee F, Bohnet I, Frey B (1997) Fairness and competence in democratic decisions. Public Choice 91(1):89–105

Oxoby R (2006) Paretian dictators: constraining choice in a voluntary contribution game. Department of Economics, University of Calgary

Roemer JE (1998) Equality of opportunity. Harvard University Press, Cambridge, Mass

Ruffle B, Sosis R (2006) Cooperation and the in-group-out-group bias: a field test on Israeli kibbutz members and city residents. J Econ Behav Organ 60(2):147–163

Schkade DA, Payne JW (1994) How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation. J Environ Econ Manage 26:88–109

Selten R (1978) The equity principle in economic behavior. In: Gottinger H, Leinfellner W (eds) Decision theory and social ethics, issues in social choice. Reifel Publishing, Dordrecht, pp 289–301

Sen AK (1970) Collective Choice and Social Welfare, North-Holland, Amsterdam

Sen AK (1979) Personal utilities and public judgements: or what's wrong with welfare economics. Econ J 89:537–558

Sobel RS, Wagner GA (2004) Expressive voting and government redistribution: Testing Tullock's Charity of the uncharitable. Public Choice 119:143–159

Stahl D, Haruvy E (2009) Testing theories of behavior for extensive-form two-player two-stage games. Exp Econ 12(2):242–251

Sugden R (2004) The opportunity criterion: consumer sovereignty without the assumption of coherent preferences. Am Econ Rev 94(4):1014–1033

Thaler RH, Sunstein CR (2008) Nudge: improving decisions about health, wealth, and happiness. Yale University Press, New Haven

Tyran J-R (2004) Voting when money and morals conflict: an experimental test of expressive voting. J Public Econ 88(7):1645–1664

World Bank Development Indicators (2005) http://devdata.worldbank.org/wdi2005/Cover.htm